Arthritis Care & Research

Aims and Scope

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

Volume 74 No. 8 August 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 • AUGUST 2022 • NO. 8

Rheumatology Practice

Review: Telemedicine in Adult Rheumatology: In Practice and in Training
Megan M. Lockwood, Rachel S. Wallwork, Kaitlin Lima, Anisha B. Dua, Philip Seo, and Marcy B. Bolster 1227

Psoriatic Arthritis

Preferences for Biologic Treatment in Patients With Psoriatic Arthritis: A Discrete Choice Experiment
Daniel Sumpton, Ayano Kelly, Jonathan C. Craig, Geraldine Hassett, Barry Kane, Michael Oliffe, Allison Tong, and Martin Howell 1234

Incidence of and Risk Factors for Heart Failure in Patients With Psoriatic Disease: A Cohort Study
Sahil Koppikar, Keith Colaco, Paula Harvey, Shadi Akhtari, Vinod Chandran, Dafna D. Gladman, Richard Cook, and Lihi Eder 1244

Comparative Efficacy of Different Triage Methods for Psoriatic Arthritis: Results From a Prospective Study in a Rapid Access Clinic
Samantha Sarabia, Chandra Farrer, Jensen Yeung, Dana Jerome, Richard J. Cook, and Lihi Eder 1254

Systemic Lupus Erythematosus

Brief Report: Predictors of Initial Hydroxychloroquine Receipt Among Medicaid Beneficiaries With Incident Systemic Lupus Erythematosus
Katherine P. Pryor, Chang Xu, Jamie E. Collins, Karen H. Costenbader, and Candace H. Feldman 1263

Polychondritis

Physician Global Assessment as a Disease Activity Measure for Relapsing Polychondritis
Emily Rose, Marcela A. Ferrada, Kaitlin A. Quinn, Wendy Goodspeed, Laurent Arnaud, Aman Sharma, Hajime Yoshifuji, Jeff Kim, Clint Allen, Arlene Sirajuddin, Marcus Chen, and Peter C. Grayson 1269

Gout

Relationship Between Changing Body Mass Index and Serum Uric Acid Alteration Among Clinically Apparently Healthy Korean Men
Jiwon Hwang, Mi Yeon Lee, Joong Kyong Ahn, and Hoon-Suk Cha 1277

Spondyloarthritis

Relationship Between Body Mass Index, Disease Activity, and Exercise in Ankylosing Spondylitis
Jean W. Liew, Milena A. Gianfrancesco, Susan R. Heckbert, and Lianne S. Gensler 1287

Vasculitis

Brief Report: Increased Mortality for Individuals With Giant Cell Arteritis: A Population-Based Study
Lillian Barra, Janet E. Pope, Priscila Pequeno, Jodi M. Gatley, and Jessica Widdifield 1294

Scleroderma

Barriers and Facilitators to Physical Activity for People With Scleroderma: A Scleroderma Patient-Centered Intervention Network Cohort Study
Sami Harb, Sandra Peláez, Marie-Eve Carrier, Linda Kwakkenbos, Susan J. Bartlett, Marie Hudson, Luc Mouthon, Maureen Sauvé, Joep Welling, Ian Shrier and Brett D. Thombs, for the SPIN Physical Activity Enhancement Patient Advisory Team and SPIN Investigators 1300

Pediatrics

Comprehensive Assessment of Quality of Life, Functioning, and Mental Health in Children With Juvenile Idiopathic Arthritis and Noninfectious Uveitis
Joseph McDonald, Amy Cassedy, Mekibib Altaye, Jennifer Andringa, Ashley M. Cooper, Carolyn Drews-Botsch, George Engelhard Jr., Theresa Hennard, Gary N. Holland, Kirsten Jenkins, Scott R. Lambert, Jessi Lipscomb, Courtney McCracken, Deborah K. McCurdy, Najima Mwase, Sampath Prahalad, Jessica Shantha, Erin Stahl, Virginia Miraldi Utz, A. Adrienne Walker, Steven Yeh, and Sheila T. Angeles-Han 1311

Rheumatoid Arthritis

Brief Report: Trends in Permanent Work Disability Associated With Rheumatoid Arthritis in the United States, 1999–2015
Michael M. Ward 1321

Decision Needs and Preferred Strategies for Shared Decision-Making in Rheumatoid Arthritis: Perspectives of Canadian Urban Indigenous Women
Valerie Umaefulam, Terri-Lynn Fox, and Cheryl Barnabe 1325

Differences in the Association Between Oral Glucocorticoids and Risk of Preterm Birth by Data Source: Reconciling the Results
Kristin Palmsten, Gretchen Bandoli, Gabriela Vazquez-Benitez, and Christina D. Chambers 1332

Osteoarthritis

Development of a Medicare Claims–Based Model to Predict Persistent High-Dose Opioid Use After Total Knee Replacement
Chandrasekar Gopalakrishnan, Rishi J. Desai, Jessica M. Franklin, Yinzhu Jin, Joyce Lii, Daniel H. Solomon, Jeffrey N. Katz, Yvonne C. Lee, Patricia D. Franklin, and Seoyoung C. Kim 1342

Societal Cost of Opioid Use in Symptomatic Knee Osteoarthritis Patients in the United States
Jamie L. Huizinga, Elizabeth E. Stanley, James K. Sullivan, Shuang Song, David J. Hunter, A. David Paltiel, Tuhina Neogi, Robert R. Edwards, Jeffrey N. Katz, and Elena Losina 1349

Application of Heterogeneity of Treatment-Effects Methods: Exploratory Analyses of a Trial of Exercise-Based Interventions for Knee Osteoarthritis
Cynthia J. Coffman, Liubov Arbeeva, Todd A. Schwartz, Leigh F. Callahan, Yvonne M. Golightly, Adam P. Goode, Kim M. Huffman, and Kelli D. Allen 1359

Brief Report: Identification of Radiographic Foot Osteoarthritis: Sensitivity of Views and Features Using the La Trobe Radiographic Atlas
Hylton B. Menz, Shannon E. Munteanu, Michelle Marshall, Martin J. Thomas, Trishna Rathod-Mistry, George M. Peat, and Edward Roddy 1369

Relationship Between Patient-Reported Readiness for Total Knee Arthroplasty and Likelihood of a Good Outcome at One-Year Follow-Up
Gillian A. Hawker, Barbara L. Conner-Spady, Eric Bohm, Michael J. Dunbar, C. Allyson Jones, Bheeshma Ravi, Tom Noseworthy, Linda J. Woodhouse, Peter Faris, Donald Dick, James Powell, Paulose Paul, and Deborah A. Marshall, on behalf of the BEST-Knee Study Team 1374

Association Between Baseline Meniscal Symptoms and Outcomes of Operative and Nonoperative Treatment of Meniscal Tear in Patients With Osteoarthritis
Lindsey A. MacFarlane, Heidi Yang, Jamie E. Collins, Robert H. Brophy, Brian J. Cole, Kurt P. Spindler, Ali Guermazi, Morgan H. Jones, Lisa A. Mandl, Scott Martin, Robert G. Marx, Bruce A. Levy, Michael Stuart, Clare Safran-Norton, John Wright, Rick W. Wright, Elena Losina, and Jeffrey N. Katz 1384

Presence of Magnetic Resonance Imaging–Defined Inflammation Particularly in Overweight and Obese Women Increases Risk of Radiographic Knee Osteoarthritis: The POMA Study
Frank W. Roemer, Ali Guermazi, Michael J. Hannon, Tomoko Fujii, Patrick Omoumi, David J. Hunter, Felix Eckstein, and C. Kent Kwoh 1391

Cover image: : The figure on the cover (from Sumpton et al, page 1234) shows an excerpt from an example of a choice set presented to patients as part of a discrete choice experiment focused on biologic medications for use in psoriatic arthritis.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1227–1233
DOI 10.1002/acr.24569

American College of Rheumatology
Empowering Rheumatology Professionals

REVIEW

Telemedicine in Adult Rheumatology: In Practice and in Training

Megan M. Lockwood,[1] Rachel S. Wallwork,[2] Kaitlin Lima,[3] Anisha B. Dua,[3] Philip Seo,[2] and Marcy B. Bolster[1]

Many rheumatology providers, including fellows-in-training, responded to the immediate need for maintaining patient access to care via telerheumatology during the COVID-19 pandemic. The rapidity of this transition did not permit an intentional approach to integrating fellow education and training into virtual patient care. Virtual patient care has since become an integrated, and perhaps, an embedded part of rheumatology practice that will likely endure beyond the COVID-19 pandemic. Thus, the development of best practices in telerheumatology, including those for fellow education and training as these new entrants prepare to enter our workforce, will benefit the entire specialty. In this work, we seek to describe current models for training learners in virtual patient care, characterize existing barriers to virtual care models, and offer strategies to integrate telerheumatology into curriculum development and training.

Introduction

Telemedicine has been defined as "the use of medical information that is exchanged from one site to another through electronic communication to improve a patient's health" (1). There are several interactions to consider within telemedicine, including clinician-to-clinician, clinician-to-patient, and patient-to-mobile health technology, each of which can provide synchronous or asynchronous care. Traditionally, rheumatologists have used telemedicine to provide care for patients with limited access to subspecialists, a care gap accentuated by the geographic maldistribution of rheumatologists in the US (2), particularly in rural and underserved communities. Due to the COVID-19 pandemic, patients' access to health care changed substantially. Rheumatologists have been forced to rapidly adopt and adapt their practices to telemedicine. Limited clinic capacity, high-risk patient populations, and redeployment of rheumatologists to other responsibilities have dramatically reduced access to care even in urban areas served by large academic centers. Providers and hospital systems have had to rely heavily on telemedicine to sustain patient care while protecting patients, providers, and staff from viral transmission.

Many rheumatology providers, including fellows-in-training, responded to the immediate need for maintaining patient access to care via telerheumatology during the COVID-19 pandemic. The rapid transition did not permit an intentional approach to integrating fellow training into virtual clinical practice. As the pandemic continues, virtual visits continue to be an ongoing part of rheumatology patient care, and the incorporation of telerheumatology into clinical practice will likely endure beyond the COVID-19 pandemic. We recognize the importance of integrating telemedicine into the training space to enhance the preparation of fellows-in-training entering clinical practice. The development of best practices in telerheumatology, specifically in training, will benefit the entire specialty of rheumatology. The thoughtful and intentional integration of education, assessment, development, and determination of competency in the virtual rheumatology space will enhance care, rather than merely replace in-person visits. Fellows-in-training, as new entrants into practice, will bring knowledge and skills in virtual care to those already in practice. The goals of the present review are to provide background for current barriers in telerheumatology and in existing virtual clinical models within rheumatology, to examine current models for training learners in virtual care delivery, and to offer strategies to incorporate telerheumatology into training and curriculum development.

[1]Megan M. Lockwood, MD, Marcy B. Bolster, MD: Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts; [2]Rachel S. Wallwork, MD, Philip Seo, MD, MHS: Johns Hopkins University School of Medicine, Baltimore, Maryland; [3]Kaitlin Lima, MD, Anisha B. Dua, MD, MPH: Northwestern University Feinberg School of Medicine, Chicago, Illinois.

Drs. Lockwood and Wallwork contributed equally to this work.

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

Address correspondence to Marcy B. Bolster, MD, Division of Rheumatology, Allergy, and Immunology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, 55 Fruit Street, Bulfinch 165, Boston, MA 02114. Email: mbolster@mgh.harvard.edu.

Submitted for publication October 8, 2020; accepted in revised form February 2, 2021.

Barriers to Telerheumatology

Presently, research investigating the effectiveness of virtual visits and identification of patients best suited for virtual visits is lacking, yet telemedicine has rapidly progressed, creating a significant gap between research and implementation (3,4). The largest systematic review of telerheumatology to date identified only 20 published studies (3), with variable study design and outcomes, leaving one unable to draw clear conclusions. We thus need more rigorous studies comparing the diagnostic accuracy and health outcomes between care delivered virtually versus in-person (5).

Telerheumatology is unlikely to optimize care for all patients. The quality of care provided by telerheumatology is likely to depend on the specific disease and level of disease activity. Studies assessing what rheumatic diseases at which stage can be safely managed virtually are critical. A previous randomized controlled trial demonstrated that for patients with rheumatoid arthritis whose disease was in remission or who had low disease activity, telerheumatology was noninferior compared to in-person visits for maintaining remission (6). However, similar studies in patients with systemic autoimmune diseases or vasculitis are lacking (3).

A survey of rheumatologists found that 19% of patients were unsuitable for telemedicine due to diagnostic uncertainty or high disease complexity (7). Patients must be appropriately triaged to virtual or in-person visits to avoid delays in diagnostic evaluation and treatment initiation (7). Alternatively, telemedicine visits may serve as an in-depth triage to determine the need for an urgent in-person evaluation. Fellows-in-training often become outpatient providers for patients with high-acuity, complex medical problems from a hospital inpatient rheumatology consult service; these patients in particular may have evolving disease processes requiring in-person evaluation for close monitoring and multidisciplinary management, and it may be inappropriate for these patients to receive care via telerheumatology. Additional robust studies are essential to guide appropriate patient selection for telerheumatology and to assist with establishing need and intervals for in-person visits.

Prior to the COVID-19 pandemic, one of the greatest barriers to telerheumatology was the lack of reimbursement. The US Centers for Medicare & Medicaid Services broadened its coverage of telemedicine during the pandemic and increased reimbursement rates to a similar level as those of in-person visits (8). Significant uncertainty exists in regard to telemedicine reimbursement rates after the COVID-19 pandemic, which will certainly impact the potential for continued implementation and further expansion of virtual care. The projected incorporation of telemedicine in patient care may influence the emphasis placed on developing and implementing telerheumatology in fellowship training.

Technology itself can create barriers to care. Currently, telemedicine visits are conducted using audiovisual or telephone technology, each of which offers unique limitations to the medical evaluation. A patient's internet bandwidth may be insufficient for the telehealth platform, leading to poor voice transmission or blurry video quality. Moreover, telephone encounters, limited to a patient's report of symptoms and lacking a physical examination, inherently restrict diagnostic abilities. Even with clear video quality, the examination remains limited. Only some patients have home equipment to collect vital signs, and even then, providers remain unable to confirm the accuracy of data collected by patients.

Telemedicine is not immune to health care access disparities and, in fact, may compound existing inequities. A study on cardiology patients revealed that the rapid adoption of telemedicine during the COVID-19 pandemic has excluded a greater proportion of female, non–English-speaking, older, and economically disadvantaged patients as these patients are less likely to complete a virtual visit (9). If a patient overcomes such access inequities to attend a video visit, the setting in which patients situate themselves can impact the evaluation. Patients may, for example, opt to conduct a video visit in a setting such as a car to gain privacy, which may significantly limit the ability to perform a physical examination. Additionally, patients may find it difficult to secure a private space to discuss personal information, particularly in the setting of sheltering-in-place, as many family members may share the same household.

Existing Virtual Clinical Models Within Rheumatology

Telerheumatology has expanded significantly during the COVID-19 pandemic, largely without rheumatology prototypes, for patients who previously had adequate access to care. There are, however, existing and well-established models for telerheumatology practice designed for remote and underserved communities. One of the largest and well-developed telerheumatology programs is based at the Alaska Native Medical Center. In this program, rheumatologists connect virtually with patients at 200 access locations staffed by community health aides who are trained to perform specific medical tasks (10). Notably, telerheumatology is generally exclusively used for follow-up visits as community focus groups and providers believed initial in-person visits were essential to allow providers to perform a physical examination and to establish the provider–patient relationship. It is important to recognize that this model differs from the current pandemic telemedicine model in which patients connect directly from home. Providers in the pandemic landscape must work with the patient's available technology and lack the assistance of a trained presenter. There are no data to compare the direct-to-patient telemedicine model to a satellite site model with a trained presenter.

Similar to the Alaska Native Medical Center program, rheumatologists at Dartmouth-Hitchcock Medical Center use telerheumatology to reach patients in remote parts of New Hampshire and Vermont. Patients travel to satellite clinics where a presenter, either a nurse or medical assistant, facilitates the visit (7). Clinicians connect to the visit from their offices with video software

that meets compliance standards set by the US Health Insurance Portability and Accountability Act (HIPAA) of 1996. Kulcsar and colleagues demonstrated that patients who drove to the satellite telerheumatology clinic saved an average 200 driving miles and $66.90 in transportation costs per remote rheumatology clinic visit (7). These findings are particularly notable as the study was conducted in the Northeastern US, which has the greatest density of rheumatologists of any region in the country (2,11). The time and financial savings to patients are potentially even greater in underserved regions (12). The cost savings for the health care system is less clear. Economic evaluation requires analyzing the complex interplay between direct cost savings, such as reduction in travel, with overall costs, such as increased administrative overhead for scheduling and telemedicine triage. Existing studies suggest that telehealth does not always confer overall cost savings in the short-to-medium term; however, data specific to the COVID-19 pandemic telehealth model are lacking (13).

During the pandemic, patients and providers are experiencing firsthand the benefits of virtual care, including travel cost and time savings, high patient satisfaction, and greater access to specialists—particularly for those patients living in rural areas (3,7,12,14,15). Telemedicine will likely continue to play a significant role in patient care after the COVID-19 pandemic, and it is therefore essential to train rheumatologists to use telemedicine appropriately, effectively, and safely. Unfortunately, currently there is a paucity of guidance and educational resources to train rheumatologists on the best practices or performance of physical examinations in the virtual care setting.

Models for Training Learners in Telehealth

A telerheumatology curriculum does not exist for fellows-in-training and has yet to be introduced into the American College of Rheumatology Core Curriculum (16). Prior to the COVID-19 pandemic, of the 24 rheumatology fellows-in-training surveyed within the Carolinas Fellows Collaborative (which includes Duke University, University of North Carolina, Wake Forest School of Medicine, and Medical University of South Carolina) and the Massachusetts General Hospital fellowship training programs in 2019, none had been exposed to telemedicine training in medical school, internal medicine residency training, or in their current fellowship training programs (unpublished observations).

Despite the lack of a published telemedicine curriculum in rheumatology, other specialties provide models from which we may learn. Neurology, a specialty with strong reliance on the physical examination, has published a telemedicine curriculum. University of California, San Francisco piloted the first formalized, outpatient teleneurology curriculum for third-year and fourth-year neurology residents in 2016 (17). The curriculum involved formal didactics and participation in synchronous provider–patient video visits, as well as asynchronous provider-to-provider electronic consultations and a pre-assessment and post-assessment. The rotation began with residents observing experienced attending physicians performing video patient encounters, allowing residents to become familiar with the technology, visit format, and virtual examination. After residents conducted their own virtual visit, they privately discussed the findings with the attending physician. The attending physician and resident then re-entered the virtual visit, the attending physician would delineate nuances, confirm salient portions of the history taking and physical examination, and generate a final plan. Although a small study, the pre-assessment and post-assessment revealed that all residents demonstrated significant improvement in telemedicine knowledge, comfort with the virtual examination, and expressed a more positive view of telemedicine following the rotation. Another study demonstrated that as vascular neurology fellows-in-training gained "telestroke" consultation experience, there was a decrease in page-to-thrombolytic time (18).

Incorporating trainee learning into virtual care delivery is logistically feasible and educationally beneficial. In one dermatology model, the resident and attending physician conduct the visit together, with the resident leading the encounter, followed by jointly discussing their findings with the patient (19). This practice provides direct observation of the trainee, which is a powerful training tool (20,21).

Many training programs rely simply on exposure to improve virtual care skills; however, a formalized curriculum provides a basis for acquisition of the knowledge, skills, and attitudes that are requisite in achieving competency in this form of clinical care delivery. With the implementation of a formal curriculum arises the need for performance assessment. Teleneurology has developed a set of competencies and milestones (22), and these will similarly be essential for other specialties. The Association of American Medical Colleges (AAMC) has developed a set of telehealth competencies that spans the professional development spectrum from medical student to faculty member (23). The adaptation of best practices from other specialties as well as the development of competencies should drive curriculum design and learning in rheumatology training programs (Table 1).

Strategies to Integrate Telerheumatology Into Training

There are several facets specific to fellowship training that should be considered in order to optimize telerheumatology care delivery. We must recognize the learner's experience when implementing a telerheumatology curriculum. Trainees must have foundational knowledge and skills in obtaining the appropriate rheumatic disease history and performing the physical examination before applying these skills to a virtual encounter. It is difficult for learners to appreciate important, nuanced physical examination findings entirely within the virtual learning space, such as subtle synovitis or palpable purpura. The comprehensive physical examination often includes subtleties that may be best

appreciated only after significant experience and may be unachievable through the virtual examination. It is critical that fellows-in-training continue to have the opportunity to develop competence and proficiency in the hands-on evaluation of patients. We cannot expect our trainees to appreciate nuances within a virtual visit before being able to recognize these subtleties in the traditional bedside evaluation. In fact, the American Academy of Neurology notes that comprehensive bedside neurology training is an essential component to safe teleneurology implementation (24). The rapidity of the transition to telerheumatology, with the onslaught of the COVID-19 pandemic, did not permit assessment of fellow readiness to transition to virtual care. We now have the unique opportunity to be more intentional in our approach to training fellows in the performance of history taking and examination skills, professionalism in the virtual space, communication with primary care and specialty providers, "webside manner" (the virtual analog of bedside manner), and demonstration of empathy via virtual formats. Additionally, virtual technology offers an additional tool for innovative approaches to feedback and assessment, such as with the use of direct observation, which has not previously been as readily available and facile in the in-person clinical setting.

Curriculum Implementation

The core competencies for all educational programs, as identified by the Accreditation Council for Graduate Medical

Table 1. Overview of telehealth competencies and application to trainees and faculty in rheumatology, adapted from the Association of American Medical Colleges (AAMC) New and Emerging Areas in Medicine Series: Telehealth Competencies

	Examples of competency milestones by learner status		
Domain*	Medical school and residency	Subspecialty fellowship + all prior competencies	Faculty + all prior competencies
Patient safety and appropriate use of telehealth	Acknowledges the limitations of and limited evidence base for use of virtual visits. Defines their own implicit and explicit biases and the implications of these when delivering virtual services.	Informs patients of the limitations of virtual visits and obtains informed consent. Assesses barriers for the patient and within the clinical setting to optimize incorporation of virtual visits into safe delivery of care.	Role models the safe delivery of virtual care, assesses areas for improvement, and appropriately escalates care when patient safety is at risk. Appropriately triages patients for virtual clinic visits versus in-person visits to optimize safe and effective care delivery.
Data collection and assessment via telehealth	Obtains relevant rheumatic history from patient and/or patient's support system and performs appropriate modified physical examination.	Demonstrates proficiency in performance and interpretation of modified physical examination. Integrates information obtained from history and physical examination into appropriate differential diagnosis and treatment plan for acute and chronic rheumatic diseases.	Role models and teaches appropriate history taking, modified physical examination, patient assessment, and management of plan development.
Communication via telehealth	Assesses patient environment including their support network such as family or friends who may participate in the virtual encounter.	Cultivates a professional and compassionate webside manner including camera at eye level, attention to eye contact, tone, and nonverbal cues. Appropriately engages patient's support networks in the virtual visit, including participation in the physical examination.	Role models empathetic webside manner and incorporation of patient support networks into the virtual visit. Leads team by coordinating preclinic huddle to strategize best practices for precepting patient encounters with learners. Provides timely and effective feedback to trainees.
Ethical practices and legal requirements for telehealth	Describes local and institutional policies on safe and secure virtual care delivery. Describes ethical challenges unique to virtual care delivery.	Promotes the ethical delivery of virtual services and actively seeks to expand access for patients.	Role models obtaining informed consent for a virtual visit, including maintenance and protection of patient privacy. Identifies and provides solutions for potential ethical breaches at a systems level.
Technology for telehealth	Identifies technology necessary to conduct synchronous and asynchronous virtual encounters. Describes strategies to adapt virtual care delivery based on patient's access and comfort level.	Distinguishes spectrum of technology options for virtual encounters and adapts practice to patient comfort. Troubleshoots technology failures including those occurring with patient-owned devices.	Implements emerging evidence-based technology into clinical practice and teaches learners to incorporate new applications of technology. Implements technological support programs into practice models.

* AAMC telehealth competency domains (ref. 23).

Education, include professionalism, patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, and systems-based practice. The AAMC has created a set of telehealth competencies that is applicable across the continuum of undergraduate, graduate, and continuing medical education (25). The development of telehealth core competencies, customized for rheumatology, will inform us of the domains for training, including patient care, communication, technology implementation and troubleshooting, ethics, and recognition of care delivery disparities.

History-taking skills are largely similar in virtual and in-person settings. The trainee can elicit the patient's history in a televisit, and when applicable, faculty preceptors can tease out additional details. The physical examination, a crucial tool for the rheumatologist, offers an important example of an area in which adaptations are needed. The virtual physical examination benefits from developing creative ways to assess the patient without touch (Table 2). The patient can be asked to record vital signs (weight, temperature, blood pressure, pulse, pulse oximetry) prior to the televisit, contingent upon availability of necessary equipment and confidence in the patient's ability to accurately record such data. Many examination techniques, such as joint range of motion, can be gleaned through a video visit. Fellows-in-training may ask the patient's caregiver or partner (if available) to perform certain portions of the physical examination. By using modeling, preceptors can demonstrate performance of a virtual examination for the fellow to enrich learning from different attending physicians' practices. Similarly, fellows may offer ideas on examination techniques to faculty preceptors, as we find ourselves in a virtual space in which all are learning and adapting. Additionally, measures of disease activity, such as the Routine Assessment of Patient Index Data 3, are conducive to a virtual encounter.

Several telemedicine competencies will require specific skills not currently included in the traditional fellowship curriculum. For example, the ability to effectively triage patients to virtual or in-person visits is an important skill, as recently included in the AAMC telehealth competencies (Table 1). We suggest that until the fellow has demonstrated proficiency in this competency, all triage decisions should be discussed with and approved by the faculty preceptor.

Several professionalism nuances exist in telemedicine. Webside manner will become an essential skill for our workforce; thus, we should be deliberate on how we teach this. Considerations for the virtual video visit, taking place from the provider's home or office must include professional video interaction, personal appearance, and appropriate virtual background. It will also be important for trainees to adapt to technology challenges and failures, in the moment, with patients who have variable technological skills.

Fellows-in-training must cultivate a professional and empathetic webside manner in order to successfully interact with patients during telemedicine visits. Providers often conduct visits from home, and therefore need to create a professional environment, characterized by a well-lit, quiet, private area with a neutral background. This may be difficult to achieve for trainees with young children, pets, or cramped living spaces (26). The camera should be oriented at eye level, and the provider should look directly into the camera as much as possible, to provide eye contact and help the patient feel connected. During the visit, providers should pause for several seconds after a patient finishes talking to ensure the patient has completed their thought and avoid over-talking. Empathy, easily demonstrated during in-person visits by simply touching the patient's shoulder or handing tissues to a tearful patient, must be adapted and more explicitly expressed during virtual visits.

Table 2. Telehealth examination modifications

Traditional examination	Modified telehealth examination performed by the patient or family member/home care provider, if available
Vital signs	Provide weight with an at-home scale. Provide blood pressure reading using an at-home device if available. Measure pulse, utilizing a clock and/or wearable device. Measure oxygen saturation with at-home pulse oximeter.
Scalp and/or temporal artery tenderness	Palpation of the scalp and temples to report tenderness.
Musculoskeletal	Perform range of motion assessment of joint in front of the camera (wrists, elbows, shoulders, and knees). Perform hypermobility maneuvers. Demonstrate for the patient self-palpation of the joints and entheses to determine the presence of joint and soft tissue tenderness. Fist formation to assess for fullness between metacarpophalangeal joints as well as fist formation ability. Demonstrate ability to fully flex fingers to palms to evaluate proximal interphalangeal joint swelling. The patient may send photos of the hands and feet in advance of the visit.
Motor strength	Ask the patient to arise from a chair without use of the hands. Ask a family member (if present) to assess strength.
Skin	In advance of the visit, ask the patient to send a picture of rash or nails. Ask the patient to open the mouth and measure the aperture with their fingers. Ask a family member to report rash presence on scalp or back of patient.

Physical space is typically not shared by trainees and faculty during virtual care clinics; therefore, coordination of the planned workflow is essential prior to the start of the virtual clinic. A preclinic virtual "huddle" can be used to determine the timing and mechanism for contacting and inviting the faculty preceptor to join the visit. Prior to the encounter, the fellow and attending physician should determine if precepting will occur with the patient in the virtual room, such as may happen with in-clinic bedside evaluations, or if the fellow and attending physician will speak privately before rejoining the patient. Consideration should be given to address trainee questions prior to inviting the faculty member into the virtual visit. Additionally, time offered for patient care questions after the visit is important. Feedback can be ad hoc or scheduled, and its timing and interval should be delineated. Assessment strategies to optimize patient–provider interactions, ensure excellence and safety in patient care delivery, and achievement of entrustable professional activities such that the trainee is deemed ready for independent practice are requisite within the virtual as well as in-person patient care space. While assessment of trainees may occur following patient encounters, there is also a need for metrics in evaluation, use of simulation to promote learning as in the observed structured clinical evaluation (OSCE), and integration of virtual care knowledge and skills in the rheumatology in-training and board certification examinations.

Conclusions

Telerheumatology training will adapt and change with time, but the infrastructure that we develop now will be foundational for this evolving era of patient care. Clear outcome measures with evidence-based practices must be developed. We have the opportunity to optimize virtual care delivery such that we may use this tool to enhance the care of patients with rheumatic diseases across all settings, rather than merely replacing the in-person visit. We must also recognize that training in telerheumatology is not limited to our fellowship programs. In addition to graduate medical education, it is imperative to consider the importance of telemedicine continuing medical education for providers in practice. Within the context of the COVID-19 pandemic, telerheumatology has become increasingly relevant. We must think critically to best harness this technology to move our specialty forward, and this not only begins, but pivots, in the training space.

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. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017;377:1585–92.
2. FitzGerald JD, Battistone M, Brown CR Jr, Cannella AC, Chakravarty E, Gelber AC, et al. Regional distribution of adult rheumatologists. Arthritis Rheum 2013;65:3017–25.
3. McDougall JA, Ferucci ED, Glover J, Fraenkel L. Telerheumatology: a systematic review. Arthritis Care Res (Hoboken) 2017;69:1546–57.
4. Solomon DH, Rudin RS. Digital health technologies: opportunities and challenges in rheumatology. Nat Rev Rheumatol 2020;16:525–35.
5. Liddy C, Moroz I, Mihan A, Nawar N, Keely E. A systematic review of asynchronous, provider-to-provider, electronic consultation services to improve access to specialty care available worldwide. Telemed J E Health 2019;25:184–98.
6. 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.
7. Kulcsar Z, Albert D, Ercolano E, Mecchella JN. Telerheumatology: a technology appropriate for virtually all. Semin Arthritis Rheum 2016;46:380–5.
8. Lowey NM. H.R.6074: Coronavirus Preparedness and Response Supplemental Appropriations Act, 2020. US House of Representatives. URL: https://www.congress.gov/bill/116th-congress/house-bill/6074.
9. Eberly LA, Khatana SA, Nathan AS, Snider C, Julien HM, Deleener ME, et al. Telemedicine outpatient cardiovascular care during the COVID-19 pandemic: bridging or opening the digital divide? Circulation 2020:510–2.
10. Bernstein S. Rheumatologists can now treat patients via telemedicine. The Rheumatologist. URL: https://www.the-rheumatologist.org/article/rheumatologists-can-now-treat-patients-via-telemedicine/.
11. 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.
12. Davis P, Howard R, Brockway P. Telehealth consultations in rheumatology: cost-effectiveness and user satisfaction. J Telemed Telecare 2001;7 Suppl 1:10–1.
13. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020;22:e17298.
14. Wood PR, Caplan L. Outcomes, satisfaction, and costs of a rheumatology telemedicine program: a longitudinal evaluation. J Clin Rheumatol 2019;25:41–4.
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) 2019;72:1404–9.
16. American College of Rheumatology. Core curriculum outline for rheumatology fellowship programs. 2015. URL: https://www.rheumatology.org/Portals/0/Files/Core-Curriculum-Outline.pdf.
17. Afshari M, Witek NP, Galifianakis NB. Education research: an experiential outpatient teleneurology curriculum for residents. Neurology 2019;93:170–5.
18. Jagolino-Cole AL, Bozorgui S, Ankrom CM, Bambhroliya AB, Cossey TD, Trevino AD, et al. Benchmarking telestroke proficiency: page-to-needle time among neurovascular fellows and attendings. Stroke 2017;48:2618–20.
19. Oldenburg R, Marsch A. Optimizing teledermatology visits for dermatology resident education during the COVID-19 pandemic. J Am Acad Dermatol 2020;82:e229.
20. Holmboe ES, Hawkins RE, Huot SJ. Effects of training in direct observation of medical residents' clinical competence: a randomized trial. Ann Intern Med 2004;140:874–81.
21. Peterson SJ, Nabors C, Frishman WH. Milestones: Direct observation may be the key to accelerated training. Am J Med 2013;126:1031–2.

22. Zha AM, Chung LS, Song SS, Majersik JJ, Jagolino-Cole AL. Training in neurology—adoption of resident teleneurology training in the wake of COVID-19: telemedicine crash course. Neurology 2020;95:404–7.
23. Association of American Medical Colleges. Competencies across the learning continuum series: new and emerging areas in medicine. 2020. URL: https://www.aamc.org/what-we-do/mission-areas/medical-education/cbme/competency.
24. Hatcher-Martin JM, Adams JL, Anderson ER, Bove R, Burrus TM, Chehrenama M, et al. Telemedicine in neurology: Telemedicine Work Group of the American Academy of Neurology update. Neurology 2020;94:30–8.
25. Association of American Medical Colleges. Innovations in health technology: telehealth. 2020. URL: https://www.aamc.org/what-we-do/mission-areas/patient-care/innovations-health-technology-telehealth.
26. Chua IS, Jackson V, Kamdar M. Webside manner during the COVID-19 pandemic: maintaining human connection during virtual visits. J Palliat Med 2020;23:1507–9.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1234–1243
DOI 10.1002/acr.24782

Preferences for Biologic Treatment in Patients With Psoriatic Arthritis: A Discrete Choice Experiment

Daniel Sumpton,[1] Ayano Kelly,[2] Jonathan C. Craig,[3] Geraldine Hassett,[4] Barry Kane,[5] Michael Oliffe,[6] Allison Tong,[7] and Martin Howell[7]

Objective. We aimed to assess patient preferences for the characteristics and outcomes of biologic and targeted synthetic disease-modifying antirheumatic drugs (DMARDs) to manage psoriatic arthritis.

Methods. We conducted a discrete choice experiment in patients with psoriatic arthritis from 3 rheumatology centers in Sydney, Australia. We assessed preferences for different attributes of biologic medications. The route and frequency of medications had a range of 5 levels, and the following 7 attributes had a range of 3 levels: the ability to attend to normal activities, improvements in joint pain, enthesitis and skin disease, chance of disease remission, risk of infection, and risk of severe adverse events. Multinomial logit models including a latent class model were used to calculate preferences.

Results. Of the 150 participants, 58.3% were female, with a median age of 53.5 years. The attributes in order of preference using the β coefficient in absolute values (95% confidence interval [95% CI]) were as follows: oral route compared to subcutaneous and intravenous routes (β coefficient 1.00 [fixed parameter]), avoiding severe side effects (β coefficient 0.72 [95% CI 0.50, 0.95]), increasing ability to attend to normal activities (β coefficient 0.66 [95% CI 0.36, 0.96]), avoiding infections (β coefficient 0.38 [95% CI 0.23, 0.53]), improvement in enthesitis pain (β coefficient 0.28 [95% CI 0.20, 0.36]), improvement in psoriasis (β coefficient 0.28 [95% CI 0.20, 0.36]), increasing chance of remission (β coefficient 0.27 [95% CI 0.19, 0.36]), and improvement in joint pain (β coefficient 0.26 [95% CI 0.00, 0.52]).

Conclusion. When choosing biologic medications, patients with psoriatic arthritis preferred oral medications. Patients prioritized avoiding severe complications, maintaining the ability to attend to work and normal activities, and avoiding infection over clinical measures of efficacy.

INTRODUCTION

Psoriatic arthritis is a chronic inflammatory arthritis treated with biologic and targeted synthetic disease-modifying antirheumatic drugs (DMARDs, referred to as biologic medications) in the severe forms of the disease (1,2). While their clinical efficacy is established, biologic medications have numerous routes, frequencies, and variable responses to the joint disease, enthesitis, and skin disease associated with the condition (1,2). There are no clinically useful biomarkers to guide the type, intensity, or length of therapy (3), and the characteristics and outcomes of biologic medications most important to patients when choosing treatment are unknown.

There is a growing awareness of the importance of understanding the priorities of patients with psoriatic arthritis to improve their care (3,4). Guidelines on the treatment of psoriatic arthritis also stress the need for shared decision-making (5,6). There is a need to understand the factors most important to patients when

[1]Daniel Sumpton, PhD: Sydney School of Public Health, University of Sydney, Sydney, New South Wales, Australia, Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, New South Wales, Australia, and Concord Repatriation and General Hospital, Concord, New South Wales, Australia; [2]Ayano Kelly, PhD: Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, New South Wales, Australia, Liverpool Hospital and Ingham Institute for Applied Medical Research, Liverpool, New South Wales, Australia, and Western Sydney Clinical School, University of New South Wales Medicine, Sydney, New South Wales, Australia; [3]Jonathan C. Craig, PhD: Flinders University, Adelaide, South Australia, Australia; [4]Geraldine Hassett, PhD: Liverpool Hospital and Ingham Institute for Applied Medical Research, Liverpool, New South Wales, Australia, Western Sydney Clinical School, University of New South Wales Medicine, Sydney, New South Wales, Australia, and Western Sydney University, Campbelltown, New South Wales, Australia; [5]Barry Kane, PhD: Concord Repatriation and General Hospital, Concord, New South Wales, Australia; [6]Michael Oliffe, MBBS: Liverpool Hospital and Ingham Institute for Applied Medical Research, Liverpool, New South Wales, Australia, and Western Sydney University, Campbelltown, New South Wales, Australia; [7]Allison Tong, PhD, Martin Howell, PhD: Sydney School of Public Health, University of Sydney, Sydney, New South Wales, Australia, and Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, New South Wales, Australia.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24782&file=acr24782-sup-0001-Disclosureform.pdf.

Address correspondence to Daniel Sumpton, PhD, Department of Rheumatology, Concord Hospital, Concord, New South Wales, Sydney 2170, Australia. Email: daniel.sumpton@health.nsw.gov.au.

Submitted for publication September 24, 2020; accepted in revised form September 7, 2021.

SIGNIFICANCE & INNOVATIONS

- When making decisions about biologic medications for the treatment of psoriatic arthritis, patients preferred oral treatments over subcutaneous or intravenous routes.
- Overall, participants prioritized avoiding severe adverse events, the ability to return to work and normal activities, and avoiding severe infection above clinical measures of disease improvement including improvement in joint pain, enthesitis, or skin disease.
- There were 2 distinct groups of patients with different priorities when choosing biologic medications, informing the need to discuss the characteristics and outcomes of biologic treatments with individual patients.

selecting a biologic medication, to identify the patient's main concerns and priorities, to enhance patient engagement in care, and to improve treatment adherence and outcomes that are important to patients. Biologic medications are also costly to patients, governments, or insurers, and it is important to understand the values that patients place on the qualities of medications to inform the design and approval of medications that patients are willing to accept.

Discrete choice experiments (DCEs) are routinely used to quantify patient treatment preferences and to determine the most important factors for patients drawn from real-life scenarios (7). This study aims to identify the characteristics and outcomes of biologic medications that patients with psoriatic arthritis prioritize when making choices about their treatment.

PATIENTS AND METHODS

Participant selection and recruitment. The study was conducted in 3 rheumatology centers in Sydney, Australia. Patients were eligible to participate if they were 18 years or older and were diagnosed as having psoriatic arthritis by a rheumatologist. Patients currently treated with biologic DMARDs (bDMARDs), targeted synthetic DMARDs (tsDMARDs), conventional DMARDs (cDMARDs), combination therapy, or no treatment were eligible to participate. Ethics approval was obtained from the South Western Sydney Research and Ethics Committee (HREC/15/LPOOL/560). All respondents provided informed consent.

DCEs. We used a DCE to quantify patient preferences for characteristics of biologic medications for use in psoriatic arthritis. In a DCE, participants are presented with multiple-choice sets containing different alternatives for a treatment or service. Each option describes characteristics or outcomes with varying levels (for example, a medication's ability to reduce pain with 3 levels of efficacy ranging from 20–50%) (8). Participants are asked to consider each of the qualities of a treatment and choose the alternative they prefer in each choice set. DCEs are based on random utility theory, which assumes that participants use rational decision-making to maximize their utility when making choices (9). It is assumed that people consider the attributes and levels of options and make a choice based on the outcome or characteristic they most value (10). These choices are then extrapolated to determine which characteristics or outcomes participants prefer over others.

Selection of characteristics and outcomes to describe biologic medications. The characteristics associated with treatment that patients consider when choosing biologic medications for psoriatic arthritis were identified through a systematic literature review of the experience of living with psoriatic arthritis and psoriasis (11). Also, standard clinical parameters and a review of previous DCEs in psoriatic arthritis, psoriasis, and rheumatoid arthritis were used to inform the development of characteristics (12–17). The levels of characteristics to describe the route and frequency of biologic medications were based on currently available biologic medications and the shortest frequency available for spondyloarthritis indications. The levels for outcomes of infection and rare adverse events were based on systematic reviews of the risks of biologics and registry data for psoriatic arthritis (18–21). The efficacy-related outcomes (joint pain, enthesitis, skin psoriasis, and disease remission) were based on published clinical data on expected American College of Rheumatology criteria for 20% improvement in disease activity (ACR20), ACR50, and ACR70 responses in psoriatic arthritis, Psoriasis Area and Severity Index 75 responses, and the chance of psoriatic arthritis disease remission based on a description of reaching minimal disease activity (21–23).

Characteristics of biologic medications or outcomes for their use and the different levels to describe these characteristics or outcomes were finalized after consulting with the research team, which included a patient research partner. From this process, there were 8 characteristics chosen, with 26 variable levels. The characteristics or outcomes and the range of their levels as presented to participants, shown in brackets, included the following: the ability to work or attend to normal activities (no improvement or negative effect, moderate improvement, significant improvement), route and frequency of medication (oral tablet once or twice daily, injection under the skin once every week or 2 weeks, injection under the skin once every 4 weeks, injection under the skin once every 12 weeks, infusion via a drip in the hospital every 6 weeks), the chance of remission (1 of every 10 people reach disease remission, 2 of every 10, 4 of every 10), effect on joint pain (3 of every 10 people have no joint pain, 4 of every 10, 6 of every 10), effect on tendon insertion pain (3 of every 10 people have no tendon insertion pain, 4 of every 10, 6 of every 10), effect on the skin (4 of every 10 people are clear of skin psoriasis,

6 of every 10, 8 of every 10), chance of typical infection (no increased risk, 2 times the number of infections in 1 year, 4 times the number of infections in 1 year), and the possibility of a rare but serious complication (no increased risk, 1 of 100 people, 10 of 100 people). A list of characteristics/outcomes, levels, and their description as presented to participants is shown in Table 1.

Survey design. The DCE was designed following established practices using the software Ngene V.1.2.0 (ChoiceMetrics) (24). We used a d-efficient fractional factorial design with 2 blocks of 9 questions using a multinomial logit model (MNL) with no interaction terms specified in the design. The final design indicated that statistical significance of main effects ($P < 0.05$) would be achieved with a sample size of at least 70 (25,26). The final DCE was piloted for design and comprehensibility with 10 patients who had psoriatic arthritis, resulting in some minor changes to the design of questions. Characteristic and outcome levels were represented by words, numbers, and pictograms to express the probability of an expected outcome. The survey was presented to the participants as follows: 1) introduction and explanation of the study, 2) sociodemographic and clinical data questions, 3) description and explanation of the characteristics/outcomes and levels, 4) an example choice set, and 5) the DCE. The DCE was unlabeled, and participants were asked to choose which medication they would prefer for treatment of their psoriatic arthritis of either medication A or medication B in 9 different scenarios. We collected sociodemographic information (sex, age, income, marital status) and clinical information (years since diagnosis of psoriatic arthritis and psoriasis, current and prior treatment, the experience of biologic medications and cDMARDs, self-reported patient pain, global health, and skin psoriasis, and self-reported knowledge of medications). An example choice set question is shown in Figure 1, representing 1 of 9 questions presented to each participant.

Data collection. Participants were invited to participate in an online survey through email. Data were collected from

Table 1. Choice set characteristics/outcomes and levels

Characteristic or outcome	Description given to participants	Level
Ability to work and attend to normal activities	Improvement in ability to do work, attend to regular exercise, ability to go out with friends or family	No improvement or negative effect; moderate improvement; significant improvement
Route and frequency of medication	The way the medication is given (a tablet, an injection, or in a drip) and how often it is taken (e.g., daily or every month)	Oral tablet once or twice daily; injection under the skin once every week or 2 weeks; injection under the skin once every 4 weeks; injection under the skin once every 12 weeks; infusion via a drip in the hospital every 6 weeks
Chance of disease remission	The number of people of 10 expected to have remission if taking the medication. Remission means that the patients has very few swollen and tender joints or tendons, low pain levels, minimal skin psoriasis, and feels quite well overall.	1 of every 10 people in remission; 2 of every 10 people in remission; 4 of every 10 people in remission
Effect on joint pain	The number of people of 10 expected to have no joint pain if taking the medication	3 of every 10 people have no joint pain; 4 of every 10 people have no joint pain; 6 of every 10 people have no joint pain
Effect on tendon insertion pain	The number of people of 10 expected to have no tendon insertion pain if taking the medication. For example, pain at the back of the heel, at the side of the elbows, or under the sole of the foot.	3 of every 10 people have no pain at tendon insertion; 4 of every 10 people have no pain at tendon insertion; 6 of every 10 people have no pain at tendon insertion
Effect on skin psoriasis	The number of people of 10 expected to have their skin clear of psoriasis if taking the medication.	4 of every 10 people are clear of skin psoriasis; 6 of every 10 people are clear of skin psoriasis; 8 of every 10 people are clear of skin psoriasis
Chance of typical infection	The increased number of typical and treatable infections each year, compared to not taking the medication (for example, a common cold or upper respiratory infection, or urinary tract infection), requiring antibiotics or no treatment and the drug to be temporarily stopped for a week.	No increased risk; 2 times the number of infections in 1 year; 4 times the number of infections in 1 year
Chance of rare but serious complication so that the medication is stopped for many months or permanently	For example, a severe infection (such as pulmonary tuberculosis or shingles, or severe pneumonia needing a hospital stay), severely low white cell count, a clot in the lung or leg, development of a cancer or organ damage. The complication is severe because the medication must be stopped for many months or permanently stopped or changed.	No increased risk; 1 of 100 people; 10 of 100 people

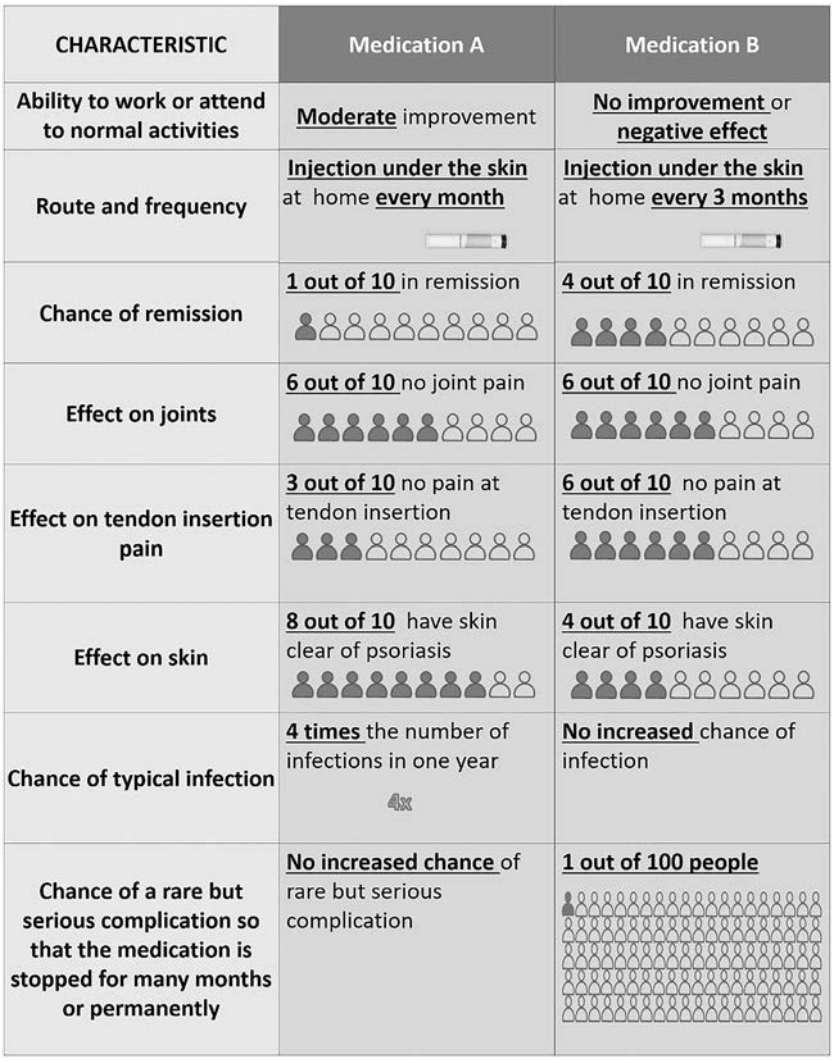


Figure 1. An example of a choice set.

March 8, 2020, to April 6, 2020. The DCE was a self-initiated survey, and self-reported characteristics were collected using Qualtrics survey software (Qualtrics XM).

Data analysis. The analysis of the DCE followed the general approach outlined by the International Society of Pharmacoeconomics and Outcomes Research and published guides to model specification (8,24,27). At study completion, data from the online surveys were extracted into an Excel spreadsheet and analyzed using NLOGIT software V.6 (Econometrics Software). All outcomes were coded as continuous variables, apart from the characteristic of frequency and route of medication, which were classified as categorical variables and coded using dummy coding. This led to individual characteristics for each route and frequency, leading to a total of 12 outcomes to describe in the classic MNL model. Outputs from the variables in the MNL model were given as β coefficients with 95% confidence intervals (95% CIs) and *P* values. Positive β coefficient values suggested that the participants preferred the characteristic or outcome, and a negative value suggested that the participants avoided the characteristic or outcome. Negative β coefficient values for adverse events and infections are presented as avoiding the outcome and given as absolute values to enhance the readability and understanding of the data.

A latent class MNL regression model was undertaken to provide an understanding of the heterogeneity of preferences among the participant sample by identifying classes of different preference structures (28). Output for each of the variables in each latent class were represented as β coefficients with 95% CIs. The latent class model takes into account the clinical and sociodemographic qualities of participants, for example, their experience of biologic medications. The likelihood of specific participant characteristics belonging to a latent class was described as β coefficients and odds ratios with 95% CIs. Membership

within a class is a latent property, and it is only possible to approximate the likelihood or probability that a respondent may be a member of a latent class. Linking to respondent covariates provides an estimation of the composition of the classes and the characteristics of individuals with different preference profiles. A trial and error process based on model fit criteria and the ability to predict the composition of individual classes was used to define the optimal number of latent classes included in the model. The final latent class model specification was determined based on the *P* value of specified parameters, log-likelihood tests, and Akaike information criteria. Demographic data were analyzed using descriptive statistics to describe the proportions of participants according to their sex, age, ethnicity, highest level of education, marital status, income level, years since diagnosis, history of psoriasis, and current and prior treatment.

RESULTS

Participant characteristics. A total of 164 patients responded to the survey from 199 invitations (82.4% response rate). One hundred fifty respondents completed the full survey (91.5% completion rate). The median age was 53.5 years (range 21–78 years), 86 participants (57.3%) were female, 71 participants (47.3%) had a university education, and 117 (78%) were married. Of the participants, 114 (76%) had psoriasis, and the median psoriatic arthritis disease duration was 5 years (range 0.2–44.0 years). A total of 75 (50%) were currently receiving treatment with a biologic, 83 (55.3%) had an experience of biologic therapy, and 41 (27.3%) had an experience of ≥2 biologics. Eighty-five (56.7%) were currently being treated with at least 1 cDMARD (methotrexate, sulfasalazine, or leflunomide). Participant characteristics are shown in Table 2.

Preferences for medication characteristics (MNL model). Preferences for the characteristics is indicated by the β coefficient values, with positive values indicating a preference for the characteristic, and negative values indicating against or avoiding the characteristic. The preferred characteristic in order of preference using the β coefficient in absolute values (95% CI) were as follows: oral route (β coefficient 1.00 [fixed-parameter]) compared to subcutaneous and intravenous routes, avoiding severe side effects (β coefficient 0.72 [95% CI 0.50, 0.95]), increasing ability attend to work and normal activities (β coefficient 0.66 [95% CI 0.36, 0.96]), avoiding infections (β coefficient 0.38 [95% CI 0.23, 0.53]), improvement in enthesitis pain (β coefficient 0.28 [95% CI 0.20, 0.36]), improvement in psoriasis (β coefficient 0.28 [95% CI 0.20, 0.36]), increasing chance of remission (β coefficient 0.27 [95% CI 0.19, 0.36]), and improvement in joint pain (β coefficient 0.26 [95% CI 0.00, 0.52]). The preferences for all characteristics are displayed in Table 3 and Figure 2.

Heterogeneity of preferences (latent class model). A latent class MNL model consisting of 2 classes was considered to

Table 2. Characteristics of the participants (n = 150)*

Characteristic	Value
Sex	
Female	86 (57.3)
Male	64 (42.7)
Age, years	
≤53	75 (50.0)
>53	75 (50.0)
Ethnicity	
Australian	92 (61.3)
Multiple†	17 (11.3)
British/Irish/Scottish	9 (6.0)
Southeast Asian	7 (4.7)
Southern and Central Asian	5 (3.3)
Other‡	20 (13.3)
Highest level of education	
University degree	71 (47.3)
Diploma, certificate, or trade	44 (29.3)
High school	35 (23.4)
Marital status	
Married or de facto	117 (78.0)
Single	17 (11.3)
Divorced	8 (5.3)
Widowed	5 (3.3)
Prefer not to answer	3 (2.0)
Income, $	
<35,000	19 (12.7)
35,000–65,000	17 (11.3)
65,001–95,000	26 (17.3)
95,001–125,000	18 (12.0)
125,001–150,000	12 (8.0)
>150,000	29 (19.3)
Prefer not to answer	29 (19.3)
Years since psoriatic arthritis diagnosis	
≤5	75 (50.0)
>5	75 (50.0)
History of psoriasis	
Psoriasis	114 (76)
No psoriasis	36 (24)
Current treatment	
Adalimumab	16 (10.7)
Golimumab	14 (9.3)
Etanercept	10 (6.7)
Certolizumab	6 (4.0)
Infliximab	1 (0.7)
Secukinumab	16 (10.7)
Ixekizumab	1 (0.7)
Ustekinumab	6 (4.0)
Risankizumab	1 (0.7)
Trial biologic medication	1 (0.7)
Tofacitinib	3 (2.0)
Methotrexate	60 (40.0)
Sulfasalazine	34 (22.7)
Leflunomide	8 (5.3)
No current systemic treatment	7 (4.7)
Experience of treatment	
Biologic therapy	83 (55.3)
Methotrexate	90 (60.0)
Sulfasalazine	58 (38.7)
Leflunomide	35 (23.3)

* Values are the number (%).
† Australian/British/Irish/Scottish (6), Australian/North African or Middle Eastern (4), Australian/Southeast Asian (2), Australian/Australian Aboriginal (1), Australian/Australian Aboriginal/Southeast Asian/British/Irish/Scottish (1), Australian/New Zealander (1), Australian/Western European (1), South America/Western European (1).
‡ North African and Middle Eastern, Pacific Islander, South American, New Zealander, Western European, Australian Aboriginal, North American, and Other (unspecified).

Table 3. Characteristic/outcome preferences for classic multinomial logit model model*

Characteristic/outcome	β coefficient (95% CI)	*P*
Oral route daily or twice a day	1.00 (fixed parameter)†	–
Avoiding risk of severe side effects	0.72 (0.50, 0.95)	0.01
Ability to work or attend to normal activities	0.66 (0.36, 0.96)	0.01
Avoiding risk of treatable infection	0.38 (0.23, 0.53)	0.01
Improvement in tendon insertion pain	0.28 (0.20, 0.36)	0.01
Improvement in skin psoriasis	0.28 (0.20, 0.36)	0.01
Chance of disease remission	0.27 (0.19, 0.36)	0.01
Improvement in joint pain	0.26 (0.00, 0.52)	0.05
SCI every 4 weeks compared to oral medication	−0.18 (−0.55, 0.19)	0.9‡
SCI every 1–2 weeks compared to oral medication	−0.23 (−0.67, 0.22)	0.9‡
SCI every 12 weeks compared to oral medication	−0.24 (−0.76, 0.28)	0.9‡
IV infusion every 6 weeks compared to oral medication	−1.02 (−1.69, −0.36)	0.01

* 95% CI = 95% confidence interval; IV = intravenous; SCI = subcutaneous injection.
† Fixed parameter for oral route as 1.0 for comparison to other routes and frequencies.
‡ Not statistically significant.

provide the best fit based on the statistical significance of characteristic coefficients. The preferences for each latent class are shown in Figure 3 and in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24782. The top 3 preferences for medication characteristics for participants in latent class 1 (LC-1) were avoiding rare but serious complications (β coefficient 1.35 [95% CI 0.83, 1.87]), the ability to work or attend to normal activities (β coefficient 1.19 [95% CI 0.53, 1.86]), and improvement in joint pain (β coefficient 0.82 [95% CI 0.23, 1.41]). Comparably, participants in latent class 2 (LC-2) had a preference for the route and frequency of medications, preferring a subcutaneous injection 1–2 times weekly (β coefficient 1.97 [95% CI 1.15, 2.80]), oral tablet once or twice daily (β coefficient 1.39 [95% CI 0.31, 2.47]), and a subcutaneous injection 4 times weekly (β coefficient 1.32 [95% CI 0.63, 2.01]). Compared to participants in LC-1, participants in LC-2 were less likely to prefer the ability to work or attend to normal activities (β coefficient 1.19 [95% CI 0.53, 1.86] in LC-1 versus β coefficient 0.54 [95% CI 0.05, 1.04] in LC-2) and avoiding the risk of a treatable infection (β coefficient 0.64 [95% CI 0.35, 0.93] in LC-1 versus β coefficient 0.33 [95% CI 0.05, 0.60] in LC-2).

The odds ratio of being in LC-1 relative to LC-2 for specific demographic characteristics and clinical features of participants is displayed in Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24782. Compared to participants in LC-2, participants in LC-1 were 21.2 times more likely to be older than 53 years, 2.9 times more likely to be female, 7.9 times more likely to have less pain (visual analog scale [VAS] <50 pain scores), 6.3 times more likely to have no experience receiving biologic therapy, 4.5 times more likely to have low perceived knowledge of medications for use in psoriatic arthritis, 3.0 times more likely to have a household income of >$66,500 per year, 3.6 times more likely to

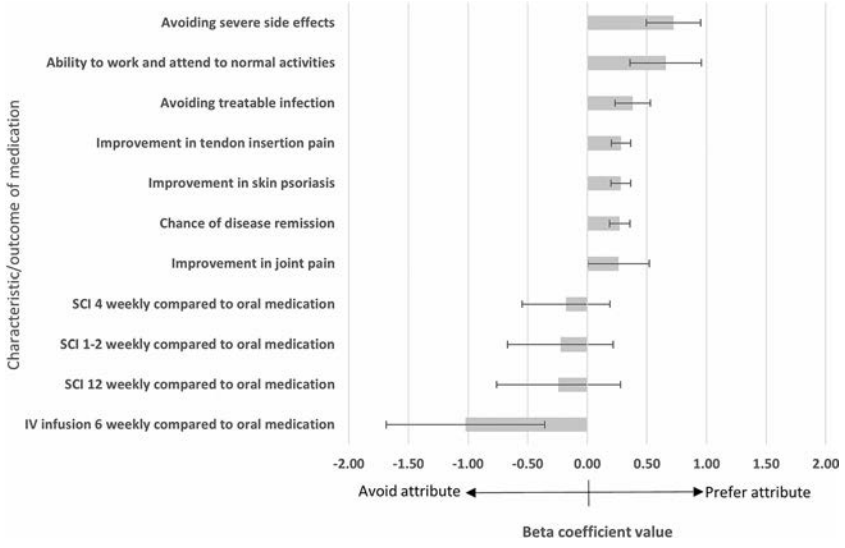


Figure 2. Characteristic/outcome preferences from the classic multinomial logit model. Error bars show the 95% confidence intervals. IV = intravenous; SCI = subcutaneous injection.

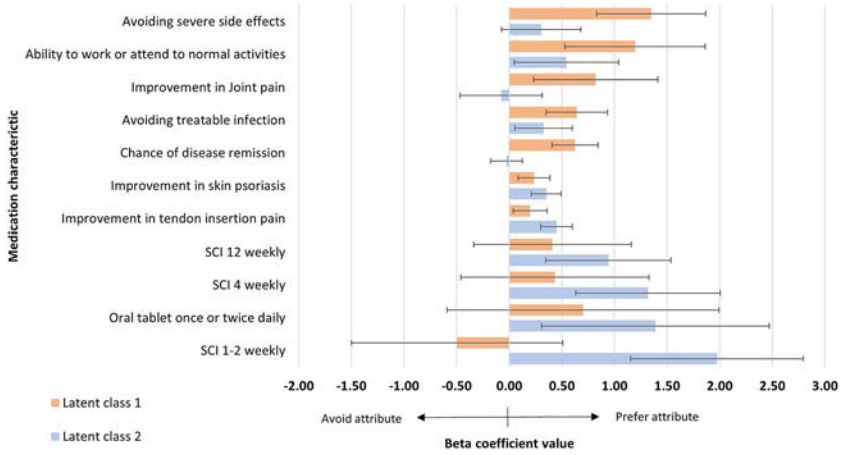


Figure 3. Preferences for medication characteristics and outcomes according to 2 latent classes. Intravenous infusion was made the reference variable when comparing the types of route and frequency in this analysis. Error bars show the 95% confidence intervals. SCI = subcutaneous injection.

have a psoriatic arthritis duration ≤5 years, and 1.9 times more likely to have skin scores on a VAS of ≥50 (better skin scores). Conversely, participants in LC-2 were, therefore, more likely to be younger than 53 years, be male, have higher pain levels (pain ≥50 on a VAS), have experience receiving biologics, have a higher perceived knowledge of medications, have a household income of <$66,500 per year, have a psoriatic arthritis duration of ≥5 years, and have skin scores on a VAS of ≤50 (worse skin scores).

DISCUSSION

For patients with psoriatic arthritis, when making decisions about selecting biologic medicines, participants prefer oral medication over subcutaneous medications of any frequency and to avoid intravenous infusions. Participants also give high priority to avoiding rare but serious complications, improving their ability to work and attend to normal activities, and avoiding the risk of treatable infection over clinical measures of improvement, including the chance of disease remission, improvement in joint pain, and improvement in joint tenderness or clearing of skin psoriasis.

The latent class model showed that there were 2 groups of participants. One group preferred avoiding serious complications, the ability to work or attend to normal activities, and improvement in joint pain (LC-1). Another group valued characteristics related to the route and frequency of medications (LC-2). The findings of the latent class model suggest that patients with less experience and knowledge of biologics and with earlier disease duration preferred avoiding serious complications of biologics but also value the ability to attend to work and social activities. Participants in LC-1 were more likely to have less pain. They were also more likely to preference joint pain as a valued outcome, suggesting that their pain may be treated by modalities other than biologic therapies. It is unclear why participants with higher income were more likely to be in LC-1 compared to LC-2, but this may relate to their preference for the ability to return to work and desire to protect their higher income. There may also be differences in health literacy between different income classes driving differences in preferences between these 2 groups. Participants in LC-2 were more experienced with biologics use and had higher perceived knowledge of their use. These participants were more likely to be comfortable with the most frequently used subcutaneous methods of treatment delivery suggestive of some experiential tolerance to subcutaneous medications.

The findings of the present study are consistent with qualitative data suggesting that fear of severe side effects and short-term side effects are significant factors for patients with psoriatic arthritis and spondyloarthropathy when choosing medications (11,29). Our findings also reflect similar DCE studies of preferences for biologics in psoriasis, showing that patients prefer avoiding severe adverse events above the value they place on treatment efficacy (16,17). Participants in our study also preferred oral tablets over intravenous or subcutaneous medications, consistent with findings from an industry-funded conjoint analysis of patient choice for biologics in psoriatic arthritis (15). While oral biologic medications have also been shown to be preferred in rheumatoid arthritis, this may be limited only to those who have no experience of subcutaneous biologic medications (30).

There are some differences between the present study and a recent DCE of Australian patients with rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. This industry-funded and commercially recruited study by Ho et al showed that patients prioritize the clinical efficacy of medications and preventing disease progression over mild-to-moderate and severe side effects (14). The differences between the present study and the study by Ho et al could be explained by the difference in population groups, the private recruitment methods used, and the detail of description of each attribute and level provided to participants in this study.

The present study has implications for the reporting of measures associated with work and social participation in clinical trials of psoriatic arthritis. Disability related to work and social functioning in psoriatic arthritis is high (31) and is a poor prognostic marker associated with increased disease activity (32). Participants in this study highly valued biologic medications because of their ability to improve their ability to work or attend social activities. Measures of these specific domains are not routinely reported in trials of biologic therapies. Clinical trials have traditionally included generic outcome measures of physical function such as the Medical Outcomes Study Short Form 36 (33) or the Health Assessment Questionnaire disability index (34). The Outcome Measures in Rheumatology (OMERACT) Psoriatic Arthritis Group has recently endorsed the Psoriatic Arthritis Impact of Disease measure (PSAID12) (35) as the preferred patient-reported measure for the health-related quality of life domain for use in clinical trials (36,37). The PSAID12 includes domains of "Work and/or leisure activities" and "Social participation" (35). The present study strengthens the conclusions from OMERACT that the PSAID12 should be included in clinical trials of psoriatic arthritis to determine the efficacy of biologics on the ability to undertake work and leisure activities. It may also be important to include the findings from the specific subdomains of "Work and/or leisure activities" and "Social participation" when reporting PSAID12 outcomes. The PSAID12 has also been validated as a tool responsive to change (38), and these subdomains related to work and social participation may be useful for clinicians to follow with regard to patient-centered responsiveness to treatment.

The present study has implications for the characteristics of biologics that rheumatologists focus on when making decisions about starting medications in time-limited consultations. Patients valued an oral route of drugs compared to subcutaneous forms and avoiding the complications of treatment over the clinical parameters of efficacy, including improvements in joint pain, tendon pain, skin disease, and disease remission. There is an apparent discrepancy between the outcomes of biologic medications most important to patients and the clinical parameters used by rheumatologists to inform measures of disease activity and treatment responsiveness (39–41). While clinical efficacy measures are useful to inform of the responsiveness to treatment, they may be less critical to patients compared to the route of medications, functional outcomes, and risk of side effects when choosing to initiate a biologic medication. The discrepancy in preference of characteristics around biologics may explain why there are reported differences in the expected results of treatment between rheumatologists and patients (42,43). The present study reinforces the need for rheumatologists to discuss the functional and social domains associated with treatment and the risk of infection and severe adverse events when making treatment decisions about biologic medications with their patients.

This study highlights the need to develop decision aids to guide shared decision-making choices and focus discussion on factors important to patients when choosing biologic medications in psoriatic arthritis. The latent class models showed that there are differences in preferred characteristics of biologic medications among patient groups, potentially related to their experience of biologics, age, level of perceived knowledge, and disease duration. Biologics-experienced patients were less concerned with the risk of adverse side effects compared to those without biologics experience. Patients also prioritized competing outcomes of medications, including enhancing social and work function as a result of medication-taking but also avoiding adverse events and infection. These findings strengthen the need for guidance when choosing biologics to strike a balance between these competing priorities. While decision aids have shown to be useful in rheumatoid arthritis at reducing decisional conflict, improving participation, and satisfaction in decision-making (44,45), there are no decision aids available for use in psoriatic arthritis. The preferences for outcomes described in the present study may be useful as a first step toward the development of a psoriatic arthritis–specific decision aid that helps align medication choice with patient knowledge and preferences.

The design of this study included guidance from a patient partner and was piloted with patients to ensure that participants comprehended the study. The design and analysis of the DCE followed established and rigorous methods. There are limitations to our study. Patients were recruited to this study based on physician diagnosis, and demographic and clinical parameters were self-reported. The present study was limited to 3 centers in Sydney, Australia. There were practical limitations on the number of attributes that could be included in the survey, and there may be attributes of importance that were not included. There were specific demographic factors that could affect the preferences for characteristics of biologics, including the proportion of participants receiving biologic therapies and the high number of university-educated participants. This may have implications for the generalizability of this study to other health settings, including settings in patients with lower education levels and lower levels of biologics access. The disease activity, function, and quality of life of participants was not measured in the study and is a limitation, as these factors may impact participant preferences.

In conclusion, when choosing biologic medications, patients with psoriatic arthritis preferred avoiding serious adverse events, the ability to work or attend to normal activities, and the risk of infection over measures of efficacy, including the chance of disease remission and improvements in joint pain, enthesitis, and psoriasis. Overall, patients prioritized oral medications over subcutaneous medications and intravenous infusions. There were differences in the preferences for biologic characteristics and outcomes based on clinical and participant characteristics. This study highlights the need for clinicians to focus on functional outcomes of treatment and adverse events and to consider individual patient preferences when discussing biologic medications.

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. Sumpton 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. Sumpton, Kelly, Craig, Hassett, Kane, Oliffe, Tong, Howell.

Acquisition of data. Sumpton.

Analysis and interpretation of data. Sumpton, Howell.

REFERENCES

1. Mahmood F, Coates LC, Helliwell PS. Current concepts and unmet needs in psoriatic arthritis. Clin Rheumatol 2018;37:297–305.
2. Winthrop KL, Weinblatt ME, Crow MK, Burmester GR, Mease PJ, So AK, et al. Unmet need in rheumatology: reports from the Targeted Therapies meeting 2018. Ann Rheum Dis 2019;78:872–8.
3. Helliwell P, Coates L, Chandran V, Gladman D, de Wit M, FitzGerald O, et al. Qualifying unmet needs and improving standards of care in psoriatic arthritis. Arthritis Care Res (Hoboken) 2014;66:1759–66.
4. Garrido-Cumbrera M, Hillmann O, Mahapatra R, Trigos D, Zajc P, Weiss L, et al. Improving the management of psoriatic arthritis and axial spondyloarthritis: roundtable discussions with healthcare professionals and patients. Rheumatol Ther 2017;4:219–31.
5. Coates LC, Kavanaugh A, Mease PJ, Soriano ER, Acosta-Felquer ML, Armstrong AW, et al. Group for research and assessment of psoriasis and psoriatic arthritis 2015 treatment recommendations for psoriatic arthritis. Arthritis Rheumatol 2016;68:1060–71.
6. Gossec L, Smolen JS, Ramiro S, de Wit M, Cutolo M, Dougados M, et al. European League Against Rheumatism (EULAR) recommendations for the management of psoriatic arthritis with pharmacological therapies: 2015 update. Ann Rheum Dis 2016;75:499–510.
7. Clark MD, Determann D, Petrou S, Moro D, de Bekker-Grob EW. Discrete choice experiments in health economics: a review of the literature. Pharmacoeconomics 2014;32:883–902.
8. Lancsar E, Fiebig DG, Hole AR. Discrete choice experiments: a guide to model specification, estimation and software. Pharmacoeconomics 2017;35:697–716.
9. McFadden D. Economic choices. Am Econ Rev 2001;91:351–78.
10. De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ 2012;21:145–72.
11. Sumpton D, Kelly A, Tunnicliffe DJ, Craig JC, Hassett G, Chessman D, et al. Patients' perspectives and experience of psoriasis and psoriatic arthritis: a systematic review and thematic synthesis of qualitative studies. Arthritis Care Res (Hoboken) 2020;72:711–22.
12. Rothery C, Bojke L, Richardson G, Bojke C, Moverley A, Coates L, et al. A discrete choice experiment to explore patients' willingness to risk disease relapse from treatment withdrawal in psoriatic arthritis. Clin Rheumatol 2016;35:2967–74.
13. Augustovski F, Beratarrechea A, Irazola V, Rubinstein F, Tesolin P, Gonzalez J, et al. Patient preferences for biologic agents in rheumatoid arthritis: a discrete-choice experiment. Value Health 2013;16:385–93.
14. Ho KA, Acar M, Puig A, Hutas G, Fifer S. What do Australian patients with inflammatory arthritis value in treatment? A discrete choice experiment. Clin Rheumatol 2020;39:1077–89.
15. Xu Y, Sudharshan L, Hsu MA, Koenig AS, Cappelleri JC, Liu WF, et al. Patient preferences associated with therapies for psoriatic arthritis: a conjoint analysis. Am Health Drug Benefits 2018;11:408–17.
16. Kromer C, Schaarschmidt ML, Schmieder A, Herr R, Goerdt S, Peitsch WK. Patient preferences for treatment of psoriasis with biologicals: a discrete choice experiment. PLoS One 2015;10:e0129120.
17. Schaarschmidt ML, Herr R, Gutknecht M, Wroblewska K, Gerdes S, Sticherling M, et al. Patients' and physicians' preferences for systemic psoriasis treatments: a nationwide comparative discrete choice experiment (PsoCompare). Acta Derm Venereol 2018;98:200–5.
18. Yiu ZZ, Exton LS, Jabbar-Lopez Z, Mohd Mustapa MF, Samarasekera EJ, Burden AD, et al. Risk of serious infections in patients with psoriasis on biologic therapies: a systematic review and meta-analysis. J Invest Dermatol 2016;136:1584–91.
19. Singh JA, Cameron C, Noorbaloochi S, Cullis T, Tucker M, Christensen R, et al. Risk of serious infection in biological treatment of patients with rheumatoid arthritis: a systematic review and meta-analysis. Lancet 2015;386:258–65.
20. Ritchlin CT, Stahle M, Poulin Y, Bagel J, Chakravarty SD, Kafka S, et al. Serious infections in patients with self-reported psoriatic arthritis from the Psoriasis Longitudinal Assessment and Registry (PSOLAR) treated with biologics. BMC Rheumatol 2019;3:52.
21. Lu C, Wallace BI, Waljee AK, Fu W, Zhang Q, Liu Y. Comparative efficacy and safety of targeted DMARDs for active psoriatic arthritis during induction therapy: a systematic review and network meta-analysis. Semin Arthritis Rheum 2019;49:381–8.
22. Ramiro S, Smolen JS, Landewé R, van der Heijde D, Dougados M, Emery P, et al. Pharmacological treatment of psoriatic arthritis: a systematic literature review for the 2015 update of the EULAR recommendations for the management of psoriatic arthritis. Ann Rheum Dis 2016;75:490–8.
23. Coates LC, Moverley AR, McParland L, Brown S, Navarro-Coy N, O'Dwyer JL, et al. Effect of tight control of inflammation in early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. Lancet 2015;386:2489–98.
24. Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health 2013;16:3–13.
25. Rose JM, Bliemer MC. Sample size requirements for stated choice experiments. Transportation 2013;40:1021–41.
26. De Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. Patient 2015;8:373–84.
27. Hauber AB, González JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, et al. Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR Conjoint Analysis Good Research Practices Task Force. Value Health 2016;19:300–15.
28. Zhou M, Thayer WM, Bridges JF. Using latent class analysis to model preference heterogeneity in health: a systematic review. Pharmacoeconomics 2018;36:175–87.
29. Kelly A, Tymms K, Tunnicliffe DJ, Sumpton D, Perera C, Fallon K, et al. Patients' attitudes and experiences of disease-modifying antirheumatic drugs in rheumatoid arthritis and spondyloarthritis: a qualitative synthesis. Arthritis Care Res (Hoboken) 2018;70:525–32.
30. Harrison M, Marra C, Shojania K, Bansback N. Societal preferences for rheumatoid arthritis treatments: evidence from a discrete choice experiment. Rheumatology (Oxford) 2015;54:1816–25.
31. Tillett W, de-Vries C, McHugh NJ. Work disability in psoriatic arthritis: a systematic review. Rheumatology (Oxford) 2012;51:275–83.
32. Wallenius M, Skomsvoll JF, Koldingsnes W, Rødevand E, Mikkelsen K, Kaufmann C, et al. Work disability and health-related quality of life in males and females with psoriatic arthritis. Ann Rheum Dis 2009;68:685–9.

33. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31:247–63.
34. Mease P, Strand V, Gladman D. Functional impairment measurement in psoriatic arthritis: importance and challenges. Semin Arthritis Rheum 2018;48:436–48.
35. Gossec L, de Wit M, Kiltz U, Braun J, Kalyoncu U, Scrivo R, et al. A patient-derived and patient-reported outcome measure for assessing psoriatic arthritis: elaboration and preliminary validation of the Psoriatic Arthritis Impact of Disease (PsAID) questionnaire, a 13-country EULAR initiative. Ann Rheum Dis 2014;73:1012–9.
36. Leung YY, Orbai AM, Ogdie A, Coates LC, de Wit M, Callis Duffin K, et al. The GRAPPA-OMERACT Psoriatic Arthritis Working Group at the 2018 Annual Meeting: report and plan for completing the core outcome measurement set. J Rheumatol Suppl 2019;95:33–7.
37. Højgaard P, Klokker L, Orbai AM, Holmsted K, Bartels EM, Leung YY, et al. A systematic review of measurement properties of patient reported outcome measures in psoriatic arthritis: a GRAPPA-OMERACT initiative. Semin Arthritis Rheum 2018;47:654–65.
38. Holland R, Højgaard P, Tillett W, Gossec L, de Wit M, Christensen R, et al. Evidence for Psoriatic Arthritis Impact of Disease (PsAID12) as Core instrument to measure health-related quality of life in psoriatic arthritis: a systematic review of psychometric properties. J Psoriasis Psoriatic Arthritis 2020;5:12–22.
39. Schoels M, Aletaha D, Funovits J, Kavanaugh A, Baker D, Smolen JS. Application of the DAREA/DAPSA score for assessment of disease activity in psoriatic arthritis. Ann Rheum Dis 2010;69:1441–7.
40. Coates LC, Lubrano E, Perrotta FM, Emery P, Conaghan PG, Helliwell PS. What should be the primary target of "treat to target" in psoriatic arthritis? J Rheumatol 2019;46:38–42.
41. Coates LC, Fransen J, Helliwell PS. Defining minimal disease activity in psoriatic arthritis: a proposed objective target for treatment. Ann Rheum Dis 2010;69:48–53.
42. 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.
43. Sacristán JA, Dilla T, Díaz-Cerezo S, Gabás-Rivera C, Aceituno S, Lizán L. Patient-physician discrepancy in the perception of immune-mediated inflammatory diseases: rheumatoid arthritis, psoriatic arthritis and psoriasis: a qualitative systematic review of the literature. PLoS One 2020;15:e0234705.
44. Nota I, Drossaert CH, Taal E, Vonkeman HE, Haagsma CJ, van de Laar MA. Evaluation of a patient decision aid for initiating disease modifying anti-rheumatic drugs. Arthritis Res Ther 2016;18:252.
45. Hazlewood GS, Marshall DA, Barber CE, Li LC, Barnabe C, Bykerk V, et al. Using a discrete-choice experiment in a decision aid to nudge patients towards value-concordant treatment choices in rheumatoid arthritis: a proof-of-concept study. Patient Prefer Adherence 2020; 14:829–38.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1244–1253
DOI 10.1002/acr.24578

Incidence of and Risk Factors for Heart Failure in Patients With Psoriatic Disease: A Cohort Study

Sahil Koppikar,[1] Keith Colaco,[2] Paula Harvey,[1] Shadi Akhtari,[1] Vinod Chandran,[3] Dafna D. Gladman,[3] Richard Cook,[4] and Lihi Eder[1]

Objective. To assess the incidence and risk factors for heart failure in patients with psoriatic disease and to describe their electrocardiographic and echocardiographic findings.

Methods. A cohort analysis was conducted involving patients with psoriatic disease followed prospectively from 1978 to 2018. Participants were assessed according to a standard protocol every 6 to 12 months. The primary outcome was the time to first event of heart failure, further classified into ischemic and nonischemic heart failure (secondary outcomes). The association between cardiovascular risk factors, measures of disease activity, and heart failure events was assessed using Cox proportional hazards regression. Electrocardiographic and echocardiographic findings associated with heart failure events were described.

Results. A total of 1,994 patients with psoriatic disease were analyzed, with 64 incident heart failure events (38 ischemic, 26 nonischemic). The incidence rate of first heart failure event was 2.85 per 1,000 patient-years. In all events, the most common electrocardiographic findings were atrial fibrillation (22%) and bundle branch blocks (29%). Echocardiogram revealed 37% reduced ejection fraction and 63% preserved ejection fraction. In multivariable analysis, independent risk factors for all heart failure events were ischemic heart disease, adjusted mean tender joint count, adjusted mean swollen joint count, adjusted mean erythrocyte sedimentation rate, adjusted mean C-reactive protein level, and physical function (by Health Assessment Questionnaire) (all $P < 0.05$). Minimal disease activity state was protective for all heart failure ($P < 0.05$).

Conclusion. Increased risk of heart failure is associated with a combination of known cardiovascular risk factors and measures of disease activity, particularly in nonischemic heart failure. The effect of inflammation on heart failure may be partially independent of atherosclerotic disease.

INTRODUCTION

Psoriasis is an immune-mediated skin disease with a prevalence of 2–3% in North American and European populations (1). Psoriatic arthritis (PsA) is an inflammatory musculoskeletal disease that affects 14–30% of psoriasis patients (2,3). Both conditions, collectively termed psoriatic disease, are associated with numerous comorbid conditions, including cardiometabolic, gastrointestinal, and renal diseases, as well as malignancies, infections, and mood disorders (4).

Among the comorbid conditions, cardiovascular diseases are of particular importance as they directly impact the patients' mortality (5). Several studies have suggested that psoriasis and PsA are independent risk factors for major adverse cardiovascular events, including myocardial infarction (MI), stroke, and cardiovascular death (6–9). The increased cardiovascular morbidity in psoriatic disease can be partially attributed to the high prevalence of metabolic abnormalities (9,10), such as impaired glucose tolerance, atherogenic lipid profiles, and unhealthy lifestyle habits (smoking, physical inactivity) that are common in these patients. These factors, in addition to systemic inflammation, may contribute to atherogenesis and cardiovascular events.

Most of the studies have focused on the link between psoriatic disease and atherosclerotic disease, hypertension, obesity,

Supported by a Young Investigator Operating grant from the Arthritis Society and Early Researcher Award from the Ontario Ministry of Research, Innovation, and Science. The Psoriatic Arthritis program is supported by the Krembil Foundation.

[1]Sahil Koppikar, MD, Paula Harvey, MD, PhD, Shadi Akhtari, MD, Lihi Eder, MD, PhD: University of Toronto and Women's College Hospital, Toronto, Ontario, Canada; [2]Keith Colaco, MSc: Women's College Hospital, Toronto, Ontario, Canada; [3]Vinod Chandran, MD, PhD, Dafna D. Gladman, MD: University of Toronto, Krembil Research Institute, Toronto Western Hospital, and University of Toronto, Toronto, Ontario, Canada; [4]Richard Cook, PhD: University of Waterloo, Waterloo, Ontario, Canada.

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

Address correspondence to Lihi Eder, MD, PhD, Women's College Research Institute, Room 6326, Women's College Hospital, 76 Grenville Street, Toronto, Ontario, M5S 1B2 Canada. Email: lihi.eder@wchospital.ca.

Submitted for publication April 13, 2020; accepted in revised form February 9, 2021.

SIGNIFICANCE & INNOVATIONS

- The findings of this study demonstrate that the increased risk of heart failure in psoriatic disease is associated with a combination of known cardiovascular risk factors and high disease activity, particularly in nonischemic heart failure, suggesting that the effect of inflammation on heart failure may be independent of atherosclerotic disease.
- Independent risk factors of heart failure events were ischemic heart disease, adjusted mean swollen joint count, adjusted mean tender joint count, adjusted mean erythrocyte sedimentation rate, adjusted mean C-reactive protein level, and physical function (by Health Assessment Questionnaire), while minimal disease activity state was protective against heart failure events.
- This study demonstrates that psoriatic disease and heart failure are linked, independent of cardiovascular risk factors, and emphasizes the importance of controlling disease activity and inflammation to reduce heart failure risk in these patients.

and metabolic syndrome. Heart failure is another major cardiovascular event with a large global burden of disease. However, data are limited describing the association between psoriatic disease and heart failure. A recent meta-analysis found that patients with PsA have a 32% increased risk of developing heart failure compared to the general population (8). Patients with psoriasis have a similarly increased risk between 20% and 60% (11,12). The association between psoriasis and cardiomyopathy has been described, with the majority of patients developing dilated cardiomyopathy (13).

More recently, echocardiographic studies have identified subclinical myocardial dysfunction in patients with psoriatic disease without cardiovascular risk factors (14,15). Although the cause of heart failure has been traditionally attributed to atherosclerotic disease, systemic inflammation is now being recognized as an independent risk factor for heart failure development, especially with preserved ejection fraction. Sustained inflammation may directly cause myocardial hypertrophy through a proinflammatory cytokine milieu, contributing to ventricular stiffness and diastolic dysfunction. A recent paradigm of "epicardial adipose inflammation" suggests that chronic systemic inflammation activates epicardial fat to produce adipocytokines that are transmitted directly to underlying tissues, leading to volume stress, fibrosis, and impaired cardiac distensibility (16,17). Furthermore, arrhythmias and valvular dysfunction are often found in psoriatic disease, conceivably being another nonischemic mechanism of heart failure development (18,19). However, few studies have thus far attempted to elucidate the link between cardiometabolic abnormalities, psoriatic disease–related factors, and heart failure development to determine the independent effect of psoriatic disease activity on heart failure.

In this cohort study we aimed to estimate the cumulative incidence of heart failure in patients with psoriatic disease, identify independent risk factors for developing heart failure events, and describe the electrocardiographic (ECG) and transthoracic echocardiographic (TTE) findings in patients experiencing heart failure events.

PATIENTS AND METHODS

Patients and setting. A cohort analysis was conducted in patients followed from 1978 to 2018 at the University of Toronto Psoriatic Disease clinic. The clinic was established in 1978 and enrolls patients with psoriatic disease who are all followed with the same protocol. Among the PsA patients, 98% meet the Classification of Psoriatic Arthritis classification criteria (20). The psoriasis patients are enrolled based on a dermatologist confirmed diagnosis of psoriasis without arthritis. Patients attending the clinic are enrolled in an ongoing prospective study aimed at assessing prognostic factors in psoriatic disease. Each patient is assessed at 6–12-month intervals according to a standard protocol (21). As part of the study protocol, information is prospectively collected, including patient demographic characteristics, lifestyle habits, medical history, medication use, disease-related outcomes, laboratory findings, and imaging studies. Laboratory tests for lipid profile and inflammatory markers are performed every 6–12 months as part of the study protocol. All data are stored in a web-based computerized database. Patients who had <1 year of follow-up or who developed a heart failure event prior to the first clinic visit were excluded from the study. All subjects' written consent was obtained according to the Declaration of Helsinki. The study has been approved by the University Health Network and Women's College Research Institute ethics boards.

Primary event definition. The primary end point was defined as the occurrence of the first event of heart failure exacerbation. Heart failure events were further classified as ischemic heart failure (ischemic cardiovascular disease defined as angina, MI, or revascularization occurring prior to or at the same time as the first heart failure event) or nonischemic heart failure (no prior ischemic cardiovascular disease). Potential heart failure events were first identified by searching the cohort database and linking to provincial mortality and hospitalization databases. Subsequently, complete medical records pertaining to the heart failure event were obtained, where available, from the patient's primary care provider and specialists. Each identified potential heart failure event was adjudicated by reviewing data from hospital admissions, death certificates, and medical records from relevant specialists. Uncertain cases were discussed with a cardiologist (SA) to determine whether to consider them as heart failure events.

Based on the level of evidence in the medical records, heart failure events were classified as either definite, probable, or

possible. A definite heart failure event was defined by any of the following: record of typical clinical symptoms and signs on physical examination, record of typical diagnostic test, including chest radiographs or echocardiogram, discharge summary indicating heart failure as an admission diagnosis, or a heart failure event, documented by a cardiologist; provincial database documentation of heart failure admission and a history of heart failure documentation by any physician or documentation in the cohort database; and death record of heart failure as the primary cause of death. A probable heart failure event was defined by any of the following: provincial database documentation of heart failure only, or history of heart failure documentation by a nonspecialist (i.e., rheumatologist). A possible heart failure event was defined by any of the following: only cohort database documentation with no additional supportive documents available; positive test results but no additional documentation; cause of death reported as heart failure only by nonmedical personnel (i.e., family member). Events not meeting any of the aforementioned criteria were excluded (i.e., inconsistencies between documentation in cohort database and medical records).

Risk factors of heart failure. Both traditional cardiovascular risk factors and psoriatic disease–related variables were assessed as risk factors of heart failure events from the time of entry into the cohort until the last visit prior to the heart failure event (in patients who developed heart failure) or the visit prior to the last date known to be alive (in patients who remained event-free). Traditional cardiovascular risk factors were defined based on the use of medications or findings on physical examination or laboratory tests. The following traditional cardiovascular risk factors were assessed: sex, smoking (current, past, or never), diabetes mellitus, prior history of hypertension based on patient report and/or the use of antihypertensive medications, measured systolic blood pressure, body mass index (BMI), levels of triglycerides, and total cholesterol. In addition, an established history of ischemic heart disease (angina, MI, or coronary revascularization) was considered as an additional risk factor of heart failure.

The following psoriatic disease–related variables were assessed: tender joint count (TJC) and swollen joint count (SJC), clinically damaged joint count (defined as the presence of limitation of range of movement of >20% of the range not related to the presence of joint effusion, the presence of joint deformity, subluxation, loosening or ankylosis), Psoriasis Area and Severity Index (PASI), erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) level, Health Assessment Questionnaire (HAQ) disability index, patient pain score, patient global assessment (PtGA) of arthritis, and minimal disease activity (MDA) state (defined by the presence of at least 5 of the following 7 items: TJC ≤1, SJC ≤1, PASI score of ≤1 or body surface area covered by psoriasis ≤3%, patient pain visual analog scale [VAS] score of ≤15, PtGA VAS score of ≤20, HAQ score of ≤0.5, and tender entheseal points ≤1) (22).

Chart reviews. Structured chart abstraction was performed by a single investigator (SK) for any heart failure with available records to identify clinical, ECG, and TTE findings. Clinical variables that were assessed included documented cause of underlying heart failure (ischemia, nonischemic cardiomyopathy, valvular abnormalities, arrhythmia, other) during the heart failure event, revascularization procedure during the heart failure event (percutaneous coronary intervention or coronary artery bypass), and death from a heart failure event. ECG variables assessed were left and right bundle branch blocks, nonspecific intraventricular conduction delay, any degree of atrioventricular heart block, left and right ventricular hypertrophy, atrial flutter, atrial fibrillation, evidence of previous MI (pathologic Q waves), nonspecific ST-T wave abnormalities, and the presence of a pacemaker. Echocardiographic variables that were assessed included ejection fraction, systolic dysfunction (ejection fraction <40%), preserved ejection fraction (ejection fraction >40%), concentric remodeling, left or right ventricular hypertrophy, left or right ventricular dilation, regional or global wall motion abnormalities, right and left atrial dilation, and valvular abnormalities. If ECG or TTE reports were not available at the time of the heart failure event, the closest recorded measurements from within 3 months after the event were used.

Statistical analysis. Descriptive statistics were computed for baseline covariates (at cohort entry) with continuous variables summarized by their mean ± SDs and categorical variables summarized by proportions. The time from the date of birth to the date of the first heart failure event was the response of interest; individuals who were event-free at the date they were last known to be alive were censored at their corresponding age. Non-heart failure death was considered as a competing event. Nonparametric estimates of the cumulative incidence function for heart failure events were obtained, with non-heart failure death as a competing risk. In addition, nonischemic heart failure was considered as a competing risk for ischemic heart failure and vice versa.

Cox proportional hazards models were fitted with age defining the time scale and the age at diagnosis of PsA or psoriasis (depending on the diagnosis) as the left-truncation time (23). Since the study period spans over 4 decades, patients were assigned 1 of the following 3 indicators based on the year they entered into the cohort: period 1 (1978–1990), period 2 (1991–2005), and period 3 (2006–2017) to adjust for potential secular effects of the cohort entry date. This indicator was included as a covariate in the regression model.

To explore the effects of time-dependent covariates, Cox regression models were fitted with covariates updated at each clinic visit. In separate Cox regression models, we computed the cumulative mean value of the covariates and used these as time-dependent covariates to study the effect of sustained elevation of the covariate over time. Both standard time-dependent and time-dependent cumulative mean values were assessed for their

association with the event in separate regression models to account for the dynamic nature of these variables over time. The following variables were considered as time-varying risk factors of incident heart failure: PASI score, TJC and SJC, damaged joint count, pain score, HAQ score, MDA, CRP level, and ESR.

The initial univariate model included each of these risk factors as a single covariate in the regression model. Subsequently, each of the above factors was included in a separate multivariable regression model, adjusting for sex, hypertension, diabetes mellitus, BMI, ischemic heart disease, total cholesterol, triglyceride level, and smoking. All of these variables (except sex) were considered as time-varying covariates. A composite of any heart failure event was considered as the primary outcome, while secondary outcomes were based on the time to the first ischemic heart failure event and the time to the first nonischemic heart failure event, which were analyzed separately.

Multiple imputation (using proc MI and proc mianalyze in SAS) was used to impute missing data in conjunction with the Cox model. The full conditional specification and predictive mean matching methods were specified as methods of imputation. The imputation model included the demographic variables, laboratory test results, medications, duration of disease, and measures of skin and joint disease activity as outcomes of interest. Five imputed data sets were used in the analysis.

RESULTS

A total of 2,205 patients who had >1 visit and were enrolled into the cohorts from January 1, 1978 to December 31, 2018

Table 1. Baseline characteristics of patients by diagnosis (at study entry)*

	No heart failure (n = 1,930)	All heart failure (n = 64)	Ischemic heart failure (n = 38)	Nonischemic heart failure (n = 26)
PsA, no. (%)	1,349 (69.9)	60 (93.8)	36 (94.7)	24 (92.3)
Cutaneous psoriasis, no. (%)	581 (30.1)	4 (6.2)	2 (5.3)	2 (7.7)
Age, years	45.0 ± 13.2	53.6 ± 12.7	56.1 ± 11.6	49.9 ± 13.5
Male, no. (%)	1,082 (56.1)	35 (54.7)	25 (65.8)	10 (38.5)
Smoking status, no. (%)				
Current	440 (22.8)	18 (28.1)	13 (34.2)	5 (19.2)
Past	506 (26.2)	16 (25)	11 (28.9)	6 (23.1)
Caucasian	1,608 (83.3)	58 (90.6)	35 (92.1)	23 (88.5)
Year of study entry, no. (%)				
1978–1995	334 (91.5)	31 (8.5)	18 (4.9)	13 (3.6)
1996–2005	448 (95.1)	23 (5)	14 (3)	9 (1.9)
2006–2017	1,154 (99.1)	10 (0.9)	6 (0.5)	4 (0.3)
Diabetes mellitus, no. (%)	103 (5.3)	11 (17.2)	7 (18.4)	4 (15.4)
Hypertension, no. (%)	306 (15.9)	25 (39.1)	15 (39.5)	10 (38.5)
Body mass index, kg/m^2	28.3 ± 6	30.3 ± 5.7	30.2 ± 6.2	30.6 ± 5
Cholesterol, mmol/liter	5.0 ± 1.1	5.2 ± 1.2	5.3 ± 1.2	5.1 ± 1.2
Triglycerides, mmol/liter	1.7 ± 1.1	2 ± 1.1	2 ± 1	2 ± 1.1
Ischemic heart disease	33 ± 1.7	9 ± 14.1	9 ± 23.7	0 ± 0
Duration of PsA, years	6.2 ± 7.7	9.7 ± 11.6	9.4 ± 10.8	10.1 ± 12.9
Duration of psoriasis, years	15.8 ± 13.2	17.7 ± 15.2	16.2 ± 15.3	19.9 ± 15
Tender joint count (0–68)	4.8 ± 7.8	8.5 ± 10.2	7.0 ± 7.4	10.6 ± 13.2
Swollen joint count (0–66)	2.4 ± 4.3	3.3 ± 6.3	2.1 ± 3	5.2 ± 9
Damaged joint count (0–68)	1.7 ± 5.6	4 ± 7.6	3.3 ± 7.7	5.1 ± 7.4
ESR, mm/hour	19.2 ± 19.3	28.8 ± 23	25.3 ± 23.1	33.9 ± 22.2
C-reactive protein level, mg/liter	9 ± 15.2	12.9 ± 22.4	13.1 ± 23.6	12.7 ± 21
PASI (0–72)	5.2 ± 7.1	6.7 ± 9.8	6.4 ± 11	7.1 ± 7.9
HAQ (0–3)	0.5 ± 0.6	0.8 ± 0.7	0.8 ± 0.7	0.9 ± 0.7
Pain (0–10)	3.6 ± 2.9	4.6 ± 2.7	4.6 ± 2.8	4.7 ± 2.8
PtGA arthritis (0–10)	3.5 ± 3	4.6 ± 2.9	4.5 ± 3	4.8 ± 2.7
MDA state, no. (%)	792 (41)	13 (20.3)	9 (23.6)	5 (19.2)
Methotrexate, no. (%)	270 (14)	3 (4.7)	3 (7.9)	0 (0)
Prednisone, no. (%)	52 (2.7)	5 (7.8)	3 (7.9)	2 (7.7)
Leflunomide, no. (%)	20 (1)	1 (1.6)	0 (0)	1 (3.9)
Sulfasalazine, no. (%)	51 (2.6)	4 (6.3)	4 (10.5)	0 (0)
TNF inhibitor, no. (%)	80 (4.2)	0 (0)	0 (0)	0 (0)
IL-17 inhibitor, no. (%)	1 (0.1)	0 (0)	0 (0)	0 (0)
IL-12/23 inhibitor, no. (%)	12 (0.6)	0 (0)	0 (0)	0 (0)

* Values are the mean ± SD unless indicated otherwise. All baseline characteristics were measured at time of entry into the cohort. BP = blood pressure; ESR = erythrocyte sedimentation rate; HAQ = Health Assessment Questionnaire; IL = interleukin; MDA = minimal disease activity; PASI = Psoriasis Area and Severity Index; PsA = psoriatic arthritis; PtGA = patient global assessment; TNF = tumor necrosis factor.

Table 2. Estimated cumulative incidence of developing a heart failure event in psoriatic disease*

	All heart failure			Ischemic heart failure			Nonischemic heart failure		
Age, years	Overall	Male	Female	Overall	Male	Female	Overall	Male	Female
50	0.3	0.2	0.4	0.1	0.1	0.0	0.3	0.1	0.4
60	1.5	1.3	1.6	0.6	1.0	0.2	0.8	0.3	1.4
70	4.1	5.5	2.5	2.3	3.8	0.6	1.7	1.6	1.9
80	10.2	10.3	9.8	5.9	6.0	5.3	4.2	4.0	4.4
90	22.5	26.2	20.4	16.4	21.9	13.1	5.8	4.0	7.1

* Values are the percentage.

were identified. Seven patients were excluded because they had developed a heart failure event prior to the first visit in the clinic, and 204 patients were excluded due to <1 year of follow-up. A total of 1,994 patients with psoriatic disease were analyzed. The 2 cohorts had a total of 22,437 person-years of follow-up, with a mean of 11.3 ± 8.7 years per person. The characteristics of the study population at baseline are summarized in Table 1.

Cumulative incidence of heart failure events in psoriatic disease. During the follow-up period, a total of 64 new heart failure events occurred (38 ischemic, 26 nonischemic). As expected, the risk of developing a heart failure event increased with age. The incidence rate of first heart failure event was 2.85 per 1,000 patient-years. The cumulative incidence of all heart failure events was 4.1% by age 70 years, 10.2% by age 80 years, and 22.5% by age 90 years. The cumulative incidence of ischemic heart failure events was numerically higher than nonischemic heart failure events in each age group. The rise in heart failure events in female patients lagged by a decade compared to male patients (eighth decade in women versus seventh decade in men) (Table 2 and Figure 1).

Risk factors of heart failure events in univariate analysis. In univariate analysis, traditional cardiovascular risk

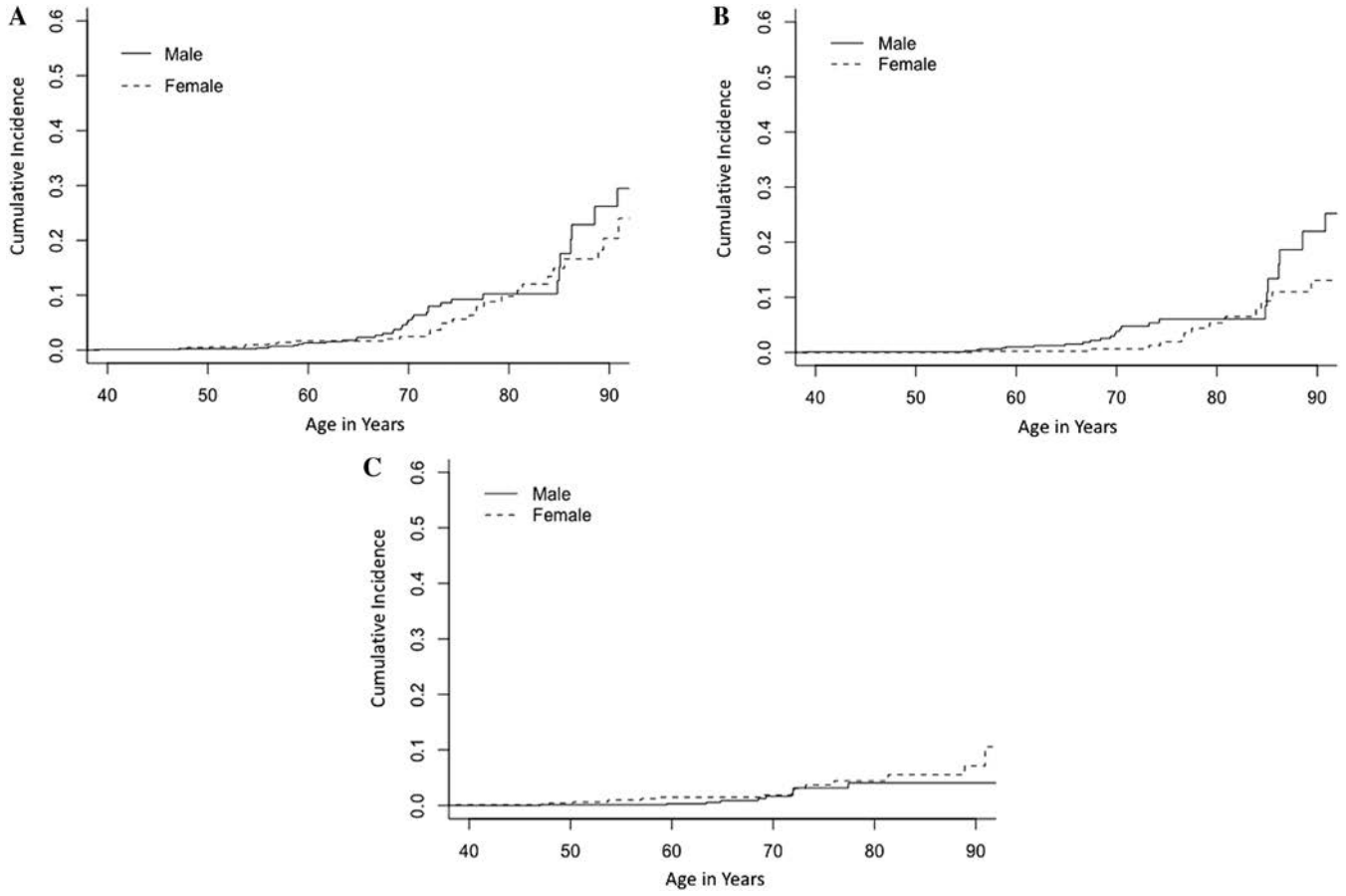


Figure 1. Cumulative incidence of experiencing various events by sex in patients with psoriatic disease. **A**, All heart failure, **B**, Ischemic heart failure, and **C**, nonischemic heart failure.

Table 3. Estimates from fitting univariable Cox proportional hazards model with age as the time scale (n = 1,991 [n = 64 events])*

Variable	All heart failure (n = 64 events)		Ischemic heart failure (n = 38 events)		Nonischemic heart failure (n = 26 events)	
	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*
Male	1.30 (0.78–2.14)	0.31	2.11 (1.07–4.16)†	0.03†	0.60 (0.27–1.33)	0.20
Smoking‡						
Current vs. never	1.40 (0.63–3.12)	0.40	3.00 (1.09–8.17)†	0.03†	0.52 (0.15–1.82)	0.30
Past vs. never	1.13 (0.64–2.05)	0.66	1.92 (0.86–4.26)†	0.11†	0.54 (0.21–1.44)	0.22
Diabetes mellitus‡	3.42 (1.97–5.92)†	<0.0001†	4.17 (2.12–8.21)†	<0.0001†	2.30 (0.85–6.26)	0.10
Hypertension‡	3.06 (1.80–5.31)†	<0.0001†	4.53 (2.10–9.75)†	0.0001†	1.99 (0.88–4.53)	0.09
Body mass index, kg/m^2‡	1.04 (1.00–1.09)	0.056	1.02 (0.96–1.08)	0.55	1.08 (1.01–1.15)†	0.02†
Cholesterol, mmol/liter‡	0.76 (0.59–0.96)†	0.02†	0.75 (0.55–1.01)	0.06	0.79 (0.54–1.17)	0.24
Triglycerides, mmol/liter‡	1.13 (0.91–1.43)	0.25	1.28 (0.99–1.63)	0.06	1.00 (0.68–1.46)	0.97
Ischemic heart disease‡	7.76 (4.61–13.19)†	<0.0001†	25.1 (11.5–54.5)†	<0.0001†	0.40 (0.05–3.03)	0.37
TJC (0–68)§	1.55 (1.21–1.97)†	0.0004†	1.54 (1.12–2.14)†	0.008†	1.65 (1.19–2.32)†	0.004†
AM TJC (0–68)¶	1.70 (1.27–2.29)†	0.0004†	1.61 (1.07–1.41)†	0.02†	1.92 (1.32–2.80)†	0.0007†
SJC (0–66)§	1.97 (1.12–3.46)†	0.02†	1.99 (0.98–4.10)	0.057	2.32 (1.01–5.31)†	0.05†
AM SJC (0–66)¶	2.25 (1.34–3.97)†	0.005†	1.62 (0.62–4.26)	0.33	3.46 (1.77–6.62)†	0.0003†
PASI (0–72)§	1.24 (0.82–1.90)	0.30	0.86 (0.42–1.77)	0.69	1.55 (0.97–2.48)	0.06
AM PASI (0–72)¶	1.70 (1.15–2.50)†	0.008†	1.88 (1.10–3.22)†	0.02†	1.60 (0.91–2.80)	0.10
Damaged joint count§	1.17 (0.99–1.38)	0.06	1.19 (0.96–1.45)	0.11	1.26 (0.99–1.60)	0.06
ESR, mm/hour§	1.17 (1.05–1.32)†	0.004†	1.10 (0.93–1.30)	0.24	1.28 (1.11–1.49)†	0.0009†
AM ESR, mm/hour¶	1.17 (1.01–1.38)†	0.035†	1.12 (0.91–1.36)	0.29	1.28 (1.06–1.57)†	0.01†
CRP level, mg/liter§	1.12 (0.98–1.27)	0.09	1.10 (0.91–1.32)	0.31	1.13 (0.88–1.43)	0.33
AM CRP level, mg/liter¶	1.26 (1.03–1.51)†	0.02†	1.25 (1.02–1.52)†	0.03†	1.26 (0.92–1.72)	0.13
HAQ (0–3)‡	2.36 (1.61–3.45)†	<0.0001†	2.66 (1.70–4.17)†	<0.0001†	2.22 (1.13–4.39)†	0.02†
Pain score (0–10)‡	1.16 (1.01–1.34)†	0.04†	1.12 (0.98–1.30)	0.09	1.26 (1.05–1.51)†	0.01†
PtGA arthritis (0–10)‡	1.13 (1.00–1.30)	0.06	1.11 (0.97–1.27)	0.11	1.22 (1.03–1.46)†	0.02†
MDA state‡	0.37 (0.15–0.90)†	0.03†	0.41 (0.17–0.99)†	0.047†	0.24 (0.06–0.87)†	0.03†
AM time in MDA state#	0.37 (0.14–0.97)†	0.04†	0.34 (0.12–0.97)†	0.04†	0.29 (0.07–1.15)	0.08

* 95% CI = 95% confidence interval; AM = adjusted mean; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; HAQ = Health Assessment Questionnaire; HR = hazard ratio; MDA = minimal disease activity; PASI = Psoriasis Area and Severity Index; PtGA = patient global assessment; SJC = swollen joint count; TJC = tender joint count.
† Statistically significant.
‡ Time-varying covariate.
§ 10 units increase; time-varying covariate.
¶ 10 units increase.
# Time-dependent mean variable.

factors, including diabetes mellitus, hypertension, and hypercholesteremia, as well as a history of ischemic heart disease, were associated with all heart failure events (Table 3). In addition, the following measures of psoriatic disease activity were associated with all heart failure events: TJC, adjusted mean TJC, SCJ, adjusted mean SJC, adjusted mean PASI, ESR, adjusted mean ESR, adjusted mean CRP level, HAQ score, and pain scores. Patient global assessment was an additional risk factor of nonischemic heart failure. Achieving MDA and adjusted mean time spent in MDA state were protective against all heart failure.

Risk factors of heart failure in multivariable analysis. The following variables were independent risk factors for all heart failure events in the multivariable analysis (Table 4): ischemic heart disease (hazard ratio [HR] 5.52 [95% confidence interval (95% CI) 2.97–10.17]), TJC (HR 1.46 [95% CI 1.12–1.90]), adjusted mean TJC (HR 1.50 [95% CI 1.07–2.11]), adjusted mean SJC (HR 1.93 [95% CI 1.02–3.67]), ESR (HR 1.19 [95% CI 1.05–1.35]), adjusted mean ESR (HR 1.27 [95% CI 1.08–1.50]), adjusted mean CRP level (HR 1.28 [95% CI 1.04–1.58]), and HAQ score (HR 1.99 [95% CI 1.35–2.94]), while MDA was protective for all heart failure (HR 0.39 [95% CI 0.17–0.89]). When the analysis was restricted to nonischemic heart failure, the adjusted effect size of disease activity measures numerically increased (Table 4). The following variables were risk factors of nonischemic heart failure after controlling for traditional heart failure risk factors: pain score, TJC, adjusted mean TJC, adjusted mean SJC, PASI score, ESR, and adjusted mean ESR. As expected, the strongest risk factor of ischemic heart failure was prior ischemic heart disease (HR 19.29 [95% CI 7.84–47.94]), and additional independent risk factors included adjusted mean ESR, adjusted mean CRP level, and HAQ score.

ECG and echocardiographic findings. A total of 41 of the 64 patients had available clinical information on the cause of heart failure events. The underlying etiologies of nonischemic heart failure (n = 26) were nonischemic dilated cardiomyopathy (n = 7), infection (n = 4), hypertension (n = 3), arrhythmias (n = 3), valvular abnormalities (n = 2), drug-induced heart failure (n = 1), alcohol-induced cardiomyopathy (n = 1), and chemotherapy

Table 4. Regression estimates from fitting multivariable Cox proportional hazards models with age as the time scale (n = 1,991 [n = 64 events])*

Variable	All heart failure (n = 64 events)†		Ischemic heart failure (n = 38 events)‡		Nonischemic heart failure (n = 26 events)§	
	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*
TJC (0–68)¶	1.46 (1.12–1.90)#	0.005#	1.35 (0.87–2.07)	0.17	1.62 (1.13–2.29)#	0.008#
AM TJC (0–68)**	1.50 (1.07–2.11)#	0.02#	1.13 (0.64–1.99)	0.67	1.80 (1.20–2.75)#	0.005#
SJC (0–66)¶	1.62 (0.90–2.91)	0.11	1.52 (0.58–3.38)	0.31	2.46 (0.99–6.17)	0.05
AM SJC (0–66)**	1.93 (1.02–3.67)#	0.04#	1.36 (0.47–3.93)	0.57	4.35 (2.03–9.30)#	0.0002#
PASI (0–72)¶	1.17 (0.77–1.78)	0.46	0.82 (0.39–1.75)	0.62	1.66 (1.03–2.70)#	0.04#
AM PASI (0–72)**	1.43 (0.92–2.22)	0.11	1.34 (0.69–2.61)	0.39	1.70 (0.96–3.00)	0.07
Damaged joint count¶	1.13 (0.94–1.32)	0.22	1.05 (0.82–1.32)	0.71	1.27 (0.98–1.67)	0.08
ESR, mm/hour¶	1.19 (1.05–1.35)#	0.005#	1.10 (0.90–1.33)	0.35	1.28 (1.09–1.49)#	0.002#
AM ESR, mm/hour**	1.27 (1.08–1.50)#	0.004#	1.28 (1.01–1.63)#	0.046#	1.27 (1.01–1.60)#	0.049#
CRP level, mg/liter¶	1.12 (0.97–1.29)	0.14	1.08 (0.84–1.47)	0.52	1.14 (0.89–1.46)	0.31
AM CRP level, mg/liter**	1.28 (1.04–1.58)#	0.02#	1.34 (1.01–1.87)#	0.04#	1.26 (0.91–1.71)	0.15
HAQ (0–3)††	1.99 (1.35–2.94)#	0.0005#	2.18 (1.26–3.78)#	0.006#	2.12 (0.97–4.66)	0.06
Pain score (0–10)††	1.14 (0.99–1.31)	0.07	1.08 (0.94–1.28)	0.32	1.23 (1.01–1.52)#	0.04#
PtGA arthritis (0–10)††	1.12 (0.99–1.26)	0.07	1.06 (0.93–1.22)	0.37	1.20 (0.99–1.46)	0.07
MDA state††	0.39 (0.17–0.89)#	0.03#	0.42 (0.17–1.04)	0.06	0.26 (0.06–1.07)	0.06
% of time in MDA state‡‡	0.44 (0.17–1.15)	0.10	0.46 (0.12–1.68)	0.24	0.34 (0.07–1.71)	0.19

* 95% CI = 95% confidence interval; AM = adjusted mean; BMI = body mass index; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; HAQ = Health Assessment Questionnaire; HR = hazard ratio; MDA = minimal disease activity; PASI = Psoriasis Area and Severity Index; PtGA = patient global assessment; SJC = swollen joint count; TJC = tender joint count.
† Adjusted for sex, hypertension, diabetes mellitus, body mass index (BMI), ischemic heart disease (IHD), cholesterol, triglycerides, smoking, and decade.
‡ Adjusted for sex, hypertension, diabetes mellitus, IHD, smoking, cholesterol, triglycerides, and decade.
§ Adjusted for sex, hypertension, diabetes mellitus, BMI, and decade.
¶ 10 units increase; time-varying covariate.
# Statistically significant in multivariable analysis.
** 10 units increase. Time-dependent mean variable.
†† Time-varying covariate.
‡‡ Time-dependent mean variable.

(n = 1). There were 4 patients without a documented etiology for nonischemic heart failure. Seventeen patients of the 64 had a revascularization procedure with coronary artery bypass or percutaneous coronary intervention. Three patients died due to their heart failure event. The most common ECG findings were pathologic Q wave (infarcts), atrial fibrillation, and bundle branch blocks (Table 5). TTE revealed preserved ejection fraction in 63% of the patients and systolic dysfunction (reduced ejection fraction) in 37% (Table 5).

DISCUSSION

The association between psoriatic disease and cardiovascular diseases is well established. However, the literature regarding psoriatic disease and heart failure is limited. In this study we investigated the incidence and risk factors of heart failure events in a large, well phenotyped cohort of patients with psoriatic disease over a period of 40 years. We found a higher incidence of heart failure events in our population (2.85 per 1,000 patient-years) compared to the general population in the same province (1.93 per 1,000 patient-years) as reported in a recent study (24). Although the risk of developing heart failure was explained in part by traditional cardiovascular risk factors and prior ischemic heart disease, the burden of systemic inflammation and the level of disease activity were independent risk factors of heart failure events. Few population-based studies have shown that the risk of developing heart failure events in patients with psoriatic disease is higher than the general population (8,11). A meta-analysis found that patients with PsA have a 32% increased risk of developing heart failure compared to the general population (8). Patients with

Table 5. Electrocardiographic (ECG) and echocardiographic findings in psoriatic disease patients with heart failure events

ECG variable (n = 49)	Percentage of patients	Echocardiographic variable (n = 41)	Percentage of patients
Pathologic Q waves (old infarcts)	33	Reduced ejection fraction	37
Bundle branch blocks	29	Preserved ejection fraction	63
Atrial fibrillation	22	Wall motion abnormalities	61
Left ventricular hypertrophy	12	Left ventricular hypertrophy	41
Atrioventricular block	12	Valvular abnormalities	32

psoriasis have a similarly increased risk between 20% and 60% (11,12). Generally, the risk of heart failure was increased with the severity and duration of psoriasis (11,12). To date, no such risk factors have been identified in PsA. We also found that the rise in heart failure events in women lagged by a decade compared to men in our cohort, especially in ischemic heart failure. This finding is interesting because women in the general population also develop heart failure and coronary artery disease approximately 5–10 years later than men (25,26). Furthermore, women are more likely to have heart failure with preserved ejection fraction compared to heart failure with reduced ejection fraction, which may explain why we did not see sex differences in the onset of nonischemic heart failure (25–28).

We identified the fact that measures of psoriatic disease activity, both at a single time point and cumulative levels over time, were associated with the occurrence of heart failure events independent of traditional cardiovascular risk factors. Moreover, a composite outcome measure of PsA activity, MDA state, was protective from heart failure events. These findings are in line with previous observations demonstrating that exposure to an increased burden of inflammation over time is associated with more severe atherosclerosis in patients with PsA (29,30). Biomarkers of systemic inflammation predict the development of cardiovascular events in the general population and in patients with rheumatoid arthritis (RA) (31–33). Patients with psoriatic disease generally have lower levels of inflammatory markers compared with patients with RA. However, in our study, higher levels of both ESR and CRP were independently associated with the development of heart failure events. The severity of psoriasis, a previously reported predictor of cardiovascular events (34,35), was associated with nonischemic heart failure in our study.

Interestingly, we identified the fact that the strength of association of disease activity measures tended to be higher when the analysis was restricted to nonischemic heart failure compared to ischemic heart failure. This finding suggests that the effect of systemic inflammation on heart failure may be partially independent of atherosclerotic disease. Recently, the role of inflammation in the development and progression of heart failure has become clearer. Heart failure development, especially with preserved ejection fraction, has been attributed to a sustained proinflammatory cytokine signaling and elevated levels of tumor necrosis factor (TNF), interferon gamma, interleukin (IL)-1β, IL-6, IL-17, and IL-18, resulting in coronary microvascular endothelial inflammation. This proinflammatory milieu reduces nitric oxide bioavailability, cyclic guanosine monophosphate content, and protein kinase G activity in adjacent cardiomyocytes, which ultimately favors the development of hypertrophy and increases interstitial fibrosis, contributing to ventricular stiffness, functional alterations, and heart failure development (36,37).

The concept of "epicardial adipose inflammation" suggests that chronic systemic inflammation activates epicardial fat to produce adipocytokines (leptin, TNF, IL-1β, IL-6) that are transmitted directly to underlying tissues, leading to volume stress, fibrosis, and decreased chamber accommodation (16,17). Furthermore, the dysfunctional epicardial adipose tissue may abut the left ventricle, resulting in ventricular myopathy, characterized by impaired cardiac distensibility and leading to heart failure with preserved ejection fraction. Therefore, a systemic inflammatory state in patients with psoriatic disease conceivably may confer an increased risk of heart failure that is independent of traditional risk factors.

Our descriptive analysis revealed that the most common ECG findings in patients with heart failure events are the presence of arrhythmias, left ventricular hypertrophy, and previous Q wave infarcts. These findings are in line with studies that have reported ECG changes in patients with psoriatic disease (38–40), but these were not restricted to patients with heart failure. In our cohort, 63% of patients with heart failure events had preserved ejection fraction on TTE. This finding is slightly higher than previously reported values, which range from 36% to 60% (15,41,42), but all studies were limited by a small study size. Several studies have also shown that the presence of arthritis and the duration of psoriasis were significantly correlated with the presence of impaired diastolic filling, subclinical myocardial deformation, and minimal segmental ischemic injury, all hallmarks of heart failure with preserved ejection fraction (14,41,43).

Advanced echocardiographic studies, which can detect subclinical myocardial disease otherwise not detectable by conventional TTE, have demonstrated that PsA patients without clinically evident cardiovascular disease or classical risk factors have a higher prevalence of subclinical left ventricular dysfunction compared to controls (15). Patients with PsA had multilayer myocardial involvement, including thickened posterior wall, increased relative wall thickness, and a higher prevalence of concentric remodeling, which likely occurs due to a different pathologic mechanism from ischemic heart disease in the early stages (14,15,43,44). These findings provide further support to the concept of sustained chronic inflammation from psoriatic disease as an independent risk factor in the development of heart failure.

Our study had several potential limitations. This is a single-center study with a relatively small number of events, which may have limited our ability to estimate the true effect size of several disease-related variables and identify further differences between ischemic and nonischemic heart failure. Another limitation is the fact that this is not an inception cohort, and thus we could not account for disease activity prior to the first visit to the clinic. The strength of this study is the large sample of patients, with comprehensive and accurate phenotyping of patients that allowed an estimation of the inflammatory burden of disease over time. Furthermore, we have complete capture of data due to linkage with administrative databases and event confirmation through chart reviews.

In summary, we have found that a significant proportion of patients with psoriatic disease develop a heart failure event at

some point in the course of their disease. Among patients with psoriatic disease, an increased risk of heart failure is associated with a combination of traditional cardiovascular risk factors, prior ischemic heart disease, and disease activity. These results support the proposition of the independent effects of psoriatic disease on the risk of developing heart failure, especially nonischemic heart failure, mediated through inflammatory mechanisms. These results also highlight the importance of controlling all traditional cardiovascular risk factors as well as targeting for MDA, which can potentially reduce the risk of heart failure in psoriatic 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. Eder 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. Koppikar, Harvey, Akhtari, Gladman, Eder.

Acquisition of data. Koppikar, Colaco, Chandran, Gladman, Eder.

Analysis and interpretation of data. Koppikar, Chandran, Gladman, Cook, Eder.

REFERENCES

1. Kurd SK, Gelfand JM. The prevalence of previously diagnosed and undiagnosed psoriasis in US adults: results from NHANES 2003-2004. J Am Acad Dermatol 2009;60:218–24.
2. Alinaghi F, Calov M, Kristensen LE, Gladman DD, Coates LC, Jullien D, et al. Prevalence of psoriatic arthritis in patients with psoriasis: a systematic review and meta-analysis of observational and clinical studies. J Am Acad Dermatol 2019;80:251–65.
3. Ritchlin CT, Colbert RA, Gladman DD. Psoriatic arthritis. N Engl J Med 2017;376:957–70.
4. Takeshita J, Grewal S, Langan SM, Mehta NN, Ogdie A, Van Voorhees AS, et al. Psoriasis and comorbid diseases: epidemiology. J Am Acad Dermatol 2017;76:377–90.
5. Abuabara K, Azfar RS, Shin DB, Neimann AL, Troxel AB, Gelfand JM. Cause-specific mortality in patients with severe psoriasis: a population-based cohort study in the U.K. Br J Dermatol 2010;163:586–92.
6. Ogdie A, Yu Y, Haynes K, Love TJ, Maliha S, Jiang Y, et al. Risk of major cardiovascular events in patients with psoriatic arthritis, psoriasis and rheumatoid arthritis: a population-based cohort study. Ann Rheum Dis 2015;74:326–32.
7. Husted JA, Thavaneswaran A, Chandran V, Eder L, Rosen CF, Cook RJ, et al. Cardiovascular and other comorbidities in patients with psoriatic arthritis: a comparison with patients with psoriasis. Arthritis Care Res (Hoboken) 2011;63:1729–35.
8. Polachek A, Touma Z, Anderson M, Eder L. Risk of cardiovascular morbidity in patients with psoriatic arthritis: a meta-analysis of observational studies. Arthritis Care Res (Hoboken) 2017;69:67–74.
9. Eder L, Wu Y, Chandran V, Cook R, Gladman DD. Incidence and predictors for cardiovascular events in patients with psoriatic arthritis. Ann Rheum Dis 2016;75:1680–6.
10. Miller IM, Skaaby T, Ellervik C, Jemec GB. Quantifying cardiovascular disease risk factors in patients with psoriasis: a meta-analysis. Br J Dermatol 2013;169:1180–7.
11. Khalid U, Ahlehoff O, Gislason GH, Kristensen SL, Skov L, Torp-Pedersen C, et al. Psoriasis and risk of heart failure: a nationwide cohort study. Eur J Heart Fail 2014;16:743–8.
12. Yang YW, Keller JJ, Lin HC. Medical comorbidity associated with psoriasis in adults: a population-based study. Br J Dermatol 2011;165:1037–43.
13. Eliakim-Raz N, Shuvy M, Lotan C, Planer D. Psoriasis and dilated cardiomyopathy: coincidence or associated diseases? Cardiology 2008;111:202–6.
14. Shang Q, Tam LS, Sanderson JE, Lee AP, Li EK, Yu CM. Impaired left ventricular apical rotation is associated with disease activity of psoriatic arthritis. J Rheumatol 2014;41:706–13.
15. Shang Q, Tam LS, Yip GW, Sanderson JE, Zhang Q, Li EK, et al. High prevalence of subclinical left ventricular dysfunction in patients with psoriatic arthritis. J Rheumatol 2011;38:1363–70.
16. Packer M. Epicardial adipose tissue inflammation can cause the distinctive pattern of cardiovascular disorders seen in psoriasis. Am J Med 2020;133:267–72.
17. Packer M. The epicardial adipose inflammatory triad: coronary atherosclerosis, atrial fibrillation, and heart failure with a preserved ejection fraction. Eur J Heart Fail 2018;20:1567–9.
18. Eisen A, Arnson Y, Dovrish Z, Hadary R, Amital H. Arrhythmias and conduction defects in rheumatological diseases: a comprehensive review. Semin Arthritis Rheum 2009;39:145–56.
19. Liew JW, Ramiro S, Gensler LS. Cardiovascular morbidity and mortality in ankylosing spondylitis and psoriatic arthritis. Best Pract Res Clin Rheumatol 2018;32:369–89.
20. Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P, Mielants H, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006;54:2665–73.
21. Gladman DD, Shuckett R, Russell ML, Thorne JC, Schachter RK. Psoriatic arthritis (PSA): an analysis of 220 patients. Q J Med 1987;62:127–41.
22. Coates LC, Fransen J, Helliwell PS. Defining minimal disease activity in psoriatic arthritis: a proposed objective target for treatment. Ann Rheum Dis 2010;69:48–53.
23. Lawless J. Truncation and selection effect. In: Lawless J, editor. Statistical models and methods for lifetime data 362. 2nd ed. Boston: John Wiley; 2011. p. 67–70.
24. Di Giuseppe G, Chu A, Tu JV, Shanmugasegaram S, Liu P, Lee DS. Incidence of heart failure among immigrants to Ontario, Canada: a CANHEART immigrant study. J Card Fail 2019;25:425–35.
25. Pandey A, Omar W, Ayers C, LaMonte M, Klein L, Allen NB, et al. Sex and race differences in lifetime risk of heart failure with preserved ejection fraction and heart failure with reduced ejection fraction. Circulation. 2018;137:1814–23.
26. Lam CS, Arnott C, Beale AL, Chandramouli C, Hilfiker-Kleiner D, Kaye DM, et al. Sex differences in heart failure. Eur Heart J 2019;0:1–13.
27. Tromp J, Shen L, Jhund PS, Anand IS, Carson PE, Desai AS, et al. Age-related characteristics and outcomes of patients with heart failure with preserved ejection fraction. J Am Coll Cardiol 2019;74:601–12.
28. Goyal P, Paul T, Almarzooq ZI, Peterson JC, Krishnan U, Swaminathan RV, et al. Sex- and race-related differences in characteristics and outcomes of hospitalizations for heart failure with preserved ejection fraction. J Am Heart Assoc 2017;6:1–10.
29. Eder L, Thavaneswaran A, Chandran V, Cook R, Gladman DD. Increased burden of inflammation over time is associated with the extent of atherosclerotic plaques in patients with psoriatic arthritis. Ann Rheum Dis 2015;74:1830–5.

30. Egeberg A, Skov L, Hansen PR, Gislason GH, Wu JJ, Thyssen JP, et al. Duration of psoriatic arthritis as a risk factor for myocardial infarction. Rheumatol Adv Pract 2018;2:rky011.
31. Arts EE, Fransen J, den Broeder AA, Popa CD, van Riel PL. The effect of disease duration and disease activity on the risk of cardiovascular disease in rheumatoid arthritis patients. Ann Rheum Dis 2015;74: 998–1003.
32. Gonzalez-Gay MA, Gonzalez-Juanatey C, Lopez-Diaz MJ, Piñeiro A, Garcia-Porrua C, Miranda-Filloy JA, et al. HLA-DRB1 and persistent chronic inflammation contribute to cardiovascular events and cardiovascular mortality in patients with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2007;57:125–32.
33. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. N Engl J Med 2005;352:1685–95.
34. Gelfand JM, Neimann A, Shin D, Wang X, Margolis D, Troxel AB. Risk of myocardial infarction in patients with psoriasis. JAMA 2006;296: 1735–41.
35. Ahlehoff O, Gislason GH, Charlot M, Jorgensen CH, Lindhardsen J, Olesen JB, et al. Psoriasis is associated with clinically significant cardiovascular risk: a Danish nationwide cohort study. J Intern Med 2011;270:147–57.
36. Paulus WJ, Tschope C. A novel paradigm for heart failure with preserved ejection fraction: comorbidities drive myocardial dysfunction and remodeling through coronary microvascular endothelial inflammation. J Am Coll Cardiol 2013;62:263–71.
37. Shirazi LF, Bissett J, Romeo F, Mehta JL. Role of inflammation in heart failure. Curr Atheroscler Rep 2017;19:27.
38. Gaydukova I, Rebrov A, Nikitina N, Poddubnyy D. Decreased heart rate variability in patients with psoriatic arthritis. Clin Rheumatol 2012;31:1377–81.
39. Rhee TM, Lee JH, Choi EK, Han KD, Lee H, Park CS, et al. Increased risk of atrial fibrillation and thromboembolism in patients with severe psoriasis: a nationwide population-based study. Sci Rep 2017;7: 9973.
40. Chiu HY, Chang WL, Huang WF, Wen YW, Tsai YW, Tsai TF. Increased risk of arrhythmia in patients with psoriatic disease: a nationwide population-based matched cohort study. J Am Acad Dermatol 2015;73:429–38.
41. Saricaoglu H, Gullulu S, Bulbul Baskan E, Cordan J, Tunali S. Echocardiographic findings in subjects with psoriatic arthropathy. J Eur Acad Dermatol Venereol 2003;17:414–7.
42. Gorga E, Scodro M, Valentini F, D'Ortona R, Arisi M, Sciatti E, et al. Echocardiographic evaluation of diastolic dysfunction in young and healthy patients with psoriasis: a case-control study. Monaldi Arch Chest Dis 2018;88:934.
43. Bulbul Sen B, Ekiz O, Rifaioglu EN, Buyukkaya E, Karakas MF, Buyukkaya S, et al. Assessment of subclinical left ventricular dysfunction in patients with psoriasis by speckle tracking echocardiography: a speckle tracking study. Int J Dermatol 2016;55:158–64.
44. Ozden HK, Polat M, Ozturk S, Bugdayci G. Assessment of subclinical cardiac damage in chronic plaque psoriasis patients: a case control study. Arch Med Sci Atheroscler Dis 2016;6:126–32.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1254–1262
DOI 10.1002/acr.24570

Comparative Efficacy of Different Triage Methods for Psoriatic Arthritis: Results From a Prospective Study in a Rapid Access Clinic

Samantha Sarabia,[1] Chandra Farrer,[2] Jensen Yeung,[2] Dana Jerome,[2] Richard J. Cook,[3] and Lihi Eder[2]

Objective. We undertook this study to identify the optimal combination of triage methods to identify psoriatic arthritis (PsA) among psoriasis patients with musculoskeletal symptoms in a rapid access clinic and to describe their outcome after 1 year.

Methods. Patients with psoriasis and no prior diagnosis of PsA were referred for assessment of their musculoskeletal symptoms. Each patient was assessed by the following 3 triage modalities: 1) assessment by an advanced practice physical therapist; 2) targeted musculoskeletal ultrasound (MSK-US); and 3) PsA screening questionnaires. The patients were then evaluated by a rheumatologist who determined the patient's disease status and classified them into the following groups: not PsA, possibly PsA, or PsA. Patients returned for a 1-year follow-up visit and were reassessed for change in their disease status. Sensitivity and specificity were calculated for each individual modality, as well as for combinations of modalities.

Results. A total of 203 patients with psoriasis and musculoskeletal symptoms were enrolled. The percentage of patients classified as having PsA was 8.8%, and 23.6% were converted into the possibly PsA group. There was no significant difference in the individual performance of the modalities. The highest sensitivity was seen with MSK-US (89%), and the highest specificity was found with the Psoriatic Arthritis Screening and Evaluation questionnaire (79%). The addition of MSK-US data improved the performance of the modalities. A total of 9 patients were classified into the PsA group after 1 year. All patient-reported outcome measures had significantly improved at 1 year ($P < 0.001$).

Conclusion. Combining MSK-US with a screening questionnaire for PsA improved the triage of patients with suspected PsA.

INTRODUCTION

Psoriatic arthritis (PsA) is an inflammatory musculoskeletal condition that affects up to one-third of people with psoriasis (1). While earlier diagnosis of PsA has been shown to result in better long-term outcomes, delayed diagnosis remains a major problem that contributes to poor patient outcomes (2,3). A number of factors can contribute to this delay in diagnosis, including wait time for access to rheumatology care, coexistence of noninflammatory musculoskeletal diseases (e.g., osteoarthritis), presentation with nonspecific musculoskeletal symptoms, and lack of objective laboratory tests to aid in diagnosis in the early stages of the disease (3).

The limited accessibility to rheumatology care does not permit timely assessment of each psoriasis patient with musculoskeletal symptoms. To date, efforts to improve early diagnosis of PsA have concentrated primarily on developing methods for early detection of potential PsA cases in psoriasis patients. Since a considerable proportion of the psoriasis patients with musculoskeletal symptoms do not have PsA, but rather have other noninflammatory rheumatic conditions, such as osteoarthritis (4), these methods have been developed to help dermatologists and family

The Early Arthritis Clinic was supported by unrestricted educational grants from UCB, Celgene, Novartis, and Janssen. Ms. Sarabia's work was supported by a summer studentship award from the National Psoriasis Foundation. Dr. Eder's work was supported by a Young Investigator Award from the Arthritis Society.

[1]Samantha Sarabia, BSc: Queen's University Medical School, Kingston, Ontario, Canada; [2]Chandra Farrer, BScPT, Jensen Yeung, MD, Dana Jerome, MD, Lihi Eder, MD, PhD: Women's College Hospital and University of Toronto, Toronto, Ontario, Canada; [3]Richard J. Cook, PhD: University of Waterloo, Waterloo, Ontario, Canada.

Dr. Eder has received unrestricted educational grants from UCB, Celgene, Novartis, and Jansen (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Lihi Eder, MD, PhD, Women's College Research Institute, Room 6326, Women's College Hospital, 76 Grenville Street, Toronto, Ontario, Canada, M5S 1B2. Email: lihi.eder@wchospital.ca.

Submitted for publication August 10, 2021; accepted in revised form February 2, 2021.

SIGNIFICANCE & INNOVATIONS

- Musculoskeletal ultrasound was found to be the most sensitive triage tool, and the Psoriatic Arthritis Screening and Evaluation questionnaire was the most specific.
- Combining triage modalities, such as musculoskeletal ultrasound with a screening questionnaire or physical therapist assessment, improved the identification of patients with psoriatic arthritis among psoriasis patients with musculoskeletal symptoms.

physicians screen these patients and prioritize rheumatology referral of potential PsA cases (5).

In order to streamline the diagnosis of PsA, a new strategy is needed. New models of care and novel diagnostic modalities may be useful in reducing the gap between presentation of symptoms and diagnosis. Previous research has indicated that a central triage system might be beneficial in reducing wait times for diagnoses in patients with rheumatologic conditions (6).

The emergence of models of care utilizing the skills of allied health professionals in extended clinical roles to manage arthritis care has evolved partly as a result of limited access to rheumatology care (7). This strategy has been utilized to improve timely access to care by performing early triage of patients with suspected inflammatory arthritis (8). Such models of care, in which individuals were initially assessed by an allied health professional before their appointment with rheumatologists, found high levels of agreement in classifying patients as having inflammatory arthritis between the allied health professional and the rheumatologist (9,10). To the best of our knowledge, the efficacy of such a central triage model in psoriasis patients with musculoskeletal symptoms has not been evaluated to date.

Another potential diagnostic and triage tool for PsA is musculoskeletal ultrasound (MSK-US). MSK-US is becoming an increasingly common tool used to aid in the diagnosis of patients with rheumatic conditions as it becomes more portable and less expensive. MSK-US has been proven reliable in detecting a number of inflammatory lesions such as synovitis, tenosynovitis, and enthesitis (11,12). In addition, it has been shown to correlate well with magnetic resonance imaging findings and to be more reliable than the physical examination in the assessment of musculoskeletal symptoms (13,14). We previously suggested that MSK-US may provide additional valuable information to currently used modalities to aid in diagnosing PsA at an earlier stage (15). This highlights the potential use of MSK-US as point of care for patients at very early stages of their disease; however, the specific role of MSK-US as a triage tool for patients with suspected PsA has not been widely assessed.

Several screening questionnaires may also be valuable in assessing patients who are suspected to have PsA. Several questionnaire-style screening tools have been developed so that non-rheumatologists can identify psoriasis patients who might have PsA and refer them to a rheumatologist for definite diagnosis (16–19). Several screening questionnaires have undergone validation; however, despite acceptable reliability in the development phases, the performance of these questionnaires in subsequent independent study populations was only moderate (19). In addition, their use in conjunction with other tools, such as MSK-US has not been evaluated.

The primary aim of this proof-of-concept study is to assess the efficacy of the following modalities alone, or in combination, to identify PsA among psoriasis patients with musculoskeletal symptoms referred to a rapid access clinic: 1) PsA screening questionnaires, 2) clinical assessment by an advanced practice physical therapist (APP), and 3) targeted MSK-US. A second study aim was to describe the outcome of study patients after 1 year and to assess whether any of these modalities predicted the development of PsA. Overall, we aimed to develop a streamlined approach to PsA diagnosis, which is expected to reduce wait times, improve long-term outcomes, and reduce the burden on the health care system.

PATIENTS AND METHODS

Setting and study population. This is a single-center, prospective cohort study conducted in an academic medical center in Toronto, Canada. The study included patients referred to a rapid access clinic for suspected PsA. This clinic was developed in collaboration with the dermatology and family medicine departments at Women's College Hospital in Toronto to enhance earlier diagnosis of PsA among psoriasis patients by facilitating rapid access to rheumatology consultation. In Canada, patients require physician referral to access rheumatology services. The clinic typically receives referrals from dermatologists and family physicians for psoriasis patients suspected of having PsA. These referrals are sorted by the level of suspicion of PsA based on the clinical information provided in the referral letter, and patients are typically seen in the clinic within 4–8 weeks. This proof-of-concept study evaluated the potential efficacy of a novel triage method considering a novel method of self-referral of psoriasis patients in addition to the conventional referral method by a physician.

Patients with psoriasis who were experiencing musculoskeletal symptoms and did not have a prior diagnosis of PsA were enrolled. Potential participants were recruited from the dermatology clinics and the Phototherapy Education and Research Centre, as well as the family medicine clinics at Women's College Hospital in Toronto. Patient recruitment for this study consisted of both self-referral and direct-referral systems. In the self-referral system, patients with a diagnosis of psoriasis who had visited 1 of the participating clinics between February 2015 and December 2017 were identified through a search in the clinic's electronic medical records. These patients were contacted via mail invitation to participate in the study if they were currently experiencing

musculoskeletal pain. Patients were asked to respond by mail or through an online form and identify whether 1) they were experiencing back, joint, or tendon symptoms and 2) they were willing to participate in the study. Only patients who responded "yes" to both of these questions were invited to participate. In addition, participating dermatologists and family physicians had the option to directly refer patients who had psoriasis and were experiencing musculoskeletal symptoms. All participating patients were evaluated by a rheumatologist (see detailed description below) and those classified as not having PsA or possibly having PsA at baseline were reassessed after 1 year to determine whether there had been a change in their PsA status. The study was approved by the Women's College Hospital Ethics Board, and all patients gave their informed consent (REB #2016-0043).

Data collected. At the initial visit, all patients were assessed in the rapid access clinic to determine whether they had PsA. The study assessed the performance of a novel central triage system as part of the clinic and included the following modalities: 1) assessment by an APP, 2) targeted MSK-US assessment, and 3) three screening questionnaires for PsA. The participants were classified by each modality as high likelihood (HL) PsA or low likelihood (LL) PsA. The HL and LL classifications were given because these modalities cannot provide a firm diagnosis, unlike assessment by a rheumatologist.

Following this assessment, each participant was evaluated by a rheumatologist, irrespective of the results of the central triage methods, to determine whether the participant had PsA. The rheumatologist and APP examinations were performed independently of each other and without knowledge of MSK-US and screening questionnaires results. The physical therapist was also blinded to the radiographic and laboratory test results. The assessment included taking the medical history and performing a musculoskeletal examination of 68 and 66 joints for tenderness and swelling, respectively. Additionally, enthesitis was assessed according to the Spondyloarthritis Research Consortium of Canada enthesitis index (20), and the presence and number of dactylitic digits were recorded. In addition, the extent of psoriasis was assessed by the Psoriasis Area and Severity Index, and the presence of psoriatic nail lesions was recorded. The following laboratory tests were performed: C-reactive protein level, erythrocyte sedimentation rate, and rheumatoid factor. A standard set of radiographs was obtained for each participant, including the hands, feet, and whole spine (including the sacroiliac joints), and the results were reported by the hospital's musculoskeletal radiologists. The APP (CF) had previously completed the Advanced Clinician Practitioner in Arthritis Care program and had 10 years of experience in working with rheumatic patients. She was also specifically trained to perform the musculoskeletal assessments required for this study. The rheumatologist classified each patient at the end of the visit to one of the following 3 categories: 1) PsA, 2) not PsA, or 3) possibly PsA. The latter category included patients for whom the diagnosis was suspected based on typical symptoms (e.g., prolonged morning stiffness, history of joint swelling) but could not be confirmed after completing the clinical assessment and reviewing the results of the laboratory tests and radiographs.

Patients were asked to complete the following 3 screening questionnaires for PsA: the Psoriatic Arthritis Screening and Evaluation (PASE), the Psoriasis Epidemiology Screening Tool (PEST), and the Toronto Psoriatic Arthritis Screen 2 (ToPAS-2) (16,17). These are validated questionnaires developed as screening tools for PsA, consisting of questions on particular features typical of PsA. The suggested cutoff levels for each questionnaire were used to classify patients into LL-PsA or HL-PsA groups. Patient-reported outcomes were also collected (at baseline and 1 year) using the following questionnaires: level of pain (visual analog scale [VAS] scored 0–10), the Dermatology Life Quality Index (DLQI), the Functional Assessment of Chronic Illness Therapy (FACIT), the Health Assessment Questionnaire (HAQ), and the patient global assessment (PtGA) of arthritis (VAS 0–10).

The 1-year follow-up visit included only a rheumatologist assessment that was conducted in a similar fashion to the first visit. Patients were again converted into the following groups: 1) PsA, 2) not PsA, or 3) possibly PsA. The rheumatologist was blinded to the MSK-US data from the first visit. The participants were again asked to complete patient-reported outcome questionnaires. Patients who declined to return for a follow-up visit were allowed to complete the questionnaires over the phone.

Ultrasound assessment. All the ultrasound assessments were performed by a single rheumatologist (LE) who has 8 years of experience in MSK-US. The intrarater reliability of this sonographer for evaluating musculoskeletal inflammation in psoriasis patients has been shown in the past (21). A MyLab Twice scanner (Esaote), equipped with a 6–18 MHz linear transducer was the ultrasound device. A targeted examination of the peripheral joints was performed in addition to a standardized assessment of 14 entheses. To determine which peripheral joints to assess, the patients were first assessed by the APP who identified tender and or swollen joints for a targeted MSK-US assessment.

The sonographer scanned these clinically affected joints, as well as the patient's contralateral side, and was blinded to the affected side. The peripheral joints were assessed for the presence of the following lesions: 1) synovitis: defined as synovial hypertrophy in grayscale (GS) and intraarticular power Doppler (PD) (22); 2) peritenonitis: defined as peritendon swelling (GS) and positive peritendinous (PD) in the extensor tendons in the hands and feet; and 3) tenosynovitis, which is synovial inflammation in tendons with tendon sheet (GS and PD) (23). In addition, enthesitis was assessed in the following 7 entheseal sites bilaterally: quadriceps tendons insertions to the patella, patellar tendons insertions to the patella and tibial tuberosity, Achilles tendons and

plantar fascia insertions into the calcaneus, triceps tendon insertions to the olecranon process, and common extensor tendon insertion to the lateral epicondyle. The presence of GS and PD entheseal lesions were assessed according to the Outcome Measures in Rheumatology definition (24). We considered the presence of both GS and PD abnormalities in the joint, peritendon, or tendon sheath an indication of sonographic synovitis, peritenonitis, and tenosynovitis, respectively. The presence of GS lesions, including hypoechogenicity and/or entheseal thickening and at least grade 2 PD was considered as sonographic enthesitis. Active sonographic musculoskeletal inflammation was defined as evidence of synovitis, peritenonitis, tenosynovitis, or enthesitis by MSK-US.

Two separate ultrasound definitions for active sonographic inflammation were considered for the analysis. Rule 1 defines a positive ultrasound as at least 1 site (joint, tendon, enthesis) with active sonographic musculoskeletal inflammation. Rule 2 defines a positive ultrasound as at least 2 sites with active sonographic MSK inflammation.

Statistical analysis. Descriptive statistics included mean and SDs for continuous variables and frequencies (%) for categorical variables. For the primary analysis, patients with PsA were compared to those with no PsA and possibly PsA (combined). Since the continuous variables were not normally distributed (by observing the density distribution histogram), we performed the nonparametric test, the Wilcoxon test, to compare continuous variables. Fisher's exact test was used to compare categorical variables between the groups, respectively.

Sensitivity, specificity, and area under the curve (AUC) were calculated for individual modality, with the rheumatologist assessment as the gold standard. Subsequently, we evaluated the performance of different combinations of modalities with MSK-US using the above-described metrics.

RESULTS

Patient enrollment. A total of 1,692 psoriasis patients received a mailed invitation to participate in the study (1,646 [97.2%] from dermatology clinics and 48 (2.8%) from family medicine) (Figure 1). In addition, 71 and 2 patients were directly referred from dermatology and family medicine, respectively. A total of 203 patients participated in the study, of whom 135 (66.5%) were enrolled as a result of self-referrals, and 68 (33.5%) were enrolled as a result of direct physician referrals. Of the 135 patients from self-referrals, 113 (83.7%) were identified through dermatology records and 22 (16.3%) through family medicine records. Of the 68 patients from direct physician referrals, 66 (97.1%) were referred by dermatologists and 2 (2.9%) were referred by family physicians.

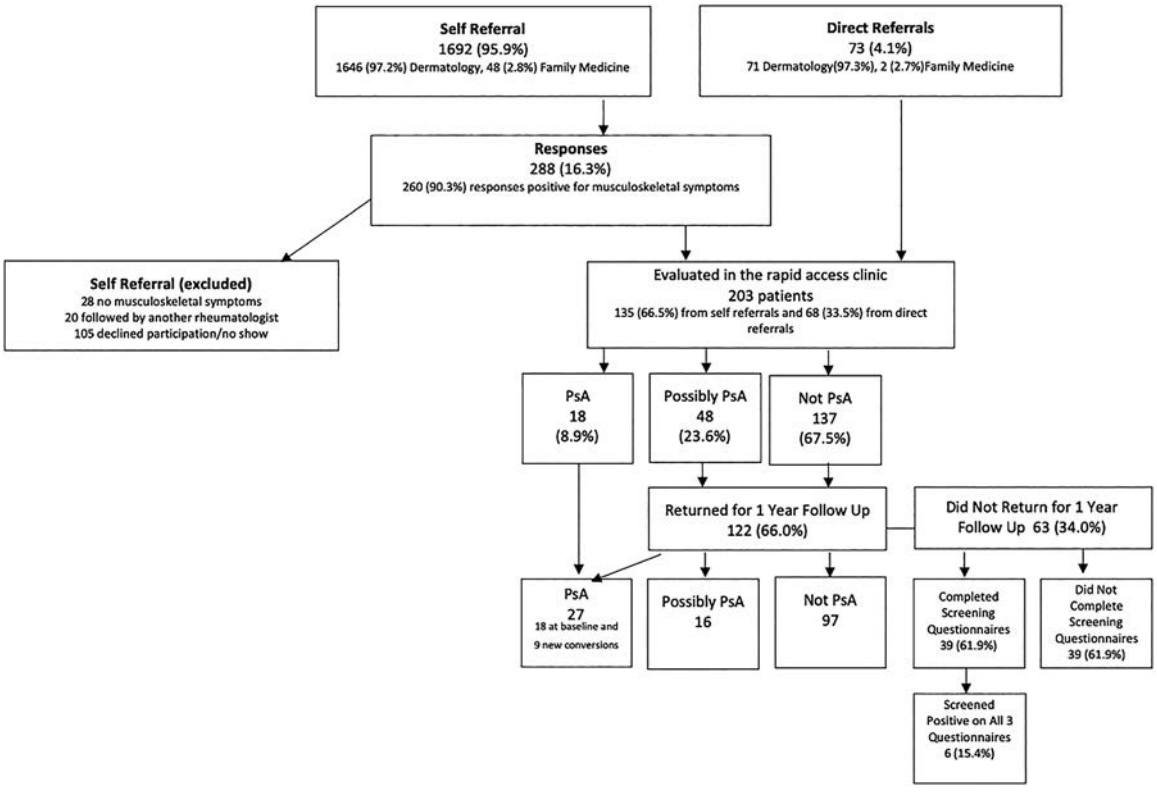


Figure 1. Patient recruitment and change in psoriatic arthritis (PsA) status from baseline to 1 year.

Comparison of patient characteristics by disease status. At baseline, 18 patients (8.9%) were converted into the PsA group, 48 (23.6%) into the possibly PsA group, and 137 (67.5%) into the not PsA group (Figure 1). The characteristics of the study patients are shown in Table 1. Patients in the PsA group were more likely to test positive on the PASE (61.1%), PEST (76.5%), and ToPAS-2 (72.2%) questionnaires. In addition, the PsA group was more likely to have positive MSK-US findings (rule 1: 88.9%; rule 2: 55.6%). All 18 patients who were diagnosed with PsA met the Classification of Psoriatic Arthritis (CASPAR) classification criteria for PsA (25). Of the patients who were diagnosed with PsA at baseline, 9 patients (50%) had polyarthritis, 6 patients (33.3%) had oligoarthritis, and 3 patients (16.6%) had purely axial involvement. One-third of the PsA patients had clinical enthesitis and 27.8% had dactylitis.

Performance of individual triage modalities in identifying PsA at baseline. The performance of each individual triage modality to detect a confirmed diagnosis of PsA is presented in Table 2. Overall, the sensitivities and specificities ranged widely from 56% to 89%, and 44% to 79%, respectively; however, the confidence intervals were wide due to the small number of patients with PsA. The most sensitive modality was MSK-US rule 1 (sensitivity 89%), but the specificity was low (44%). A more stringent definition of positive MSK-US (rule 2) increased the specificity to 77%, but this definition reduced the sensitivity to 56%. Examination by the physical therapist was able to detect PsA with a sensitivity of 83%; however, the specificity was relatively low at 54%. There was no significant difference in sensitivity and specificity among the 3 screening questionnaires, with PEST having the highest nominal sensitivity at 77%, followed by ToPAS-2 at 72% and PASE at 61%. PASE had the highest nominal specificity at 79%, followed by ToPAS-2 at 72% and PEST at 70%. The AUC ranged from 0.63 (physical therapy assessment) to 0.73 (PEST). There was no significant change in the performance of the screening methods when the patients in the possibly PsA group were combined with PsA patients and compared to those in the no PsA group (data not shown).

Performance of combinations of modalities in identifying PsA. The performance of various combinations of triage modalities are shown in Table 3. Using a combination of the MSK-US with other modalities showed a marginal increase in sensitivity and specificity, depending on the MSK-US definition and the requirement of 1 or both modalities to be positive. As expected, sensitivity tended to be highest when any of the

Table 1. Participant characteristics at baseline (n = 203)*

Variable	All	Not PsA (n = 137)	Possibly PsA (n = 48)	PsA (n = 18)	*P* (PsA vs. possibly/ not PsA)
Age, mean ± SD years	50.8 ± 14.4	52.8 ± 14.3	45.6 ± 13.8	49.4 ± 13.6	0.78
Female sex	133 (65.2)	93 (67.4)	31 (64.6)	9 (50)	0.20
Prolonged duration of musculoskeletal symptoms (>2 years)	111 (56.1)	73 (54.9)	27 (57.5)	11 (61.1)	0.80
Duration of psoriasis, mean ± SD years	17.4 ± 15.8	17.7 ± 16.7	18.8 ± 14.9	11.7 ± 10	0.26
Use of systemic nonbiologic medications for psoriasis	7 (3.43)	0	4 (8.3)	3 (16.7)	0.02
Methotrexate	2 (1)	0	1 (2.1)	1 (5.5)	0.17
Apremilast	5 (2.45)	0	3 (6.25)	2 (11.1)	0.06
Use of biologic medication for psoriasis	22 (10.8)	9 (6.5)	7 (14.6)	6 (33)	0.006
TNF inhibitor	8 (3.9)	3 (2.2)	3 (6.25)	2 (11.1)	0.15
IL-12/IL-23 inhibitor	11 (5.4)	5 (3.6)	2 (4.17)	4 (22.2)	0.009
IL-17 inhibitor	2 (1)	1 (0.7)	1 (2.1)	0	–
Severe psoriasis (PASI >10)	35 (17.2)	18 (13)	10 (20.8)	7 (38.9)	0.02
Nail psoriasis	79 (39.1)	47 (34.6)	20 (41.7)	12 (66.7)	0.02
Physiotherapist assessment	100 (49.3)	56 (40.9)	29 (60.4)	15 (83.3)	0.002
Positive PASE	47 (24.3)	26 (20.5)	10 (20.8)	11 (61.1)	<0.001
Positive PEST	66 (33.9)	34 (26.2)	19 (39.6)	13 (76.5)	<0.001
Positive ToPAS-2	63 (32)	35 (26.7)	15 (31.3)	13 (72.2)	<0.001
Positive MSK-US: rule 1	120 (59.1)	75 (55.7)	29 (60.4)	16 (88.9)	0.01
CRP, mean ± SD mg/dl	1.9 ± 3	1.6 ± 3	1.3 ± 2.9	2.4 ± 6.9	0.04
ESR, mean ± SD mm/hour	11 ± 12	11.5 ± 11.5	10 ± 11.5	17.5 ± 23.5	0.13

* Values are the number (%) unless indicated otherwise. CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; IL-12 = interleukin-12; MSK-US = musculoskeletal ultrasound; PASE = Psoriatic Arthritis Screening and Evaluation; PASI = Psoriasis Area and Severity Index; PEST = Psoriasis Epidemiology Screening Tool; PsA = psoriatic arthritis; TNF = tumor necrosis factor; ToPAS-2 = Toronto Psoriatic Arthritis Screen 2.

Table 2. Performance of each triage method with clinical diagnosis of PsA as the outcome*

	Not PsA	Possibly PsA	PsA	Sensitivity (95% CI)	Specificity (95% CI)	AUC
MSK-US: rule 1				89 (65–99)	44 (37–51)	0.66
Positive	75	29	16			
Negative	62	19	2			
MSK-US: rule 2				56 (31–78)	77 (71–83)	0.66
Positive	31	11	10			
Negative	106	37	8			
Physical therapist assessment				83 (56–96)	54 (46–61)	0.63
Positive	56	29	15			
Negative	81	19	3			
Positive ToPAS-2				72 (47–90)	72 (65–79)	0.72
Positive	35	15	13			
Negative	96	33	5			
Positive PEST				77 (50–93)	70 (63–77)	0.73
Positive	34	19	13			
Negative	96	29	4			
Positive PASE				61 (36–83)	79 (73–85)	0.70
Positive	26	10	11			
Negative	101	38	7			

* Values are the number of patients unless indicated otherwise. The possibly psoriatic arthritis (PsA) group was combined with the not PsA group. 95% CI = 95% confidence interval; AUC = area under the curve (see Table 1 for other definitions).

modalities was required to be positive, and specificity tended to be highest when both modalities were required to be positive. The highest sensitivity achieved was 100% by MKS-US rule 1 or ToPAS-2 being positive and MSK-US rule 1 or physical therapist assessment being positive. The highest specificity was 95%, achieved by requiring both MSK-US rule 2 and PASE to be positive. One of the optimal combinations of sensitivity and specificity included requiring both PEST and MSK-US rule 1 to be positive (71% sensitivity and 81% specificity).

Change in disease status and musculoskeletal symptoms after 1 year. At 1 year, 122 of the 185 patients (75.7%) who were invited returned for a follow-up visit and an additional 39 patients (21.1%) completed questionnaires over the phone or by email. Of the 122 patients who returned for an in-person 1-year follow-up, 9 patients (7.4%) were newly converted into the PsA group, 16 patients (13.1%) were converted into the possibly PsA group, and 97 patients (79.5%) were converted into the not PsA group (Figure 1). Overall, 63 patients did not return for the 1-year assessment; of these 63 patients, 39 returned the screening questionnaires, and 24 were lost to follow-up. There were 38 completed questionnaires, of which 6 (15.4%) indicated a positive result for PsA on all 3 (PASE, PEST, and ToPAS-2) of the questionnaires completed.

The status at 1 year for all study patients (completed follow-up assessment or the diagnosis from baseline was carried forward, n = 203) was 27 in the PsA group (13.3%; 18 at baseline and 9 who were newly classified within 1 year), 34 (16.7%) in the possibly PsA group PsA, and 142 (79.3%) not PsA. With respect to the screening modality's ability to predict PsA status (at baseline or within 1 year), the individual screening modalities were found to have lower sensitivity, with no improvement in specificity (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24570). In addition, we evaluated the change in musculoskeletal symptoms, quality of life, and physical function over the course of a year

Table 3. Performance of combinations of the triage modalities*

	ToPAS		PASE		PEST		Physical therapist	
	Both tests positive	Any test positive	Both tests positive	Any test positive	Both tests positive	Any test positive	Both tests positive	Any test positive
MSK-US: rule 1								
Sensitivity (95% CI)	61 (36–83)	100 (80–100)	56 (31–78)	89 (67–99)	71 (44–90)	94 (70–100)	72 (46–90)	100 (82–100)
Specificity (95% CI)	82 (76–87)	34 (27–41)	89 (83–93)	35 (28–43)	81 (74–86)	33 (26–40)	71 (64–77)	27 (21–34)
AUC	0.79		0.78		0.78		0.75	
MSK-US: rule 2								
Sensitivity (95% CI)	44 (22–69)	83 (59–96)	39 (17–64)	79 (54–94)	47 (23–72)	82 (57–96)	44 (22–69)	94 (73–100)
Specificity (95% CI)	92 (87–95)	58 (51–65)	95 (91–98)	62 (55–69)	92 (87–96)	56 (48–63)	86 (81–91)	45 (38–52)
AUC	0.77		0.76		0.78		0.76	

* 95% CI = 95% confidence interval (see Table 1 for other definitions).

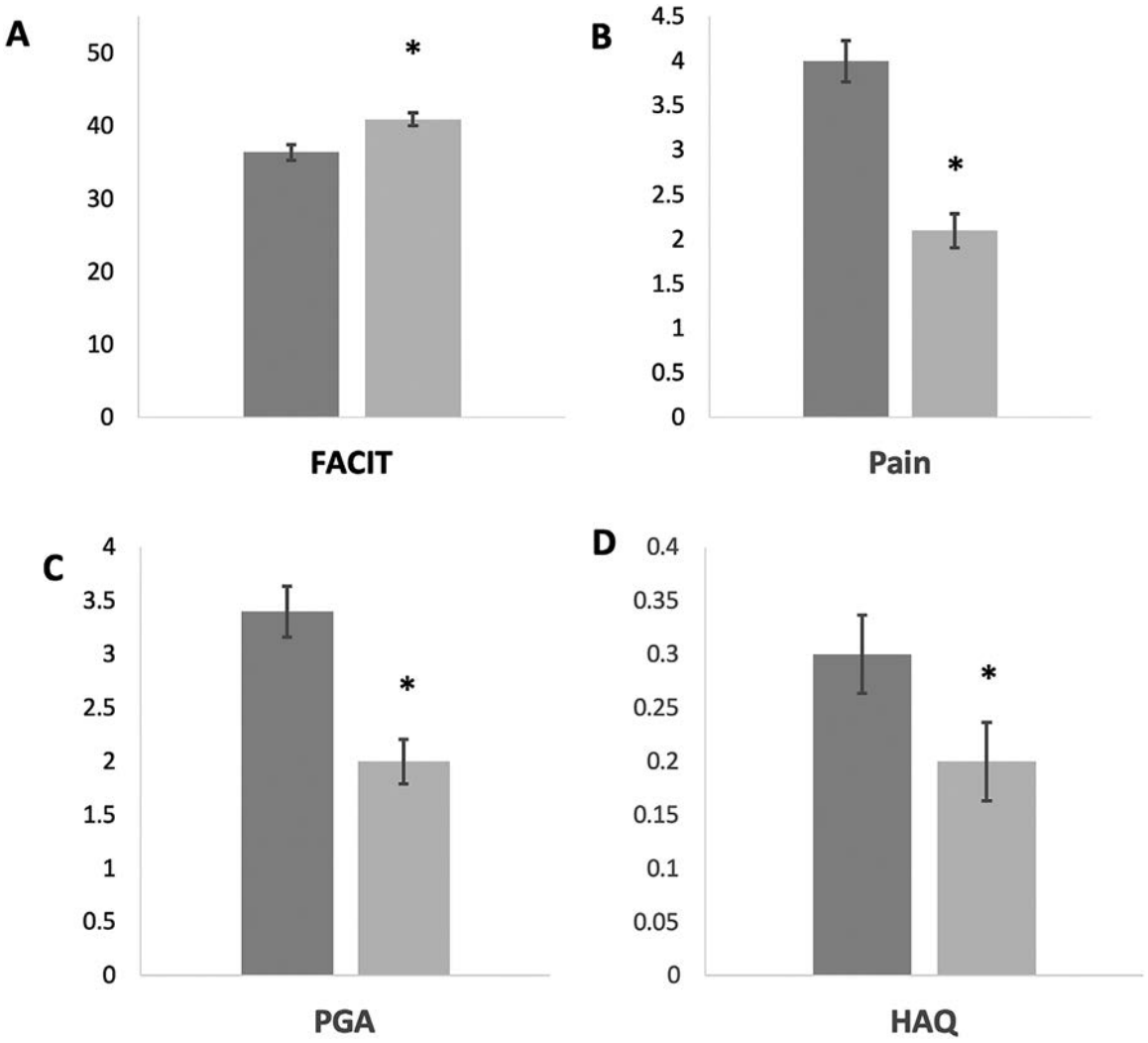


Figure 2. Evaluation of change shown from baseline (dark gray) in musculoskeletal symptoms, quality of life, and physical function over the course of a year (light gray), as indicated by **A**, Functional Assessment of Chronic Illness Therapy (FACIT); **B**, pain; **C**, Patient Global Assessment (PGA); and **D**, Health Assessment Questionnaire (HAQ) scores. An increase in score demonstrates improvement for all scales, except FACIT. * = $P < 0.001$.

(Figure 2). A significant improvement was seen in all patient-reported outcomes (FACIT, pain, PtGA-arthritis, DLQI, HAQ) compared to the baseline assessment ($P < 0.001$ for all).

DISCUSSION

In this study, we evaluated the performance of a number of components of a central triage system in a rapid access clinic for patients with psoriasis who are experiencing musculoskeletal pain. We found that none of the individual modalities had significantly superior sensitivity or specificity than the others in regard to their ability to identify patients with PsA; however, combining MSK-US with another tool improved overall performance. This suggests that MSK-US may provide additional information to aid in the initial screening of patients suspected of having PsA. In addition, we found that these triage modalities did not have any predictive value with respect to identifying patients who were newly classified as having PsA within 1 year. Finally, patient symptoms, quality of life, and function have improved over the course of 1 year. These findings help guide the role of various triage modalities in the diagnostic process of PsA.

The rapid access triage model has been used in the diagnostic process of some medical conditions, including inflammatory bowel disease and skin cancer (26,27). Triage systems of different forms have been successful in reducing wait times for rheumatologist visits. One study showed that a central referral and triage system was able to reduce wait times for moderate and urgent referrals (28). A recent Canadian study evaluated a rapid access triage system for inflammatory arthritis, comparing advanced trained practitioners to the rheumatologist assessment and found the sensitivity to be 100% and the specificity to be 93.1% (6). These results show a higher level of performance than in our study; however, their objective was to identify inflammatory arthritis in general and not specifically PsA.

In combining MSK-US results with another tool, we demonstrated an improvement in sensitivity and specificity. Solmaz et al showed that MSK-US informed the decision of dermatologists regarding the need for referring psoriasis patients to rheumatology for suspected PsA (29). Without MSK-US data, the sensitivity of a referral (versus PsA diagnosis by rheumatologist) was 95%, and specificity was 9%, but this changed to a sensitivity of 88% and specificity of 77% with access to MSK-US data. This aligns well with our results in combining the use of questionnaires and MSK-US and suggests that MSK-US can add value in a triage system. However, it should be kept in mind that the added benefit of combining MSK-US with other modalities is associated with added time and cost and requires a skilled sonographer who may not be available in every setting.

The improvement in patient-reported outcomes at 1 year is reassuring and suggests that a significant proportion of patients improve spontaneously or as a result of treatment provided for noninflammatory musculoskeletal conditions. The fact that such an improvement remains after removing PsA patients from the analysis may be in part due to treatment recommendations that may have been given to the patients at baseline, such as physical therapy, massage, or nonsteroidal antiinflammatory drugs, which may improve their musculoskeletal symptoms. In addition, the course of many arthritic conditions naturally fluctuates over time, or patients show improvement on their own.

There are some limitations to our study. This is a single-center study with a majority of the patients coming from dermatology clinics that may have limited the generalizability due to differences in referral practices and local population characteristics. However, it may be argued that the patient population seen in dermatology settings is the most relevant population on which to focus future screening efforts, since the severity of psoriasis is a risk factor for the development of PsA (30), and the prevalence of PsA found among patients attending dermatology clinics tends to be significantly higher than that found in the general population (19,31). Second, MSK-US protocol was limited to the symptomatic joints and the contralateral side, which resulted in a different number of joints scanned for each patient. This may have increased the chances of detecting positive MSK-US findings with more peripheral joints scanned, which may have led to underperformance of MSK-US in patients with oligoarthritis and in those with predominantly axial disease. In addition, there may be a participation bias as those patients with more severe joint pain may be more interested in having their joints examined and therefore more interested in participating in the study. Finally, being a proof-of-concept study, we evaluated the efficacy of the different triage methods in a research setting but did not perform any cost assessment or evaluation as to whether this model was superior to the standard of care in terms of patient outcomes.

In conclusion, we evaluated the performance of various central triage modalities in a rapid access clinic to identify PsA among patients with psoriasis who experience musculoskeletal symptoms. We found that no individual tool is superior to another, yet combining MSK-US with another modality provided the optimal performance. This conclusion, however, needs to be balanced in terms of cost effectiveness. Further research needs to be done on how the central screening modalities affect wait times and short- and long-term patient outcomes and whether it is feasible on a larger scale.

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. Eder 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. Sarabia, Farrer, Yeung, Jerome, Cook, Eder.

Acquisition of data. Sarabia, Farrer, Yeung, Jerome, Eder.

Analysis and interpretation of data. Cook, Eder.

REFERENCES

1. Eder L, Widdifield J, Rosen CF, Cook R, Lee KA, Alhusayen R, et al. Trends in the prevalence and incidence of psoriasis and psoriatic arthritis in Ontario, Canada: a population-based study. Arthritis Care Res (Hoboken) 2019;71:1084–91.
2. Haroon M, Gallagher P, FitzGerald O. Diagnostic delay of more than 6 months contributes to poor radiographic and functional outcome in psoriatic arthritis. Ann Rheum Dis 2015;74:1045–50.
3. Widdifield J, Bernatsky S, Thorne JC, Bombardier C, Jaakkimainen RL, Wing L, et al. Wait times to rheumatology care for patients with rheumatic diseases: a data linkage study of primary care electronic medical records and administrative data. CMAJ Open 2016;4:E205–12.
4. Coates LC, Aslam T, Al Balushi F, Burden AD, Burden-Teh E, Caperon AR, et al. Comparison of three screening tools to detect psoriatic arthritis in patients with psoriasis (CONTEST study). Br J Dermatol 2013;168:802–7.
5. Richard MA, Barnetche T, Rouzaud M, Sevrain M, Villani AP, Aractingi S, et al. Evidence-based recommendations on the role of dermatologists in the diagnosis and management of psoriatic arthritis: systematic review and expert opinion. J Eur Acad Dermatol Venereol 2014;28 Suppl 5: 3–12.
6. Ahluwalia V, Lineker S, Sweezie R, Bell MJ, Kendzerska T, Widdifield J, et al. The effect of triage assessments on identifying inflammatory arthritis and reducing rheumatology wait times in Ontario. J Rheumatol 2020;47:461–7.
7. Passalent LA, Kennedy C, Warmington K, Soever LJ, Lundon K, Shupak R, et al. System integration and clinical utilization of the Advanced Clinician Practitioner in Arthritis Care (ACPAC) program–trained extended role practitioners in Ontario: a two-year, system-level evaluation. Healthcare Policy 2013;8:56–70.
8. Carpenter T, Katz SJ. Review of a rheumatology triage system: simple, accurate, and effective. Clin Rheumatol 2014;33:247–52.
9. Gormley GJ, Steele WK, Gilliland A, Leggett P, Wright GD, Bell AL, et al. Can diagnostic triage by general practitioners or rheumatology nurses improve the positive predictive value of referrals to early arthritis clinics? Rheumatology (Oxford) 2003;42:763–8.
10. Caffrey A, Smart KM, FitzGerald O. Physiotherapist-led triage at a rheumatology-based musculoskeletal assessment clinic: an 18-month service evaluation of activity and outcomes. ACR Open Rheumatol 2019;1:213–8.

11. Caldarola G, D'Agostino M, Carbone A, Guerriero C, Bonomo L, Amerio P, et al. Usefulness of ultrasound imaging in detecting psoriatic arthritis of fingers and toes in patients with psoriasis. Clin Develop Immunol 2011;2011:390726.
12. Zabotti A, Bandinelli F, Batticciotto A, Scire CA, Iagnocco A, Sakellariou G, et al. Musculoskeletal ultrasonography for psoriatic arthritis and psoriasis patients: a systematic literature review. Rheumatology (Oxford) 2017;56:1518–32.
13. Sankowski AJ, Łebkowska UM, Ćwikła J, Walecka I, Walecki J. The comparison of efficacy of different imaging techniques (conventional radiography, ultrasonography, magnetic resonance) in assessment of wrist joints and metacarpophalangeal joints in patients with psoriatic arthritis. Polish J Radiol 2013;78:18.
14. Ostergaard M, Eder L, Christiansen SN, Kaeley GS. Imaging in the diagnosis and management of peripheral psoriatic arthritis: the clinical utility of magnetic resonance imaging and ultrasonography. Best Pract Res Clin Rheumatol 2016;30:624–37.
15. Sarabia S, Farrer C, Yeung J, Jerome D, Lee KA, Cook R, et al. The pattern of musculoskeletal complaints in patients with suspected psoriatic arthritis and their correlation with physical examination and ultrasound. J Rheumatol 2021;48:214–21.
16. Gladman DD, Schentag CT, Tom BD, Chandran V, Brockbank J, Rosen C, et al. Development and initial validation of a screening questionnaire for psoriatic arthritis: the Toronto Psoriatic Arthritis Screen (ToPAS). Ann Rheum Dis 2009;68:497–501.
17. Ibrahim GH, Buch MH, Lawson C, Waxman R, Helliwell PS. Evaluation of an existing screening tool for psoriatic arthritis in people with psoriasis and the development of a new instrument: the Psoriasis Epidemiology Screening Tool (PEST) questionnaire. Clin Exp Rheumatol 2009;27:469–74.
18. Dominguez P, Husni ME, Garg A, Qureshi AA. Psoriatic Arthritis Screening and Evaluation (PASE) questionnaire and the role of dermatologists: a report from the GRAPPA 2009 annual meeting. J Rheumatol 2011;38:548–50.
19. Haroon M, Kirby B, FitzGerald O. High prevalence of psoriatic arthritis in patients with severe psoriasis with suboptimal performance of screening questionnaires. Ann Rheum Dis 2013;72:736–40.
20. Maksymowych WP, Mallon C, Morrow S, Shojania K, Olszynski WP, Wong RL, et al. Development and validation of the Spondyloarthritis Research Consortium of Canada (SPARCC) Enthesitis Index. Ann Rheum Dis 2009;68:948–53.
21. Eder L, Jayakar J, Thavaneswaran A, Haddad A, Chandran V, Salonen D, et al. Is the Madrid Sonographic Enthesitis Index useful for differentiating psoriatic arthritis from psoriasis alone and healthy controls? J Rheumatol 2014;41:466–72.
22. Terslev L, Naredo E, Aegerter P, Wakefield RJ, Backhaus M, Balint P, et al. Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce. Part 2: reliability and application to multiple joints of a standardised consensus-based scoring system. RMD Open 2017;3:e000427.
23. Bruyn GA, Hanova P, Iagnocco A, d'Agostino MA, Moller I, Terslev L, et al. Ultrasound definition of tendon damage in patients with rheumatoid arthritis: results of a OMERACT consensus-based ultrasound score focusing on the diagnostic reliability. Ann Rheum Dis 2014;73:1929–34.
24. Terslev L, Naredo E, Iagnocco A, Balint PV, Wakefield RJ, Aegerter P, et al. Defining enthesitis in spondyloarthritis by ultrasound: results of a Delphi process and of a reliability reading exercise. Arthritis Care Res (Hoboken) 2014;66:741–8.
25. Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P, Mielants H, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006;54:2665–73.
26. Reinglas J, Restellini S, Gonczi L, Kurti Z, Verdon C, Nene S, et al. Harmonization of quality of care in an IBD center impacts disease outcomes: importance of structure, process indicators and rapid access clinic. Dig Liver Dis 2019;51:340–5.
27. Graves J, Fleischman MH, Goldstein GD. Derm access: a new triage system to rapidly identify suspicious skin lesions. Dermatol Surg 2006;32:1486–90.
28. Hazlewood GS, Barr SG, Lopatina E, Marshall DA, Lupton TL, Fritzler MJ, et al. Improving appropriate access to care with central referral and triage in rheumatology. Arthritis Care Res (Hoboken) 2016;68:1547–53.
29. Solmaz D, Bakirci S, Al Onazi A, Al Osaimi N, Fahim S, Aydin SZ. Musculoskeletal ultrasound can improve referrals from dermatology to rheumatology for patients with psoriasis. Br J Dermatol 2020;182:804–6.
30. Eder L, Haddad A, Rosen CF, Lee KA, Chandran V, Cook R, et al. The incidence and risk factors for psoriatic arthritis in patients with psoriasis: a prospective cohort study. Arthritis Rheumatol 2016;68:915–23.
31. Ibrahim G, Waxman R, Helliwell PS. The prevalence of psoriatic arthritis in people with psoriasis. Arthritis Rheum 2009;61:1373–8.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1263–1268
DOI 10.1002/acr.24572

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Predictors of Initial Hydroxychloroquine Receipt Among Medicaid Beneficiaries With Incident Systemic Lupus Erythematosus

Katherine P. Pryor, Chang Xu, Jamie E. Collins, Karen H. Costenbader, and Candace H. Feldman

Objective. Although hydroxychloroquine/chloroquine (HCQ/CQ) form the cornerstone of systemic lupus erythematosus (SLE) treatment, not all patients receive this, which may contribute to disparities in outcomes. The present study was undertaken to investigate factors associated with first dispensing of HCQ/CQ.

Methods. Using Medicaid insurance claims from 2000 to 2010, we identified individuals ages 18–65 years with incident SLE (≥3 SLE International Classification of Diseases, Ninth Revision codes separated by ≥30 days without prior SLE codes or HCQ/CQ use for 24 months). The primary outcome was first dispensing of HCQ/CQ within 24 months of the first SLE code. We used Cox proportional hazards regression models to examine the association between sociodemographic factors, comorbidities, health care utilization, and medication use and HCQ/CQ dispensing within 24 months of diagnosis.

Results. We identified 9,560 Medicaid beneficiaries with incident SLE; 41% received HCQ (n = 3,949) or CQ (n = 14) within 24 months of diagnosis. Younger patients were more likely to receive HCQ/CQ. Black, Asian, Hispanic, and American Indian/Alaska Native individuals were more likely to receive HCQ/CQ than White individuals. Alcohol and nicotine use, chronic pain, diabetes mellitus, and end-stage renal disease were associated with lower dispensing. Appointments and preventive care services were associated with higher rates, and more hospitalizations with lower rates.

Conclusion. Only 41% of Medicaid beneficiaries with SLE received HCQ/CQ within 24 months of diagnosis. Greater outpatient and preventive care increased receipt. All non-White race/ethnicities had higher rates of first dispensing. Time to initial HCQ/CQ dispensing may not explain racial/ethnic disparities in adverse outcomes, highlighting the need to consider other care quality-related issues and medication adherence challenges.

INTRODUCTION

Hydroxychloroquine/chloroquine (HCQ/CQ) is the backbone of high-quality systemic lupus erythematosus (SLE) care. HCQ/CQ may improve survival, reduce flare rates, thrombosis, and progression to end-stage renal disease (ESRD) (1). Despite a largely reassuring safety profile, in particular for initial use in early SLE, and its importance in reducing flares and potentially avoidable SLE-related adverse outcomes (1), there are variations in HCQ/CQ receipt. Prior studies in international academic medical center cohorts suggest that between 67% and 80% of individuals with SLE receive HCQ/CQ, in stark contrast to the 36.4% seen in a cohort of Medicaid beneficiaries with incident lupus nephritis (2).

Medicaid is the largest US public insurance for low-income individuals and provides coverage to roughly 1 in 5 Americans. Patients with SLE enrolled in Medicaid experience a disproportionate burden of SLE and adverse outcomes relative to the general population (3,4). Differences in receipt of HCQ/CQ may exacerbate disparities in adverse events. It is also possible that

The content herein is solely the responsibility of the authors and does not necessarily represent the views of the funding source.

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant K23-AR-071500) and the Rheumatology Research Foundation (Investigator award to Dr. Feldman). Dr. Feldman also receives research support from the HHS Office of Minority Health and Brigham and Women's Hospital.

Katherine P. Pryor, MD, MS, Chang Xu, MS, Jamie E. Collins, PhD, Karen H. Costenbader, MD, MPH, Candace H. Feldman, MD, MPH, ScD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Dr. Collins has received consulting fees from Boston Imaging Core Labs (less than $10,000). No other disclosures relevant to this article were reported.

Address correspondence to Katherine P. Pryor, MD, MS, Brigham and Women's Hospital, Department of Medicine, 75 Francis Street, Boston, MA 02115. Email: kpryor@bwh.harvard.edu.

Submitted for publication November 16, 2020; accepted in revised form February 2, 2021.

SIGNIFICANCE & INNOVATIONS

- Less than one-half of Medicaid beneficiaries with incident systemic lupus erythematosus (SLE) receive hydroxychloroquine/chloroquine (HCQ/CQ), the backbone of high-quality care, within 24 months of diagnosis.
- Preventive care and outpatient visits, markers of access to health care, were associated with higher rates of initial HCQ/CQ dispensing, whereas frequent hospitalizations, high-risk lifestyles, and complex comorbidities were associated with lower rates.
- In this Medicaid population, individuals from all non-White racial/ethnic groups had higher rates of HCQ/CQ dispensing compared to White individuals, suggesting that initial HCQ/CQ dispensing differences do not explain racial/ethnic disparities in adverse outcomes.
- Significant differences in characteristics of those who receive versus do not receive HCQ/CQ suggest that caution is warranted in the interpretation of observational studies comparing outcomes between HCQ/CQ users and nonusers.

significant differences between who is and is not initially prescribed HCQ lead to bias in our interpretation of the associations between HCQ/CQ use and outcomes in observational studies. For example, the protective effects attributed to HCQ/CQ use may be related to characteristics of individuals who receive HCQ/CQ compared to those who do not. While propensity score-matched analyses help address measured confounding, bias from unmeasured factors (e.g., healthy behaviors) remains.

The aim of this study was to investigate factors associated with first dispensing of HCQ/CQ among Medicaid beneficiaries with incident SLE. We hypothesized that individuals receiving HCQ/CQ would have better health care access, as measured by outpatient visits and preventive care (e.g., cancer screening and vaccinations), and fewer comorbidities compared to individuals who do not receive HCQ/CQ within 2 years of SLE diagnosis.

PATIENTS AND METHODS

Study population and outcome of interest. We utilized Medicaid data, MAX, with demographic information, pharmacy dispensing data, and billing claims from 2000 to 2006 (47 states) and 2007 to 2010 (29 states). We identified individuals ages 18–65 years with incident SLE (≥3 SLE International Classification of Diseases, Ninth Revision [ICD-9] codes separated by ≥30 days within 24 months), as previously described, and used the first SLE code to define the index date (4). To capture new use of HCQ, we required 24 months of continuous enrollment without SLE codes and without HCQ/CQ dispensings prior to the index date of incident SLE. The primary outcome was the first dispensing of HCQ/CQ on or within 24 months of the index date.

Baseline covariates. Sociodemographic factors were assessed at the index date and included age, sex (male, female), race/ethnicity (White, Black, Asian, American Indian/Alaska Native, Hispanic, and >1 race/ethnicity), and region of the US (Midwest, Northeast, South, and West). All other covariates were assessed in the 12 months prior to the index date. We measured health care utilization as outpatient visits (0, 1–5 visits, or >5 visits), emergency department visits (0, ≥1), and hospitalizations (0, 1, ≥2). Preventive care use was defined as visits with codes for annual physical examinations, vaccinations, Pap tests, mammograms, colonoscopies, and bone density scans (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24572) and categorized as 0, 1–2, or ≥2 visits. Medication use, also measured in the 12 months prior to the index date, included glucocorticoids, immunosuppressives (azathioprine, mycophenolate mofetil, cyclophosphamide, methotrexate, leflunomide, rituximab, and sulfasalazine), prescription nonsteroidal antiinflammatory drugs (NSAIDs), and opioids. Comorbidities included alcohol use disorder, cardiovascular disease, diabetes mellitus, chronic pain, smoking, renal disease, ESRD, and pregnancy, all defined using ICD-9 codes during the baseline 12-month period (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24572).

Statistical analysis. We used Cox proportional hazards models to examine the associations (with hazards ratios [HR] and 95% confidence intervals [95% CIs]) between baseline sociodemographic factors, comorbidities, health care utilization, preventive care, and medication use and time to first HCQ/CQ dispensing in the first 24 months following SLE diagnosis. Models were additionally adjusted by calendar year of the index date (2000–2003, 2004–2006, 2007–2010). Analyses were conducted in SAS, version 9.4, and the Mass General Brigham Institutional Review Board approved this study. Medicaid Analytic eXtract (MAX) data were obtained through a Data Use Agreement through the Centers of Medicare and Medicaid Services; cell sizes <11 were suppressed in accordance with their policies.

RESULTS

We identified 9,560 Medicaid beneficiaries with incident SLE without prior HCQ use. The mean ± SD age was 35.7 ± 11.7 years for HCQ/CQ users and 39.9 ± 12.2 years for nonusers. In all, 44.8% of HCQ/CQ users were Black compared to 39.4% of nonusers; 19.8% of users were Hispanic compared to 14.8% of nonusers (Table 1). There were 3,963 individuals (41%) who received HCQ (n = 3,949) or CQ (n = 14) within 24 months of meeting criteria for SLE in this data set, and this proportion did not increase with later index dates. Seventy-two percent received glucocorticoids within 24 months of diagnosis; of these, 51.1% also received HCQ/CQ.

Table 1. Baseline characteristics of individuals with incident systemic lupus erythematosus (n = 9,560) with and without ≥1 dispensing of hydroxychloroquine (HCQ) or chloroquine (CQ) within the first 2 years of diagnosis*

Characteristic	HCQ/CQ users (n = 3,963)	HCQ/CQ nonusers (n = 5,597)
Age, mean ± SD years	35.7 ± 11.7	39.9 ± 12.2
Female	3,778 (95.3)	5,228 (93.4)
Race/ethnicity		
Asian	134 (3.4)	104 (1.9)
Black	1,776 (44.8)	2,206 (39.4)
Hispanic	783 (19.8)	829 (14.8)
American Indian/Alaska Native	44 (1.1)	58 (1.0)
White	1,128 (28.5)	2,221 (39.7)
Other	98 (2.5)	179 (3.2)
Region		
Midwest	753 (19.0)	1,054 (18.8)
Northeast	1,012 (25.5)	1,118 (20.0)
South	1,226 (30.9)	2,192 (39.2)
West	972 (24.5)	1,233 (22.0)
Year of index date		
2000–2003	1,041 (26.3)	1,287 (23.0)
2004–2006	1,723 (43.5)	2,243 (40.1)
2007–2010	1,199 (30.3)	2,067 (36.9)
Medication use		
Steroid use	1,339 (33.8)	1,506 (26.9)
Immunosuppressive agent†	146 (3.7)	377 (6.7)
NSAID use	2,190 (55.3)	2,208 (39.5)
Opioid use	69 (1.7)	190 (3.4)
Comorbidities		
Alcohol use disorder	265 (6.7)	554 (9.9)
Cardiovascular disease	2,901 (73.2)	4,176 (74.6)
Diabetes mellitus	1,045 (26.4)	1,873 (33.5)
Chronic pain	606 (15.3)	1,165 (20.8)
Smoking	810 (20.4)	1,416 (25.3)
Renal disease	756 (19.1)	1,044 (18.7)
End-stage renal disease	194 (4.9)	366 (6.5)
Pregnancy	968 (24.4)	1,096 (19.6)
Preventive care‡		
1–2	2,088 (52.7)	2,860 (51.1)
>2	1,401 (35.4)	1,619 (28.9)
Health care utilization		
1–5 outpatient visits	1,403 (35.4)	1,838 (32.8)
>5 outpatient visits	1,750 (44.2)	2,313 (41.3)
1 hospitalization	535 (13.5)	770 (13.8)
≥2 hospitalizations	315 (8.0)	542 (9.7)
≥1 emergency department visit	2,154 (54.4)	2,839 (50.7)

* Values are the number (%) unless indicated otherwise. NSAID = nonsteroidal antiinflammatory drug.
† Immunosuppressive agents included azathioprine, mycophenolate mofetil, cyclophosphamide, methotrexate, leflunomide, rituximab, and sulfasalazine.
‡ Preventive care included annual physical examinations, vaccinations, pap tests, mammograms, colonoscopies, and bone density scans.

Table 2. Baseline factors associated with first hydroxychloroquine (HCQ) or chloroquine dispensing among Medicaid beneficiaries with incident systemic lupus erythematosus (n = 9,560)*

	HR (95% CI)†
Male sex (Ref. = female)	0.88 (0.75–1.02)
Age (Ref. = 51–65 years)	
18–24	2.33 (2.05–2.64)‡
25–31	1.99 (1.76–2.24)‡
32–38	1.74 (1.54–1.95)‡
39–45	1.31 (1.17–1.48)‡
46–50	1.15 (1.01–1.32)‡
Race/ethnicity (Ref. = White)	
Asian	1.61 (1.34–1.94)‡
American Indian/Alaska Native	1.49 (1.10–2.02)‡
Hispanic	1.36 (1.24–1.50)‡
Black	1.29 (1.20–1.40)‡
Other	1.05 (0.85–1.29)
Region (Ref. = Northeast)	
Midwest	0.94 (0.85–1.04)
West	0.90 (0.82–0.99)‡
South	0.76 (0.70–0.83)‡
Medication use (Ref. = no use)	
Prescription NSAID use	1.53 (1.43–1.64)‡
Glucocorticoid use	1.27 (1.18–1.36)‡
Opioid use	0.75 (0.59–0.96)‡
Immunosuppressive agents use§	0.48 (0.40–0.57)‡
Comorbidities	
Renal disease (excluding ESRD)	1.07 (0.97–1.17)
Cardiovascular disease	1.06 (0.98–1.15)
Smoking	0.90 (0.83–0.98)‡
Diabetes mellitus	0.84 (0.77–0.90)‡
Chronic pain	0.83 (0.76–0.91)‡
Pregnancy	0.80 (0.74–0.87)‡
Alcohol use disorder	0.74 (0.65–0.85)‡
ESRD	0.72 (0.61–0.84)‡
Preventive care (Ref. = 0)¶	
1–2	1.29 (1.16–1.43)‡
>2	1.46 (1.30–1.64)‡
Health care utilization (Ref. = none)	
1–5 outpatient visits	1.08 (0.99–1.18)
>5 outpatient visits	1.10 (1.00–1.21)‡
1 hospitalization	0.95 (0.87–1.05)
≥2 hospitalizations	0.87 (0.77–0.98)‡
≥1 emergency department visit	0.99 (0.92–1.06)

* Included 3,949 HCQ recipients and 14 chloroquine recipients. Numerical values represent the sum of different preventive care types. 95% CI = 95% confidence interval; ESRD = end-stage renal disease; HR = hazard ratio; NSAID = nonsteroidal antiinflammatory drug; Ref. = reference.
† HRs from multivariable Cox proportional hazards regression model additionally adjusted for calendar year of index date and emergency department visits.
‡ Statistically significant association.
§ Immunosuppressive agents included azathioprine, mycophenolate mofetil, cyclophosphamide, methotrexate, leflunomide, rituximab, and sulfasalazine.
¶ Preventive care during the baseline period included annual physical examinations, vaccinations, pap tests, mammograms, colonoscopy, and bone density scans.

Multivariable-adjusted Cox models demonstrated that younger patients (age 18–24 years) were 2.3 times more likely to receive HCQ/CQ compared with 51–65 year-olds (Table 2). All racial/ethnic groups examined (Asian, American Indian/Alaska Native, Hispanic, and Black individuals) were more likely to receive HCQ/CQ than were White individuals. Relative to the Northeast, living in the South or the West was associated with less HCQ/CQ use, particularly in the South (HR 0.76 [95% CI 0.70–0.83]).

We explored underlying factors that could increase hesitancy to use HCQ/CQ. We examined baseline tamoxifen use and

glucose-6-phosphate dehydrogenase (G6PD) deficiency and both were not statistically significant. We examined ophthalmologic conditions including retinal disease, macular degeneration, and cataracts (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24572) during the baseline period and found an increased risk of HCQ/CQ dispensing (HR 1.32 [95% CI 1.20–1.46]). We also found that 51% of these codes occurred in individuals ages 51–65 years. We did not include these variables in our final model, as the point estimates did not change significantly for other factors.

Complex medical comorbidities, including alcohol use disorder, chronic pain, diabetes mellitus, ESRD, and smoking were associated with a lower rate of dispensing, but no association was seen with cardiovascular disease or non-ESRD renal disease. Notably, pregnancy was associated with lower rates of receipt. Individuals were also less likely to receive HCQ/CQ if they were taking immunosuppressive medications but were more likely if they were using NSAIDs (HR 1.53 [95% CI 1.43–1.64]) or glucocorticoids (HR 1.27 [95% CI 1.18–1.36]). Receiving 1–2 preventive care services was associated with 29% higher rate of receipt, >2 with a 46% higher rate. Individuals with >5 outpatient visits were more likely to receive HCQ/CQ, while those with ≥2 hospitalizations were less likely.

DISCUSSION

HCQ/CQ use is central to SLE treatment; however, we observed that only 41% of Medicaid beneficiaries with SLE received HCQ/CQ within 2 years of diagnosis (2). While in line with our prior study among individuals enrolled in Medicaid with incident lupus nephritis, this is in contrast to the nearly 80% who receive HCQ/CQ in other cohorts such as Systemic Lupus International Collaborating Clinics and LUpus in MInorities, NAture versus nurture (LUMINA) (5,6). Medicaid beneficiaries with SLE are a vulnerable population who experience a disproportionate burden of SLE and adverse outcomes in part due to lack of access to high-quality, coordinated subspecialty care (3,4,7). Progression to ESRD and mortality are markedly higher among Medicaid beneficiaries, likely in part due to lower quality of care provided to Medicaid beneficiaries with incident SLE (8,9). Reduced rates of HCQ/CQ receipt within the Medicaid population may contribute to the poorer outcomes in this population. However, within this vulnerable population, there were clear differences between those who receive HCQ/CQ and those who did not. We found that patients who were younger and those who were Black, Asian, Hispanic, and American Indian/Alaska Native were more likely to receive HCQ/CQ compared to individuals who were older or White (2). Geography seems to play a role; patients in the Northeast had a higher likelihood of receiving HCQ/CQ compared to the South and the West, possibly in part related to the concentration of rheumatologists.

We did not find statistically significant associations between factors that may be relative contraindications to HCQ/CQ use including tamoxifen use and G6PD deficiency. Regarding baseline ophthalmologic conditions, we paradoxically observed an increased risk of use associated with having a baseline ICD-9 code for retinal disease, macular degeneration, or cataracts. We suspect that this is related to surveillance bias due to increased screening prior to initiation of HCQ in general, and possibly among patients with known baseline eye disease. Some of these codes, specifically those for toxic maculopathy, may have been used as rule out codes in patients receiving baseline eye exams prior to HCQ use. However, 51% of these codes related to ophthalmologic conditions were found in individuals ages 51–65. It is possible that prevalence of eye disease, or concern for eye disease, explains in part the lower rates of dispensing in this age group.

Complex comorbidities were also associated with lower rates of receipt. Notably, the association of ESRD, but not earlier stages of renal disease with lower rates of HCQ prescribing, is important because it highlights an active decision among clinicians who may not prescribe HCQ/CQ in the setting of ESRD either due to quiescent disease or increased toxicity risk. However, we recognize as a limitation that detailed data in a Medicaid data set from patients with ESRD is challenging, as many become dually eligible for Medicare soon after developing ESRD, and we did not have a Data Use Agreement that would have allowed us to link our data to the US Renal Data System (USRDS). We acknowledge that this can affect dispensing data as well and thus, we may have underestimated the number of patients with SLE-related ESRD receiving HCQ. However, our findings support lower use in this population possibly due to increased concern for toxicity, as well as the thought that patients with ESRD may have less active SLE. Finally, it is also notable that pregnancy was associated with decreased rates of HCQ/CQ receipt. We suspect that these low proportions may in part be due to only more recent recognition of the safety and benefit of HCQ in pregnancy, with few randomized studies published during the earlier years of our study (10).

Preventive services and outpatient visits, surrogates for access to care, were associated with higher rates of HCQ/CQ dispensing. In a study of Medicaid beneficiaries with incident lupus nephritis, treat-and-release emergency department (ED) visits, a surrogate for inconsistent ambulatory care, were associated with lower receipt of recommended treatments, including HCQ/CQ compared to beneficiaries with more outpatient visits. Receipt of an antimalarial in this population was 36.4% at 90 days but increased to 45.8% by 1 year, suggesting that delays in access to care also contribute to lower rates of receipt (2). Additionally, receipt of an antimalarial in this population was also associated with lower odds of recurrent acute care use (defined as ED visits and hospitalizations) (11). We do not have available data on specialty of physician seen but suspect that difficulties in access to rheumatologists partly explains lower rates of dispensing in this vulnerable population.

Although Black and Hispanic patients appeared to be more likely to receive HCQ following a SLE diagnosis, they experience worse SLE-related outcomes than their White counterparts. It is possible that this may be more related to fragmented care following diagnosis and also to multilevel factors that may limit adherence to SLE-related medications. In a study of 10,268 Medicaid patients who newly initiated HCQ, <20% of patients were adherent, with lower rates seen in younger patients, Black and Hispanic patients, and in patients with higher acute care use (ED visits and hospitalizations), diabetes mellitus, and antidepressant use. Lower odds of adherence were also observed in zip codes with higher percentages of Black residents, even after controlling at the individual level for race, comorbidities, and utilization factors (12). In a recent study by Sun et al, reasons for nonadherence (particularly to mycophenolate mofetil) varied by race. Black patients were less likely to report high adherence, refills, and composite adherence to all SLE medications compared to White patients (13). Nonadherence among Black patients was associated with lower survey ratings on compassionate respectful interactions with providers and poorer survey scores on mental health. Among White patients, lower adherence was associated with greater medication burden and fibromyalgia pain scores (13). It is also possible that non-White individuals may present with more active disease possibly due to initial delays in care and diagnosis. White patients may have either milder SLE on presentation or increased rates of misclassification, which could explain the lower rates of HCQ prescribing with similar numbers of SLE-related ICD-9 codes. Further studies that include SLE activity-related data are needed to explore this question.

Our findings also suggest that those who received HCQ/CQ within 2 years may be systematically different from those who did not (14,15). In the LUMINA cohort, Alarcón et al analyzed the impact of HCQ on survival independent of sociodemographic and clinic characteristics. They constructed propensity scores and found that individuals with milder disease, and/or higher socioeconomic status, were the most likely to receive HCQ (100%) versus not receive HCQ (39.6%) (15). Although HCQ was noted to improve survival, after adding propensity scores to the model, the confidence interval became much wider (15). While propensity scores can help adjust for measured confounding by indication (factors that may be associated with the decision to treat with HCQ versus not treat that also may be associated with the outcome of interest), they cannot account for unmeasured confounding. This may be particularly significant when comparing HCQ users to nonusers, as users may be healthier and have more consistent access to high-quality care. These factors can be captured in part by measured covariates, but other healthy behaviors that may correspond with this (e.g., regular exercise, healthy diet) may not be.

This study has several strengths. While prior studies have examined medication adherence and associations between HCQ use and outcomes, few have specifically explored predictors of initial HCQ dispensing. Our study was conducted in a large, racially and ethnically diverse population-based cohort of individuals with a high burden of SLE, comorbid conditions, and adverse outcomes. Thus, exploring potentially modifiable strategies to improve outcomes is especially important. Our study also has limitations. ICD-9 codes were used to identify incident SLE cases, and misclassification is possible. We lacked measures of SLE disease activity and time from initial SLE symptoms to diagnosis. We also lacked data on other healthy behaviors and on access to subspecialists. While we had data on first dispensing of HCQ, we do not know how many individuals may have been given a prescription for HCQ but never filled it. Although our data are from 2000–2010 and new trends may have emerged, over the course of the decade examined we did not see a significant trend toward increased use.

Overall, our findings raise the question of whether receipt (or nonreceipt) of timely and high-quality SLE care, including HCQ/CQ, may exacerbate disparities in adverse events. Somewhat unexpectedly, while prior studies have demonstrated significantly poorer quality metrics and a higher burden of adverse outcomes among Black and Hispanic individuals with SLE compared to White individuals, we did not see this for initial HCQ/CQ dispensing. This finding highlights opportunities for downstream interventions after HCQ/CQ receipt that may more directly reduce disparities in outcomes. Although additional research can help to further elucidate areas where intervention is needed, it is clear that these approaches must at the least improve access to care and medication adherence for the most vulnerable populations.

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. Feldman 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. Pryor, Xu, Collins, Costenbader, Feldman.

Acquisition of data. Xu, Costenbader, Feldman.

Analysis and interpretation of data. Pryor, Xu, Collins, Costenbader, Feldman.

REFERENCES

1. Lee SJ, Silverman E, Bargman JM. The role of antimalarial agents in the treatment of SLE and lupus nephritis. Nat Rev Nephrol 2011;7:718–29.
2. Yazdany J, Feldman CH, Liu J, Ward MM, Fischer MA, Costenbader KH. Quality of care for incident lupus nephritis among Medicaid beneficiaries in the United States. Arthritis Care Res (Hoboken) 2014;66:617–24.
3. Jacobson DL, Gange SJ, Rose NR, Graham NM. Epidemiology and estimated population burden of selected autoimmune diseases in the United States. Clin Immunol Immunopathol 1997;84:223–43.

4. 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.
5. Alarcón GS, McGwin G Jr, Bastian HM, Roseman J, Lisse J, Fessler BJ, et al, for the LUMINA Study Group. Systemic lupus erythematosus in three ethnic groups. VIII. Predictors of early mortality in the LUMINA cohort. Arthritis Rheum 2001;45:191–202.
6. Bruce IN, O'Keeffe AG, Farewell V, Hanly JG, Manzi S, Su L, et al. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the Systemic Lupus International Collaborating Clinics (SLICC) Inception Cohort. Ann Rheum Dis 2015;74:1706–13.
7. Uramoto KM, Michet CJ Jr, Thumboo J, Sunku J, O'Fallon WM, Gabriel SE. Trends in the incidence and mortality of systemic lupus erythematosus, 1950–1992. Arthritis Rheum 1999;42:46–50.
8. Hanly JG, O'Keeffe AG, Su L, Urowitz MB, Romero-Diaz J, Gordon C, et al. The frequency and outcome of lupus nephritis: results from an international inception cohort study. Rheumatology (Oxford) 2016;55:252–62.
9. Feldman CH, Broder A, Guan H, Yazdany J, Costenbader KH. Sex differences in health care utilization, end-stage renal disease, and mortality among Medicaid beneficiaries with incident lupus nephritis. Arthritis Rheumatol 2018;70:417–26.
10. Levy RA, Vilela VS, Cataldo MJ, Ramos RC, Duarte JL, Tura BR, et al. Hydroxychloroquine (HCQ) in lupus pregnancy: double-blind and placebo-controlled study. Lupus 2001;10:401–4.
11. Feldman CH, Xu C, Williams J, Collins JE, Costenbader KH. Patterns and predictors of recurrent acute care use among Medicaid beneficiaries with systemic lupus erythematosus. Semin Arthritis Rheum 2020;50:1428–36.
12. Feldman CH, Costenbader KH, Solomon DH, Subramanian SV, Kawachi I. Area-level predictors of medication nonadherence among US Medicaid beneficiaries with lupus: a multilevel study. Arthritis Care Res (Hoboken) 2019;71:903–13.
13. Sun K, Eudy AM, Criscione-Schreiber LG, Sadun RE, Rogers JL, Doss J, et al. Racial disparities in medication adherence between African American and Caucasian patients with systemic lupus erythematosus and their associated factors. ACR Open Rheumatol 2020;2:430–7.
14. Petri M. Hydroxychloroquine use in the Baltimore Lupus Cohort: effects on lipids, glucose and thrombosis. Lupus 1996;5 Suppl 1:S16–22.
15. Alarcón GS, McGwin G, Bertoli AM, Fessler BJ, Calvo-Alén J, Bastian HM, et al. Effect of hydroxychloroquine on the survival of patients with systemic lupus erythematosus: data from LUMINA, a multiethnic US cohort (LUMINA L). Ann Rheum Dis 2007;66:1168–72.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1269–1276
DOI 10.1002/acr.24574

American College of Rheumatology
Empowering Rheumatology Professionals

Physician Global Assessment as a Disease Activity Measure for Relapsing Polychondritis

Emily Rose,[1] Marcela A. Ferrada,[1] Kaitlin A. Quinn,[1] Wendy Goodspeed,[1] Laurent Arnaud,[2] Aman Sharma,[3] Hajime Yoshifuji,[4] Jeff Kim,[5] Clint Allen,[5] Arlene Sirajuddin,[6] Marcus Chen,[6] and Peter C. Grayson[1]

Objective. Relapsing polychondritis (RP) is a systemic inflammatory disorder of cartilage that lacks validated disease activity measures. Our objective was to test physician global assessment (PhGA), a measure of disease activity commonly used in rheumatic diseases, in a cohort of patients with RP, which has not been done before.

Methods. Adult patients in an observational cohort of RP underwent standardized, comprehensive evaluation at approximately 6-month intervals. PhGA was scored by 3 physicians from the evaluating institution on a scale of 0–10 for each visit. A random subset of 20 visits was scored by 3 independent physicians not affiliated with the evaluating institution. Treatment change between consecutive visits was categorized as increased, decreased, or unchanged.

Results. In total, 78 patients were evaluated over 164 visits. The intraclass correlation coefficient ($ICC_{2,1}$) for the 3 raters from the evaluating institution was excellent (0.79 [95% confidence interval (95% CI) 0.73, 0.84]) but was poor in the subset of cases scored by the additional raters ($ICC_{2,1}$ 0.27 [95% CI –0.01, 0.53]). Median PhGA was 3 (range 0–7). PhGA weakly correlated with C-reactive protein level ($r_s = 0.30$, $P < 0.01$). In response to increased treatment, median PhGA decreased from 3 (interquartile range [IQR] 2, 4) to 2 (IQR 2, 3) ($P < 0.01$) but rarely went to 0.

Conclusion. Within a single center, PhGA can be used to quantify disease activity and monitor disease response in RP. Persistent disease activity despite treatment, rather than a relapsing-remitting pattern, is observed for most patients with RP. Reliability of PhGA may not generalize across different institutions. A validated disease-specific activity index is needed in RP.

INTRODUCTION

Relapsing polychondritis (RP) is a rare systemic inflammatory disorder that affects multiple organs, with a predilection for cartilaginous structures such as the ear, nose, airway, and joints (1). RP can also affect the eyes, central nervous system, vasculature, inner ear, and skin (1,2). Given the rarity of the disease, clinical assessment has not been standardized, and the disease lacks validated measures of disease activity (3).

Clinical assessment tools are useful to measure disease activity and treatment response in clinical trials and in daily practice. The Relapsing Polychondritis Disease Activity Index (RPDAI) is a proposed tool to quantify disease activity in RP (4). The RPDAI was developed using clinical vignettes rather than patient data and has not been validated in independent cohorts of patients with RP. When deriving weights for items within the RPDAI, the physician global assessment (PhGA) score was used as the gold standard to quantify disease activity, yet the PhGA has also never been systematically studied or validated for use in RP.

The PhGA is frequently used to measure disease activity and track response to treatment in a variety of rheumatic

ClinicalTrials.gov identifier: NCT02257866.

Supported by the Intramural Research Program at the National Institute of Arthritis and Musculoskeletal and Skin Diseases and the NIH Medical Research Scholars Program (a public-private partnership supported jointly by the NIH and contributions to the Foundation for the NIH from the Doris Duke Charitable Foundation, Genentech, the American Association for Dental Research, the Colgate-Palmolive Company, and other private donors).

[1]Emily Rose, BS, Marcela A. Ferrada, MD, Kaitlin A. Quinn, MD, Wendy Goodspeed, RN, Peter C. Grayson, MD, MSc: National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH, Bethesda, Maryland; [2]Laurent Arnaud, MD, PhD: Hôpitaux Universitaires de Strasbourg, Strasbourg, France; [3]Aman Sharma, MD: Postgraduate Institute of Medical Education and Research, Sector-12, Chandigarh, India; [4]Hajime Yoshifuji, MD, PhD: Graduate School of Medicine, Kyoto University, Kyoto, Japan; [5]Jeff Kim, MD, Clint Allen, MD: National Institutes on Deafness and Other Communication Disorders, NIH, Bethesda, Maryland; [6]Arlene Sirajuddin, MD, Marcus Chen, MD: National Heart, Lung, and Blood Institute, NIH, Bethesda, Maryland.

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

Address correspondence to Peter C. Grayson, MD, MSc, National Institutes of Health/NIAMS, 10 Center Drive, Building 10, 10N Rm 311D, Bethesda, MD 20892. Email: peter.grayson@nih.gov.

Submitted for publication July 21, 2020; accepted in revised form February 2, 2021.

SIGNIFICANCE & INNOVATIONS

- Physician global assessment (PhGA) can be used to monitor disease activity in relapsing polychondritis (RP).
- Acute-phase reactants are not elevated in most patients with RP and weakly correlate with PhGA.
- Persistently active disease is more common than a relapsing-remitting pattern of disease activity in RP.
- Despite treatment, few patients with RP achieve a PhGA score of 0, indicating an absence of disease activity.

diseases (5–14). Because performance characteristics of the PhGA may vary by disease, the PhGA should be tested specifically in a cohort of patients with RP. Therefore, this study sought to characterize PhGA to measure disease activity using data from a prospective, observational cohort of patients with RP.

PATIENTS AND METHODS

Study population. Patients age ≥18 years who fulfilled existing diagnostic criteria for RP were recruited into a prospective, observational cohort at the National Institutes of Health (NIH) Clinical Center from August 2016 to October 2019 (15–17). Patients were recruited at any point in the disease course from different countries with no geographic or demographic restrictions. Consecutive patients with clinical visits to the NIH in the time period stated were included. Patients were evaluated by an investigative team with expertise in RP at approximately 6-month intervals. Written informed consent was obtained from all patients, and the study was approved by local ethics review at the NIH.

Data collection. At each visit, patients underwent a standardized comprehensive evaluation that included a clinical rheumatology and otolaryngology evaluation, with direct laryngoscopy, laboratory studies, audiology, echocardiogram, chest imaging with a dynamic computed tomography scan, and any additional clinically appropriate evaluations such as ophthalmology or pulmonology. Clinical signs and symptoms related to common organ involvement due to RP were recorded as either present or absent within the past month. Blood was collected on the day of the study visit for laboratory assessments, including erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) level, and complement levels. All laboratory testing was performed in the NIH Department of Laboratory Medicine. Elevated acute-phase reactants were defined based on levels above the laboratory normal range, which included CRP level 0–4.99 mg/liter, ESR 0–42 mm/hour for females and 0–25 mm/hour for males.

PhGA scoring. From the data collected at each patient visit, de-identified clinical vignettes were created that summarized current symptoms, physical examination findings, laboratory results, imaging findings, and other diagnostics (e.g., audiogram) at each visit. Using these clinical vignettes derived from patient data, 3 practicing rheumatologists from the evaluating institution (MAF, KAQ, and PCG) with expertise caring for patients with RP scored the PhGAs for each study visit blinded to each other. The amount of direct clinical observation of patients in the cohort differed between the raters, and 1 rater only directly evaluated <10% of the patients in the cohort. PhGA was scored on a scale of 0 (defined as the absence of disease activity) to 10 (defined as maximum disease activity), based on an overall assessment of disease activity in the past month. Raters were instructed to consider only symptoms directly attributable to disease activity rather than damage. The raters agreed that a score of 0 would define clinical remission. Ratings were analyzed for discordance, defined as a difference of 3 or more points on the PhGA between any 2 raters (18). For visits with concordant ratings, a final PhGA was assigned based on averaging the 3 raters and rounding to the nearest whole number. Discordant PhGAs were adjudicated by group discussion among the raters.

To compare frequencies of organ involvement across levels of PhGA, PhGAs were categorized on the following scale: 0, 1 or 2, 3 or 4, and ≥5. To determine whether acute-phase reactants impacted the rating of PhGA, a subset of 60 randomly selected study visits were scored a second time, except that ESR and CRP values were withheld. Scores when ESR and CRP were withheld were compared to scores when this information was available. To describe the cumulative number of symptoms, a disease activity summary score was created. The disease activity summary score was defined as the total number of active disease manifestations on a scale of 0–9 that a patient experienced in the previous month of the following 9 items: auricular chondritis, nasal chondritis, chondritis of the chest wall, cardiovascular involvement, respiratory chondritis, eye involvement, hearing loss, vestibular dysfunction, and joint involvement. Each item equaled 1 point.

To assess interrater reliability of PhGA among physicians who practice outside of the evaluating institution, clinical vignettes from 20 unique patients (25% of the cohort) were randomly selected and independently scored by 3 physicians (LA, AS, and HY) from different regions of the world (France, India, and Japan, respectively). Each of these raters has >10 years of experience caring for patients with RP, and none of these raters directly evaluated any of the patients in this study.

Response to treatment. Response to treatment was analyzed in patients who provided data from at least 2 study visits. A detailed history of immunosuppressive therapy, including glucocorticoids, synthetic disease-modifying antirheumatic drugs (DMARDs; including targeted synthetic agents [JAK inhibitors]),

and biologic DMARDs was recorded at each visit. Treatment decisions were made by the referring physicians rather than the investigative team. The PhGA raters were blinded to treatment data. Treatment change between consecutive visits was categorized as increased, decreased, or unchanged. A change in treatment was defined as any of the following: a daily prednisone dose change of >5 mg, the addition or removal of a DMARD therapy, a decrease in dose of DMARD therapy ≥50%, or an increase in dose of DMARD therapy ≥50% at the time of the follow-up visit relative to the previous visit (13).

Statistical analysis. Demographic and clinical variables were expressed as median and interquartile range (IQR), or frequencies, according to data type. The intraclass correlation coefficient ($ICC_{2,1}$) and 95% confidence intervals (95% CIs) were calculated using the icc function of the R package irr to determine interrater reliability between the PhGA raters. This model was chosen to assess interrater reliability because of its feasibility with >2 raters (18). ICC results were interpreted as follows: <0.4 = poor, 0.40–0.59 = fair, 0.60–0.74 = good, and 0.75–1.0 = excellent (19). Spearman's rank correlation analysis was used to evaluate the strength of association between PhGA and continuous variables and the correlation between different raters. Organ-specific signs and symptoms of RP were compared between the proportion of visits with PhGA ≤2 and the proportion of visits with PhGA ≥3 using Fisher's exact test. Wilcoxon's signed rank test was used to compare changes in PhGA and CRP level in association with treatment status and with addition/increase or subtraction/decrease of specific categories of medications. The chi-square test was used to compare the proportion of patients who increased treatment in the first visit interval compared to other visit intervals. *P* values less than 0.05 were considered significant. All statistical analysis was performed using JMP 14, PRISM 8, or RStudio 1.2.5001.

RESULTS

Characteristics of the cohort. In all, 78 patients were evaluated over 164 study visits over a median follow-up period of 6 months (IQR 0, 16). Forty-seven patients (60.3%) had at least 1 follow-up visit, with a median follow-up interval of 6 months (IQR 5, 9). The cohort was predominantly Caucasian (82.0%) and female (83.3%). The median age at the baseline visit was 45 years (IQR 36, 55), and the median disease duration from symptom onset was 8 years (IQR 4, 16). The median CRP level was 2.2 (IQR 0.7, 6.9) mg/liter at the baseline visit and the ESR was 10.5 (IQR 5, 20.8) mm/hour. Most patients (78.2%) were receiving treatment at their first visit. Demographics and clinical characteristics of the participants at the baseline visit are shown in Table 1.

PhGA ratings. Of 164 study visits, the median for PhGA ratings by each of the 3 raters at the evaluating institution was 3 (IQR 2, 4), 2 (IQR 1, 3), and 2.5 (IQR 2, 3). For 6 of 164 visits (3.7%), the ratings by the 3 raters were discordant. After adjudicating the discordant ratings, the median assigned PhGA was 3 (IQR 2, 3), ranging from 0 to 7. The distribution of the final assigned PhGA score is shown in Figure 1A. Patients had PhGA scores of 0 at 4 visits (2.4%), scores of 1–2 at 76 visits (46.3%), scores of 3–4 at 62 visits (37.8%), and scores of ≥5 at 22 visits (13.4%).

Table 1. Baseline study population characteristics (n = 78)*

Characteristic	Value
Caucasian	64 (82.1)
Female	65 (83.3)
Age at visit, median (IQR) years	45 (36, 55)
Disease duration, median (IQR) years	8 (4, 16)
Number of visits, median (IQR)	2 (1, 3)
Follow-up period, median (IQR) months	6 (0, 16)
Laboratory	
CRP, median (IQR) mg/liter†	2.2 (0.7, 6.9)
ESR, median (IQR) mm/hour‡	10.5 (5, 20.8)
Elevated acute-phase reactants	28 (36.8)
Symptom frequencies at visit attributed to RP in 28 days	
Musculoskeletal	65 (83.3)
Arthralgia or stiffness	57 (73.1)
Tenosynovitis or synovitis	25 (33.3)
Nasal chondritis	34 (43.6)
Auricular chondritis	19 (24.4)
Chest wall chondritis	42 (53.8)
Audiovestibular	2 (2.6)
Vestibular dysfunction	0
Sensorineural hearing loss	2 (2.6)
Cutaneous	2 (2.6)
Respiratory chondritis§	44 (56.4)
Wheezing	3 (3.8)
Voice changes	17 (21.8)
Dry cough	24 (30.8)
Dyspnea	29 (37.2)
Cardiovascular involvement	1 (1.3)
Eye involvement	0
Subglottic inflammation	7 (10.2)
Tracheal wall thickness ≥3 mm	18 (24)
Tracheomalacia	31 (42.3)
Bronchomalacia	10 (13.7)
Current treatment	
No treatment	17 (21.8)
Prednisone dose, median (IQR) mg	5 (0, 20)
At least 1 synthetic DMARD	44 (56.4)
At least 1 biologic DMARD	17 (21.8)
Activity scores, median (IQR)	
Physician global assessment	3 (2, 3)
Disease activity summary score	3 (2, 3)

* Values are the number (%) unless indicated otherwise. Missing values: subglottic inflammation (n = 9), tracheal wall thickness (n = 5), tracheomalacia (n = 5), bronchomalacia (n = 5). CRP = C-reactive protein; DMARD = disease-modifying antirheumatic drug; ESR = erythrocyte sedimentation rate; IQR = interquartile range.
† Laboratory normal range: 0–4.99 (n = 77).
‡ Laboratory normal range: 0–42 mm/hour for females and 0–25 mm/hour for males (n = 76).
§ Respiratory chondritis includes symptoms, not objective findings.

Interrater reliability of PhGA ratings. When comparing ratings across all study visits by the 3 physicians from the

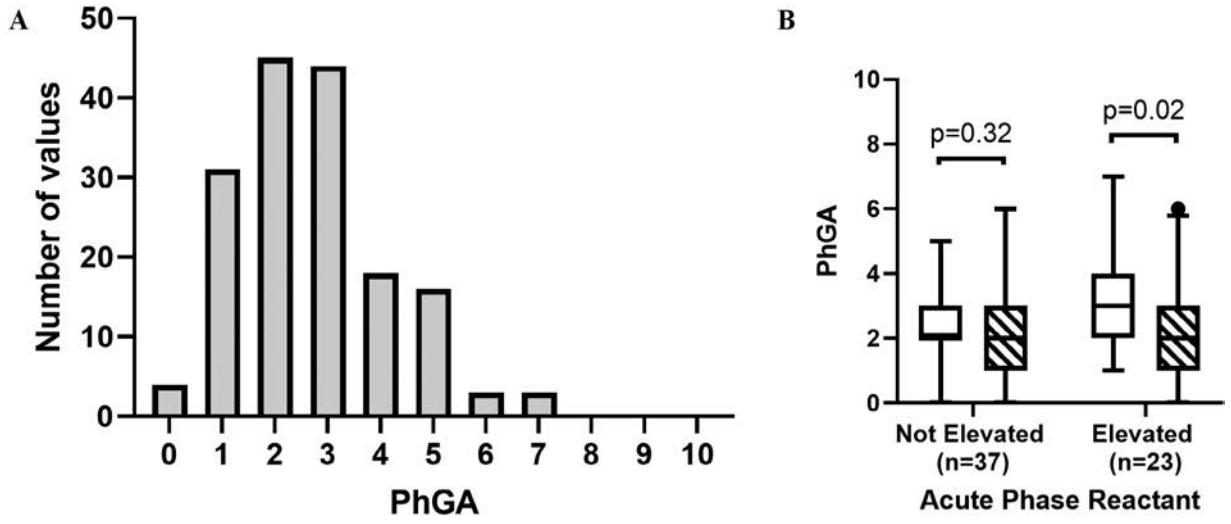


Figure 1. Distribution of physician global assessment (PhGA) scores in relapsing polychondritis. **A**, Histogram shows the distribution of PhGA scores for 164 visits (median 3 [interquartile range 2, 3]). **B**, Box and whisker plot of PhGA scores when information about erythrocyte sedimentation rate and C-reactive protein level was provided to the raters (no stripes) or withheld from raters (stripes). Whiskers show 5th to 95th percentile, 60 visits included.

evaluating institution, the $ICC_{2,1}$ was excellent at 0.79 (95% CI 0.73, 0.84). For the 20 visits randomly selected for additional review, the $ICC_{2,1}$ was excellent at 0.85 (95% CI 0.72, 0.93) for the 3 physicians at the evaluating institution but was poor at 0.27 (95% CI −0.01, 0.53) for the additional raters from outside institutions. The $ICC_{2,1}$ for all 6 raters in the subset of 20 visits was poor at 0.21 (95% CI 0.06, 0.44). A scatter plot of the ratings for all 6 raters is shown in Figure 2A. There was a significant positive correlation ($P < 0.01$) between each of the 3 raters from the evaluating institution. No correlation was observed between the outside raters with each other or with the raters from the evaluating institution (Figure 2B).

Laboratory variables and PhGA. ESR was elevated above the laboratory normal range at 22 of 162 visits (13.6%), and CRP level was elevated at 44 of 163 visits (27.0%). In each case, elevated acute-phase reactants were believed to be associated with disease activity and not related to a secondary process. The PhGA score was weakly correlated with CRP level ($r_s = 0.30$, $P < 0.01$) but not with ESR ($r_s = 0.13$, $P = 0.10$). Similar results were found when analyzing only data from the baseline visits (data not shown).

To investigate the contribution of acute-phase reactants (ESR and CRP) to PhGA scores, 60 clinical vignettes were rescored without providing values for acute-phase reactants. For visits when at least 1 of the acute-phase reactants was elevated, the median PhGA rating was 3 (IQR 2, 4) when the acute-phase reactant information was provided; however, repeat PhGA ratings were significantly lower, with a median of 2 (IQR 1, 3), when information about acute-phase reactants was omitted ($P = 0.02$) (Figure 1B). For visits with normal values for the acute-phase reactants, there was no difference in PhGA ratings when information about these tests was available (2 [IQR 2, 3]) compared to when the information was withheld (2 [IQR 1, 3]; $P = 0.32$) (Figure 1B). While CRP level positively correlated with PhGA ratings when this information was provided ($r_s = 0.25$, $P = 0.05$) (see Supplementary Figure 1A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24574), CRP level did not correlate with the PhGA score when information about acute-phase reactants was withheld ($r_s = -0.013$, $P = 0.92$) (see Supplementary Figure 1B, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24574).

Association between clinical symptoms and PhGA. To determine signs and symptoms that contribute to PhGA ratings, symptom frequencies were assessed. PhGA scores did not significantly differ based on the presence or absence of most symptoms, except that significantly greater PhGA scores were observed in association with arthralgia/stiffness and wheezing (Table 2). There was no difference in symptom frequencies assessed between visits with PhGA ≤2 and with PhGA ≥3 (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24574).

The disease activity summary score ranged from 0 to 5 with a median of 2 (IQR 1, 3). Across all visits, the disease activity summary score was not correlated to PhGA score ($r_s = 0.087$, $P = 0.27$). However, for the 60 visits that were rescored without acute-phase reactant information provided to the raters, PhGA ratings were significantly correlated with the disease activity summary score ($r_s = 0.51$, $P < 0.01$) (see Supplementary Figures 1C and 1D, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24574).

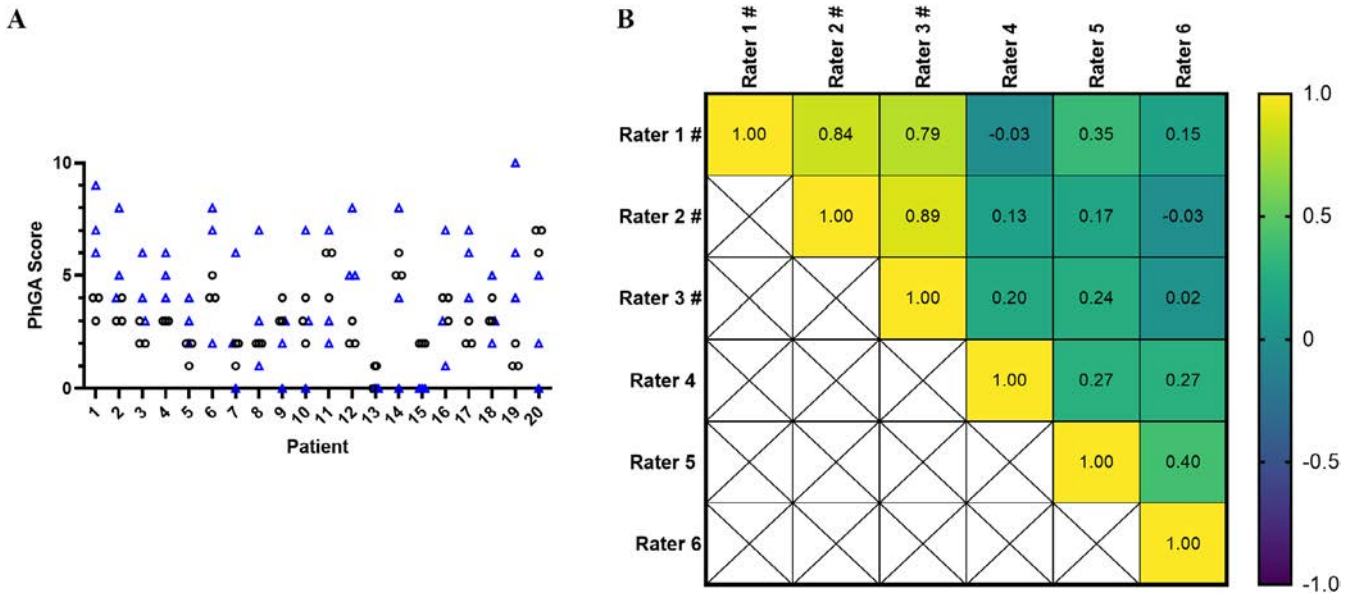


Figure 2. Physician global assessment (PhGA) scores of clinical vignettes by 6 independent raters. **A**, Scatter plot shows the PhGA scores for 20 patients by 3 raters from the evaluating institution (black circles) and 3 additional outside raters (blue triangles). **B**, Correlation matrix shows the Spearman's correlation between each the 6 raters for the 20 patient visits. # = rater from the evaluating institution. Scores were significantly correlated ($P < 0.01$) only between the 3 raters from the evaluating institution. Color bar represents the degree of correlation between raters.

Change in PhGA over time. Of the 86 follow-up visits from 47 patients, the interval changes in PhGA scores ranged from –5 to 3. Between consecutive visits, the median change was 0 (IQR –1, 1). PhGA scores increased by 2 or more between 12 visit intervals (14.0%), decreased by 2 or more between 16 visit intervals (18.6%), and remained unchanged for 30 visit intervals (Figure 3A). The median for the first 4 study visits were 3 (IQR 2, 5), 2 (IQR 1, 3), 2 (IQR 2, 3), and 2 (IQR 1, 3), respectively. There was a significant decrease between the baseline visit compared to each of the next 3 follow-up visits ($P < 0.01$). There were no significant differences in PhGA scores between any of the follow-up visits (Figure 3B).

Change of PhGA in response to treatment. Of the 86 follow-up visits from 47 patients, there was an increase in treatment over 55 visits, a decrease in treatment over 14 visits, and no change in treatment over 17 visits. The proportion of visits with increased treatment was greater in the first visit interval (74.5%) compared to all other intervals ($P = 0.03$). Patients who had an increase in treatment between visits showed a decrease in PhGA score from 3 (IQR 2, 4) to 2 (IQR 2, 3; $P < 0.01$) (Figure 3C). PhGA did not change between visits in patients who decreased treatment or who had no change in treatment. For visit intervals with no change in treatment, PhGA ratings stayed the

Table 2. Median PhGA for visits with symptom present compared to visits when symptom not present*

Symptom	No PhGA symptom	PhGA symptom	*P*
Musculoskeletal involvement	2.5 (1, 3)	3 (2, 4)	0.57
Arthralgia or stiffness	2 (1, 3)	3 (2, 4)	0.024
Tenosynovitis or synovitis	2.5 (1, 4)	3 (2, 3)	0.49
Nasal chondritis	2 (2, 3.25)	3 (1.75, 3.25)	0.91
Auricular chondritis	2 (2, 3)	3 (2, 4)	0.45
Chest wall chondritis	2 (2, 3)	3 (2, 4)	0.42
Audiovestibular	3 (2, 3)	2 (1.75, 4.25)	0.88
Vestibular dysfunction	3 (2, 3.25)	2 (2, 2)	0.45
Sensorineural hearing loss	3 (2, 3)	3 (1.4, 4.5)	0.94
Cardiovascular involvement	3 (2, 3)	3 (3, 3)	0.66
Eye involvement	3 (2, 3.5)	2 (2, 2)	0.36
Respiratory chondritis	2 (2, 3)	3 (2, 4)	0.21
Wheezing	2.5 (2, 3)	3.5 (2.25, 5)	0.045
Voice changes	3 (2, 3)	3 (1.75, 4)	0.61
Dry cough	3 (2, 3)	2 (1.75, 4)	0.99
Dyspnea	2 (2, 3)	3 (2, 4)	0.12
Choking sensation	3 (2, 3.5)	3 (2, 3)	0.87

* Values are the median (interquartile range) unless indicated otherwise. PhGA = physician global assessment.

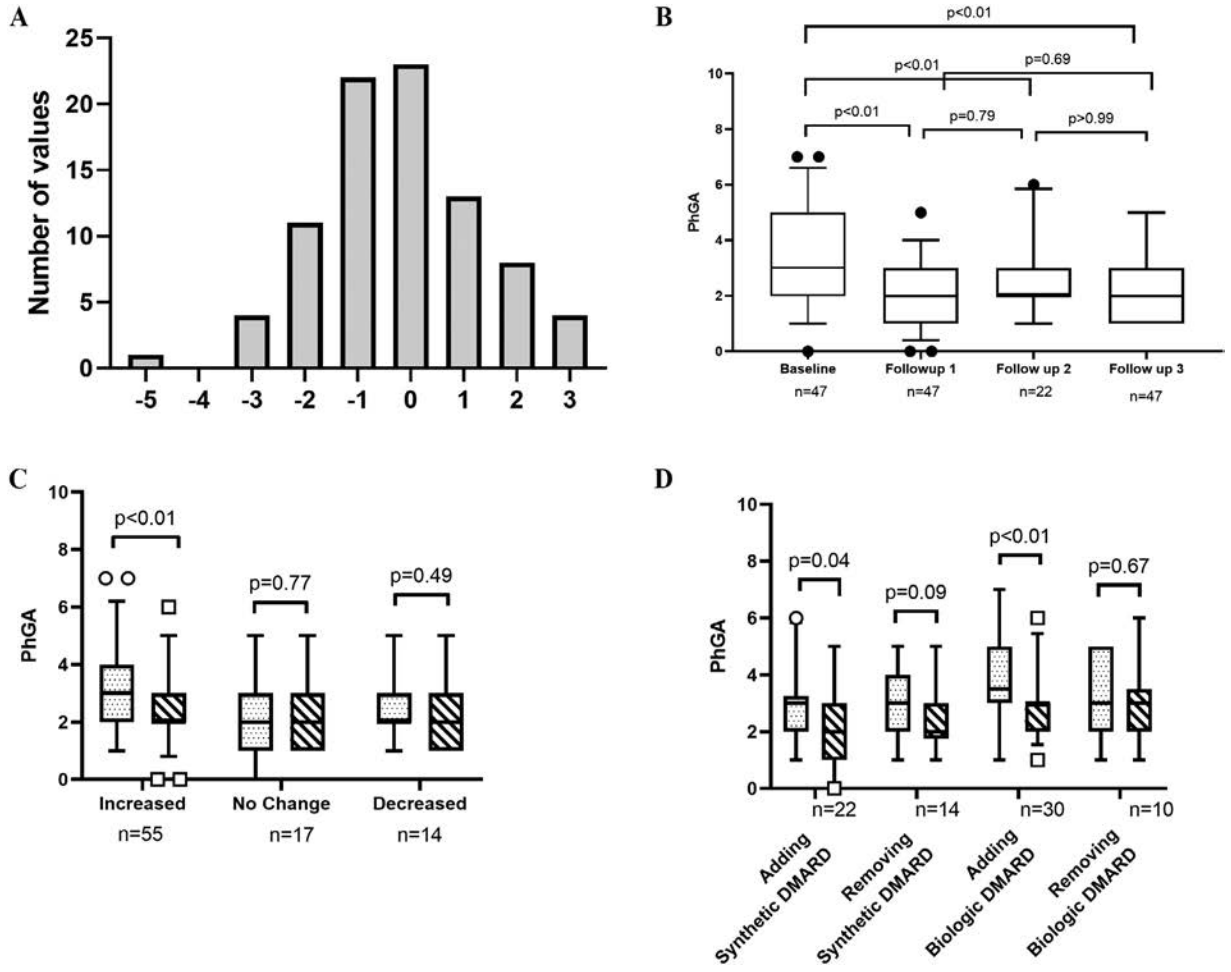


Figure 3. Interval change in physician global assessment (PhGA) score in association with treatment. **A**, Histogram shows change in PhGA scores between consecutive visits over 86 visit intervals. **B**, Box and whisker plots of PhGA scores from 4 consecutive study visits in 47 patients with at least 2 study visits. The medians for the first 4 visits were 3 (interquartile range [IQR] 2, 5), 2 (IQR 1, 3), 2 (IQR 2, 3) and 2 (IQR 1, 3), respectively. **C**, Box and whisker plots of PhGA scores from 2 consecutive study visits categorized by an increase in treatment, no change in treatment, or a decrease in treatment between visits. Dots = visit A; stripes = visit B. **D**, Box and whisker plots of PhGA scores from consecutive study visits in relation to specific treatment changes between visits. DMARD = disease-modifying antirheumatic drug; dots = visit A; stripes = visit B; black dots (**B**) = outlier values; white circles (**C** and **D**) = outlier values during initial visit; white squares (**C** and **D**) = outlier values on follow up visit.

same (2 [IQR 1, 3]; $P = 0.77$), and for visits that decreased in treatment, PhGA scores went from 2 (IQR 2, 3) to 2 (IQR 1, 3; $P = 0.49$). Patients who increased treatment had greater median PhGA scores at the visit before the treatment change compared to patients who did not change treatment: 3 (IQR 2, 4) versus 2 (IQR 1, 3; $P = 0.02$). There were no other significant differences in PhGA scores between the pre- or posttreatment groups. Similarly, there was a significant decrease in median CRP levels following an increase in treatment from 2.3 (IQR 0.6, 4.4) mg/liter to 0.9 (IQR 0.3, 4) mg/liter ($P = 0.04$) but no change in CRP level with either decreasing or not changing treatment. There were no changes in ESR between visits with increasing, decreasing, or not changing treatment. Adding either a synthetic (including targeted synthetic) DMARD or a biologic DMARD was associated with a significant decrease in PhGA ratings between visits, from 3 (IQR 2, 3.25) to 2 (IQR 1, 3) and from 3.5 (IQR 3, 5) to 3 (IQR 2, 3) ($P = 0.04$ and $P < 0.01$, respectively) (Figure 3D). Withdrawal of a synthetic or biologic DMARD was not associated with a change in PhGA score. Prednisone dose was not correlated with PhGA score ($r_s = -0.087$, $P = 0.27$).

DISCUSSION

This study was undertaken to characterize PhGA for RP, a rare disease that lacks validated outcome measures. PhGA was not a reliable measure of disease activity among raters from different institutions, highlighting a need to develop and validate disease-specific activity indices in RP. Despite this limitation, PhGA was reliable and useful to monitor disease activity in response to treatment within a single-center observational cohort study of patients with RP. Persistent disease activity was common in RP despite treatment, and most patients with

RP did not experience a relapsing-remitting pattern of disease activity.

The factors that influence PhGA ratings in RP are complex and can lead to differences in clinical assessment of disease activity. For example, accurate determination of airway inflammation based on clinical symptoms and direct observation of only the upper airway can be challenging. In this study, very few individual signs or symptoms were independently associated with PhGA score, highlighting the fact that disease activity assessment in RP is likely driven by multifaceted factors. Issues of study design may contribute to the lack of strong associations between PhGA score and individual clinical symptoms. Patients were often evaluated later in the disease course while receiving treatment, which may have blunted associations between individual symptoms and disease activity assessment. Additionally, clinical symptoms were categorized as present or absent. The degree of severity of individual symptoms may have been more strongly associated with PhGA score.

Multiple measures point to a moderate influence of ESR and CRP level, particularly when elevated, on PhGA ratings. CRP level weakly correlated with PhGA score. However, at most visits, patients had normal acute-phase reactants despite ongoing symptoms attributable to RP. Without the acute-phase reactant information, symptom-based variables appear to carry more weight in the PhGA rating. Because in RP differentiating between disease activity and damage can be difficult, knowledge of the acute-phase reactant levels influences to what extent the rater considers clinical symptoms as representative of active disease. These findings are consistent with findings in other rheumatic diseases, such as systemic lupus erythematosus, where laboratory data have been shown to significantly impact PhGA assessment (20). Together, these data show that laboratory markers interplay with other factors in the assessment of disease activity in RP.

The distribution of PhGA scores provides insight into disease activity profiles in patients with RP. In this study, most patients had PhGA scores between 1 and 3. There were few patients with PhGA scores >3 and only a small number of patients who achieved a PhGA score of 0. The small number of patients with a PhGA score of 0 points to the difficulty of completely eliminating disease activity, even with treatment. Selection bias may affect the distribution of PhGA scores, as patients on either extreme on the PhGA scale may have been less likely to enroll and travel to the NIH to participate in this study. In some cases, when clinical symptoms were minor, distinguishing between other etiologies rather than true disease activity may have been difficult, which is affirmed by the rating differences of PhGA when information about elevated acute-phase reactants is provided versus withheld. Differentiating activity from damage remains an important challenge in the assessment and care of patients with RP.

Examination of PhGA scores over time provides insight into the natural history of RP and treatment response. Based on change in PhGA ratings over consecutive visits, RP worsens in only a few patients on treatment, improves slightly in many patients with treatment, and stays the same only in those patients who already had low disease activity. Contrary to the disease name, relapsing polychondritis, patients with multiple visits did not show a relapsing-remitting pattern. Rather, there was a significant decrease in disease activity between the first and second visits, and most patients remained relatively stable over subsequent visits, with ongoing persistent mild disease activity despite treatment. Treatment escalation most often occurred after the initial study visit. Increasing treatment with either synthetic, targeted synthetic, or biologic DMARDs was associated with a lower PhGA score at subsequent follow-up visits, but PhGA score was not correlated with glucocorticoid dose. Improvement in PhGA following increased treatment suggests that aggressive pharmacologic measures for patients with active disease does help, as suggested by others (21). Alternatively, improvement in PhGA score only during the initial observation interval could represent regression to the mean or the natural history of disease.

This study has some limitations. This was not an inception cohort; therefore, patients were enrolled at various stages of disease and many had RP for years. This limitation is somewhat countered by a relatively large sample size for a rare disease, prospective data collection, and a standardized assessment protocol. Studying patients later in the disease course may have restricted the range of disease activity observed in this study. Cumulative glucocorticoid dosing was not examined, which could explain the lack of association of PhGA score and glucocorticoid dose. While this study begins to explore the association of laboratory and symptom-based factors with PhGA ratings in RP, the composite drivers of PhGA scores are yet to be determined.

Validation of a disease-specific activity index is an unmet need in RP. While PhGA is simple to implement, it is a generic measure of disease activity that lacks the nuance of a disease-specific index. A disease-specific activity index may standardize activity assessment in RP and improve agreement among investigators from different institutions. The RPDAI is a proposed disease-specific tool to quantify disease activity in RP, and efforts are currently underway to modify and validate this index (4).

Assessment of the PhGA in a cohort of patients with RP is a necessary step toward the goal of outcome measure development and validation in RP. PhGA is useful to monitor disease activity in RP but may not be reliable across different institutions. The complexity of factors influencing PhGA ratings highlights the challenges of clinical assessment in RP. Development of disease-specific outcome measures is a prerequisite to the successful conduct of much-needed randomized clinical trials in RP.

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. Grayson 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. Rose, Ferrada, Grayson.

Acquisition of data. Rose, Ferrada, Quinn, Goodspeed, Arnaud, Sharma, Yoshifuji, Kim, Allen, Sirajuddin, Chen, Grayson.

Analysis and interpretation of data. Rose, Ferrada, Quinn, Arnaud, Sharma, Yoshifuji, Grayson.

REFERENCES

1. Mathian A, Miyara M, Cohen-Aubart F, Haroche J, Hie M, Pha M, et al. Relapsing polychondritis: a 2016 update on clinical features, diagnostic tools, treatment and biological drug use. Best Pract Res Clin Rheumatol 2016;30:316–33.
2. Puechal X, Terrier B, Mouthon L, Costedoat-Chalumeau N, Guillevin L, Le Jeunne C. Relapsing polychondritis. Joint Bone Spine 2014; 81:118–24.
3. Rednic S, Damian L, Talarico R, Scire CA, Tobias A, Costedoat-Chalumeau N, et al. Relapsing polychondritis: state of the art on clinical practice guidelines. RMD Open 2018;4 Suppl 1:e000788.
4. Arnaud L, Devilliers H, Peng SL, Mathian A, Costedoat-Chalumeau N, Buckner J, et al. The Relapsing Polychondritis Disease Activity Index: development of a disease activity score for relapsing polychondritis. Autoimmun Rev 2012;12:204–9.
5. Mukhtyar C, Lee R, Brown D, Carruthers D, Dasgupta B, Dubey S, et al. Modification and validation of the Birmingham Vasculitis Activity Score (version 3). Ann Rheum Dis 2009;68:1827–32.
6. Griffiths B, Mosca M, Gordon C. Assessment of patients with systemic lupus erythematosus and the use of lupus disease activity indices. Best Pract Res Clin Rheumatol 2005;19:685–708.
7. Rohekar G, Pope J. Test-retest reliability of patient global assessment and physician global assessment in rheumatoid arthritis. J Rheumatol 2009;36:2178–82.
8. Cauli A, Gladman DD, Mathieu A, Olivieri I, Porru G, Tak PP, et al. Physician's global assessment in psoriatic arthritis: a multicenter GRAPPA study. J Rheumatol 2018;45:1256–62.
9. Hinojosa-Azaola A, Jimenez-Gonzalez A, Alcocer-Castillejos N. Patient and physician perspectives on the impact of health-related quality of life in Mexican patients with ANCA-associated vasculitis. Rheumatol Int 2018;38:631–40.
10. Rimland CA, Quinn KA, Rosenblum JS, Schwartz MN, Gribbons KB, Novakovich E, et al. Outcome measures in large-vessel vasculitis: relationship between patient-, physician-, imaging-, and laboratory-based assessments. Arthritis Care Res (Hoboken) 2020;72: 1296–304.
11. Tago M, Sawada T, Nishiyama S, Tahara K, Kato E, Hayashi H, et al. Influence of large joint involvement on patient: physician discordance in global assessment of rheumatoid arthritis disease activity analyzed by a novel joint index. Int J Rheum Dis 2018;21:1237–45.
12. Chessa E, Piga M, Floris A, Devilliers H, Cauli A, Arnaud L. Use of physician global assessment (PGA) in systemic lupus erythematosus: a systematic review of its psychometric properties. Rheumatology (Oxford) 2020;59:3622–32.
13. Banerjee S, Quinn KA, Gribbons KB, Rosenblum JS, Civelek AC, Novakovich E, et al. Effect of treatment on imaging, clinical, and serologic assessments of disease activity in large-vessel vasculitis. J Rheumatol 2020;47:99–107.
14. Kaneko Y, Takeuchi T, Cai Z, Sato M, Awakura K, Gaich C, et al. Determinants of patient's global assessment of disease activity and physician's global assessment of disease activity in patients with rheumatoid arthritis: a post hoc analysis of overall and Japanese results from phase 3 clinical trials. Mod Rheumatol 2018;28:960–7.
15. McAdam LP, O'Hanlan MA, Bluestone R, Pearson CM. Relapsing polychondritis: prospective study of 23 patients and a review of the literature. Medicine (Baltimore) 1976;55:193–215.
16. Michet CJ Jr, McKenna CH, Luthra HS, O'Fallon WM. Relapsing polychondritis: survival and predictive role of early disease manifestations. Ann Intern Med 1986;104:74–8.
17. Damiani JM, Levine HL. Relapsing polychondritis: report of ten cases. Laryngoscope 1979;89 Pt 1:929–46.
18. Trevethan R. Intraclass correlation coefficients: clearing the air, extending some cautions, and making some requests. Health Serv Outcomes Res Methodol 2017;17:127–43.
19. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6:284–90.
20. Aranow C, Askanase A, Oon S, Huq M, Calderone A, Morand EF, et al. Laboratory investigation results influence physician's global assessment (PGA) of disease activity in SLE. Ann Rheum Dis 2020; 79:787–92.
21. Moulis G, Pugnet G, Costedoat-Chalumeau N, Mathian A, Leroux G, Boutemy J, et al. Efficacy and safety of biologics in relapsing polychondritis: a French national multicentre study. Ann Rheum Dis 2018;77:1172–8.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1277–1286
DOI 10.1002/acr.24576

American College of Rheumatology
Empowering Rheumatology Professionals

Relationship Between Changing Body Mass Index and Serum Uric Acid Alteration Among Clinically Apparently Healthy Korean Men

Jiwon Hwang,[1] Mi Yeon Lee,[2] Joong Kyong Ahn,[2] and Hoon-Suk Cha[3]

Objective. Gout and hyperuricemia incidence is increasing worldwide, reflecting pandemic overweight and obesity. However, the magnitude of the association of body mass index (BMI) changes with serum uric acid (UA) level in the general population has remained unevaluated.

Methods. This retrospective cohort study enrolled 27,422 Korean men who underwent a comprehensive health check-up between 2015 and 2017. BMI change was categorized into 7 groups. The relationship between BMI change and serum UA level alteration was determined using multivariable regression models.

Results. The mean age, BMI, and serum UA level were 38.8 years, 24.7 kg/m^2, and 6.2 mg/dl, respectively. All BMI change categories had a clear dose-response relationship with the serum UA level changes. The corresponding β coefficient of serum UA level changes was 0.13 (95% confidence interval [95% CI] 0.11, 0.16), 0.25 (95% CI 0.2, 0.3), and 0.44 (95% CI 0.36, 0.52) for a BMI decrease of 0.5–1.5, 1.5–2.5, and ≥2.5, respectively. Compared with no BMI change, the multivariate odds ratios of achieving normouricemia for a BMI increase of 0.5–1.5, 1.5–2.5, and ≥2.5 were 0.88 (95% CI 0.83, 0.95), 0.67 (95% CI 0.60, 0.75), and 0.60 (95% CI 0.49, 0.74), whereas those for a BMI decrease of 0.5–1.5, 1.5–2.5, and ≥2.5 were 1.17 (95% CI 1.07, 1.27), 1.28 (95% CI 1.08, 1.52), and 1.46 (95% CI 1.13, 1.88), respectively.

Conclusion. BMI change could have a significant association with the alteration of serum UA levels of apparently healthy men. Despite its small effect size, the health risks and benefits of BMI change would be emphasized for serum UA level alteration.

INTRODUCTION

Uric acid is the end product of purine metabolism in humans because unlike most mammals, humans lack uricase (1). Thus, serum uric acid (UA) levels in humans are higher than those in other mammals. Subsequently, some advantageous roles of uric acid in humans have been hypothesized along with the discoveries of high antioxidant capacity and relevantly increased life expectancy, the maintenance of blood pressure (BP) to sustain a vertical position, and higher intelligence and neuroprotection (2).

Despite the physiologic and beneficial expectations for uric acid role, hyperuricemia is the main cause of gout, thereby regarded as a health threat. Furthermore, hyperuricemia has been associated with various diseases, including chronic kidney disease, diabetes mellitus, metabolic syndrome, fatty liver disease, stroke, coronary artery disease, and hypertension (3,4). Hence, the globally increasing prevalence and incidence of hyperuricemia and gout are important issues (5). For instance, the prevalence and incidence of gout in South Korea were predicted to further increase in 2025 (6).

The burden of gout and hyperuricemia may reflect the status of modern society in terms of aging, comorbidities, polypharmacy, and lifestyle and dietary factors including purine-rich food (7–9), and even pandemic overweight and obesity (10,11). Thus, weight control is commonly recommended for gout and asymptomatic hyperuricemia. In fact, recent American and European

[1]Jiwon Hwang, MD, PhD: Samsung Changwon Hospital, Sungkyunkwan University School of Medicine, Changwon, Republic of Korea; [2]Mi Yeon Lee, MPH, PhD, Joong Kyong Ahn, MD, PhD: Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; [3]Hoon-Suk Cha, MD, PhD: Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea.

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

Address correspondence to Joong Kyong Ahn, MD, PhD, Division of Rheumatology, Department of Internal Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University of School of Medicine, 29 Saemunan-ro, Jongno-gu, Seoul 03181, Republic of Korea (email: mdahnjk@skku.edu); or to Hoon-Suk Cha, MD, PhD, Division of Rheumatology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Republic of Korea (email: hoonsuk.cha@samsung.com).

Submitted for publication July 1, 2020; accepted in revised form February 2, 2021.

SIGNIFICANCE & INNOVATIONS

- Over 2 years, more participants experienced a gain in body mass index (BMI) instead of a loss (37.3% versus 19.6%).
- BMI change over 2 years was positively associated with serum uric acid (UA) levels: 1) serum UA level increased as BMI increased, while serum UA level decreased as BMI decreased; and 2) the more that BMI changed, the larger the serum UA level alteration, while the lesser the BMI change, the smaller the serum UA level alteration.
- More loss of BMI was likely to attain the target serum UA levels, both of normouricemia or <6 mg/dl.
- The present study suggests a possible basis for BMI change in serum UA levels among apparently healthy men.

guidelines conditionally have recommended weight loss as a nonpharmacologic measure for gout management (12,13).

Prior studies have had several aspects that failed to explain general weight and uric acid reduction. These aspects are as follows: 1) a specific intervention, such as bariatric surgery or physical activity, was performed (14,15); 2) a special group, such as men with high cardiovascular risk profiles or people with obesity, was targeted (16,17); or 3) the effects were assessed only in people with gout (18). According to a recent Chinese study, body mass index (BMI) is associated with serum UA in the general population in terms of age and sex; subjects with active BMI gain were more likely to develop hyperuricemia than their respective normal-weight controls (19).

The aim of our study was to evaluate the association of weight change (gain or loss) with subsequent alteration of serum UA levels (rise or fall) according to BMI variance in apparently clinically healthy men. A retrospective observational cohort study was conducted using health screening examination data that had been serially checked. The impact of weight change on achieving target serum UA levels was also investigated using 2 different targets: <7 and <6 mg/dl.

SUBJECTS AND METHODS

Study population. We recruited men who consecutively underwent a health examination between 2015 and 2017 in one of the Kangbuk Samsung Hospital Total Healthcare Centers located in Seoul and Suwon in Korea. This regular examination conforms to the South Korea's Industrial Safety and Health Law, which requires annual or biennial free health check-ups for all employees. More than 80% of the study participants were employees of various companies and local governmental organizations; other participants voluntarily engaged in the health check-up program at their own expense.

We screened 78,756 men; however, we excluded those who had underlying conditions such as heart disease (n = 693), hypertension (n = 9,975), diabetes mellitus (n = 3,025), stroke (n = 362), malignancy (n = 1,417), chronic kidney disease (n = 259), and chronic liver disease (n = 40), and those who were taking medication for hypertension, dyslipidemia, gout, and nephrolithiasis (n = 8,386). We also excluded those who did not answer the food frequency questionnaires (FFQ) (n = 38,583) and who had missing values for the continuous covariates (n = 14,142). Thereafter, we categorized the participants according to health check-up interval. Respectively, 27,422 and 13,603 men were eligible for the biennial examination in 2015 and 2017 and the annual examination in 2015, 2016, and 2017, consecutively. The Ethics Committee of the Kangbuk Samsung Hospital approved this study (number KBSMC 2019-09-001), and informed consent was not required because a deidentified database was used to analyze the data retrospectively.

Assessment of serum UA, BMI variance, and other covariates. Comprehensive medical check-ups including laboratory tests and anthropometric measurements (height, weight, BP, and physical activity) were skillfully executed as previously described (20,21). All enrolled participants completed a 103-item, self-reported FFQ, which is a semiquantitative form designated and validated in Korea (22).

We collected the following data: age (years); weight (kg); height (m); smoking status (never, former, and current); alcohol consumption (gm/day); health-enhancing physical activity (HEPA) level; education level (high school or college graduate); medication and medical history; daily dietary intake of total energy (kcal), total protein (gm), total fat (gm), fiber (gm), calcium (mg), and vitamin C (mg) based on the FFQ; and the laboratory results for uric acid, calcium, phosphorus, alkaline phosphatase (AP), low-density lipoprotein (LDL) cholesterol, creatinine, and high-sensitivity C-reactive protein (hsCRP). BMI was calculated as the weight (kg) divided by height squared (m^2); BMI ≥25 kg/m^2, which is the proposed cutoff for Asian populations, indicated obesity (23). HEPA was defined when either of 2 criteria were satisfied: 1) vigorous-intensity activity for ≥3 days per week accumulating in ≥1,500 metabolic equivalent task (MET) minutes/week; or 2) 7 days of any combination of walking, moderate-intensity, or vigorous-intensity activities achieving at least 3,000 MET minutes/week (24). Furthermore, the estimated glomerular filtration rate (eGFR) was calculated using the Modification of Diet in Renal Disease Study equation (25). The homeostasis model of assessment–estimated insulin resistance (HOMA-IR) level was calculated as the fasting insulin (mU/liter) × fasting glucose (mmoles/liter)/22.5.

In this study, hyperuricemia was defined as serum UA ≥7 mg/dl, and normouricemia as serum UA <7 mg/dl. We estimated serum UA level variability according to BMI variance, which was defined as the difference of BMI between the first examination

and the last examination (i.e., the BMI of the 2015 examination minus the BMI of the 2017 examination). Based on the increase or decrease of BMI variance, the study population was categorized into 7 groups as follows: increase of BMI (>2.5, 1.5–2.5, and 0.5–1.5); no change (–0.5 to 0.5); and decrease of BMI (>2.5, 1.5–2.5, and 0.5–1.5).

Statistical analysis. Data were reported as means ± SD. We used median and interquartile range for continuous variables and numbers and percentages for categorical variables. According to the BMI variance categories, the variables of the study population were compared using analysis of variance, Kruskal-Wallis H test, and chi-square test.

We assessed the association between BMI variance and serum UA variability via multiple linear regression analysis in which the serum UA level (mg/dl) alteration served as the dependent variable, and the BMI variance categorized into 7 groups served as the independent variable, after adjusting for potential confounders. We also examined the impact of BMI variance on achieving the target serum UA levels through binary logistic regression analysis in which we set 2 targets of serum UA levels and separately addressed each target as the dependent variable. The 2 targets were as follows: 1) <7 mg/dl of the universally acknowledged upper normal limit in men (normouricemia); and 2) <6 mg/dl of the officially endorsed therapeutic target in patients with gout. Of the multivariable regression analyses, model 1 was adjusted for age, education level, smoking status, daily alcohol consumption, HEPA, and systolic BP (SBP); model 2 was adjusted for the total energy, total protein, total fat, fiber, calcium, and vitamin C intake based on the FFQ in addition to the variables listed in model 1; and model 3 was adjusted for the laboratory results for calcium, phosphorus, AP, LDL cholesterol, hsCRP, eGFR, and HOMA-IR in addition to the listed variables of model 2.

Through stratified analysis, we identified the effect modification of BMI variance on target serum UA level achievement in prespecified subgroups as follows: age (<43.5 years versus ≥43.5 years), alcohol consumption (<20 gm versus ≥20 gm of alcohol per day), physical activity level (<1 times/week, ≥1 times/week but insufficient HEPA, and HEPA), BMI (<25 kg/m^2 versus ≥25 kg/m^2), education level (≤ high school graduate versus ≥ college graduate), and HOMA-IR (<2.5 % versus ≥2.5 %). Each stratum was examined using the fully adjusted model. Interactions between subgroups were tested using likelihood ratio tests to compare models with and without product terms.

All analyses were first performed for subjects who undertook biennial examination (n = 27,422) and then reperformed for those who had annual examinations in the consecutive 3 years as a sensitivity analysis (n = 13,603; for baseline characteristics, see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24576). All significance tests were 2-tailed, and *P* values less than 0.05 indicated statistical significance. To analyze the data, we used Stata software, version 16.1.

RESULTS

Baseline characteristics of the study population. Table 1 gives the characteristics of the study population according to BMI variance. The mean ± SD baseline age and serum UA level were 38.8 ± 6.6 years and 6.23 ± 1.2 mg/dl, respectively. Additionally, 25.3% of the population had hyperuricemia at baseline; the more increased the BMI, the more hyperuricemic the subject was at baseline. The average BMI was 24.7 ± 2.9 kg/m^2, and younger participants had a larger BMI variance over 2 years, either being increased or decreased. In addition, the larger the BMI variance, the more frequent the baseline obesity was. Participants with a more decreased BMI were more likely to have heavier weight, higher SBP, and more elevated LDL-C and HOMA-IR at baseline than those with less decreased BMI. Diet seemingly had insignificant tendencies, excluding energy and total fat intake; the more the BMI decreased, the higher the energy intake at baseline, while the smaller the BMI variance, the lower the total fat intake at baseline.

Relationship between BMI variance and SUA variability. Over 2 years, participants with increased BMI were higher in number than those with decreased BMI (37.3% versus 19.6%). Serum UA level variability had a graded association with BMI variance in terms of the direction and size of change; the more the BMI increased, the more the SUA level increased, while the more the BMI decreased, the more the serum UA level decreased (Table 2). The association remained significantly strong, with a comparable β coefficient size in models 1, 2, and 3 after covariate adjustment. For every 1 unit of BMI loss, the serum UA level changed as much as 0.12 ($P < 0.001$). The dose-dependent relationship between the BMI variance and serum UA variability from 2015 to 2017 is depicted in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24576. We also ascertained this inclination in those who undertook 3 consecutive examinations in 2015, 2016, and 2017; for every 1 unit of BMI loss, the serum UA level changed as much as 0.13 ($P < 0.001$) (see Supplementary Table 2 and Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24576). Then, we further tested for the impact of 1-year change of BMI during 2015–2016 and the consecutive change of BMI during 2015–2017 on changes in serum UA levels; the consecutive change of BMI, either increase or decrease, could be significantly associated with changes in serum UA levels, while the 1-year change of BMI had no significant association with changes in serum UA levels except for change of BMI decrease ≥1.5 (see Supplementary Tables 3, 4, and 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24576).

Table 1. Baseline characteristics of the study population according to the categories of body mass index (BMI) change*

Variable	Overall	Increase in BMI			No change	Decrease in BMI			*P*
		≥2.5	1.5–2.5	0.5–1.5		0.5–1.5	1.5–2.5	≥2.5	
Total no. (%)	27,422 (100)	478 (1.7)	1,822 (6.6)	7,928 (28.9)	11,815 (43.1)	4,184 (15.3)	842 (3.1)	353 (1.3)	
Serum uric acid, mg/dl	6.23 ± 1.22	6.30 ± 1.32	6.32 ± 1.84	6.19 ± 1.18	6.20 ± 1.20	6.30 ± 1.25	6.40 ± 1.23	6.62 ± 1.32	<0.001
Hyperuricemia, no. (%)	6,926 (25.3)	177 (37.0)	601 (33.0)	2,076 (26.2)	2,819 (23.9)	960 (22.9)	203 (24.1)	90 (25.5)	<0.001
Age, years	38.8 ± 6.6	34.6 ± 5.5	35.9 ± 6.1	38.3 ± 6.4	39.5 ± 6.6	39.5 ± 6.6	38.8 ± 6.2	37.5 ± 5.5	<0.001
BMI, kg/m^2	24.7 ± 2.9	25.1 ± 3.4	24.7 ± 3.2	24.4 ± 2.9	24.5 ± 2.8	25.2 ± 2.8	26.5 ± 3.0	28.4 ± 3.3	<0.001
Obesity, no. (%)	11,593 (42.3)	217 (45.4)	762 (41.8)	2,918 (36.8)	4,746 (40.2)	2,071 (49.5)	570 (67.7)	309 (87.5)	<0.001
SBP, mm Hg	113.8 ± 10.4	114.7 ± 10.4	113.3 ± 10.3	113.2 ± 10.4	113.5 ± 10.3	114.8 ± 10.6	116.8 ± 30.6	119.3 ± 10.71	<0.001
HEPA, no. (%)	4,183 (15.2)	119 (24.9)	347 (19.0)	1,276 (16.1)	1,704 (14.4)	568 (13.6)	124 (14.7)	45 (12.7)	<0.001
Calcium, mg/dl	9.59 ± 0.29	9.63 ± 0.30	9.63 ± 0.29	9.61 ± 0.29	9.58 ± 0.29	9.57 ± 0.29	9.58 ± 0.30	9.56 ± 0.34	<0.001
Phosphorus, mg/dl	3.50 ± 0.42	3.63 ± 0.44	3.54 ± 0.43	3.52 ± 0.42	3.49 ± 0.41	3.46 ± 0.41	3.47 ± 0.42	3.45 ± 0.44	<0.001
AP, mg/dl	61.66 ± 14.33	62.06 ± 16.19	62.20 ± 14.91	61.81 ± 14.49	61.44 ± 14.12	61.57 ± 14.17	61.45 ± 14.39	64.06 ± 13.67	0.429
LDL cholesterol, mg/dl	133.06 ± 30.76	128.75± 30.37	128.18 ± 30.30	131.30 ± 30.52	133.78 ± 30.90	135.23 ± 30.59	138.11 ± 30.61	141.48 ± 30.18	<0.001
eGFR, ml/minutes/ 1.73m^2	98.30 ± 14.82	100.94 ± 16.02	99.42 ± 14.55	98.12 ± 14.69	98.01 ± 14.73	98.21 ± 14.99	99.92 ± 15.67	99.70 ± 15.28	0.292
hsCRP, median (IQR) mg/dl	0.05 (0.03–0.10)	0.05 (0.03–0.10)	0.05 (0.03–0.09)	0.05 (0.03–0.09)	0.05 (0.03–0.10)	0.06 (0.03–0.11)	0.06 (0.04–0.12)	0.09 (0.05–0.16)	0.167
HOMA-IR, mean ± SD %	1.66 ± 1.10	1.56 ± 1.10	1.54 ± 1.02	1.55 ± 1.00	1.63 ± 1.08	1.83 ± 1.21	2.15 ± 1.39	2.48 ± 1.56	<0.001
Energy intake, median (IQR) kcal	1,433.2 (1,099.3–1,804.3)	1,397.9 (1,035.0–1,895.6)	1,423.7 (1,077.4–1,838.3)	1,417.7 (1,080.9–1,787.7)	1,441.8 (1,111.3–1,799.9)	1,430.3 (1,116.4–1,808.7)	1,487.9 (1,130.0–1,852.0)	1,519.0 (1,132.8–1,978.1)	0.043
Calcium intake, median (IQR) mg	240.4 (151.7–359.7)	263.5 (161.8–380.3)	246.7 (151.2–372.9)	236.4 (148.4–356.0)	242.3 (154.8–359.1)	237.3 (147.0–359.8)	244.6 (160.9–361.3)	251.4 (155.0–390.1)	0.300
Vitamin C intake, median (IQR) mg	46.05 (25.2–75.8)	44.25 (23.5–75.4)	46.6 (24.8–77.3)	45.2 (24.4–75.1)	46.6 (25.9–76.1)	45.0 (25.2–75.7)	49.0 (25.4–77.5)	46.6 (24.9–73.7)	0.922
Total protein intake, median (IQR) gm	46.9 (34.3–63.3)	49.8 (34.4–67.1)	47.8 (34.6–6.2)	46.5 (33.6–62.9)	46.8 (34.5–62.5)	47.0 (34.1–63.3)	47.9 (35.8–64.5)	50.5 (34.4–69.1))	0.832
Total fat intake, median (IQR) gm	27.2 (18.0–40.4)	30.9 (20.1–45.2)	29.2 (18.7–44.1)	27.1 (17.9–40.3)	26.8 (17.9–39.5)	26.8 (17.8–40.0)	28.5 (18.5–41.5)	30.0 (18.9–46.5)	0.005
Fiber intake, median (IQR) gm	2.8 (1.9–4.1)	2.6 (1.8–4.1)	2.8 (1.8–4.1)	2.8 (1.8–4.1)	2.8 (1.9–4.1)	2.8 (1.9–4.1)	2.9 (1.9–4.2)	2.9 (1.9–4.3)	0.092
Alcohol intake, median (IQR) gm	10 (4–23)	10 (4–21)	9 (4–21)	10 (4–21)	10 (4–24)	10 (4–24)	10 (4–26)	10 (4–24)	0.004
Highest education, no. (%)†	24,331 (88.7)	432 (90.4)	1,616 (88.7)	7,008 (88.4)	10,478 (88.7)	3,723 (89.0)	750 (89.1)	324 (91.8)	0.293
Current smoker, no. (%)	6,914 (25.2)	128 (26.8)	462 (25.4)	2,002 (25.2)	2,940 (24.9)	1,093 (26.1)	202 (24.0)	87 (24.6)	0.823

* Values are the mean ± SD unless indicated otherwise. For percentages with missing values, the percentage represents the percent of subjects with a measurement with a positive value. AP = alkaline phosphatase; eGFR = estimated glomerular filtration rate; HEPA = health-enhancing physical activity; HOMA-IR = homeostasis model assessment–estimated insulin resistance; hsCRP = high-sensitivity C-reactive protein; IQR = interquartile range; LDL = low-density lipoprotein; SBP = systolic blood pressure.
† More than or equal to college graduate.

Table 2. Relationship between body mass index (BMI) variance and serum uric acid variability*

BMI variance	No.	Model 1		Model 2		Model 3	
		β (95% CI)	*P*	β (95% CI)	*P*	β (95% CI)	*P*
BMI, 1 unit decreased	27,422	0.11 (0.10, 0.12)	<0.001	0.11 (0.1, 0.12)	<0.001	0.12 (0.11, 0.13)	<0.001
Increase							
≥2.5	478	−0.35 (−0.42, −0.28)	<0.001	−0.35 (−0.42, −0.27)	<0.001	−0.36 (−0.43, −0.29)	<0.001
1.5–2.5	1,822	−0.18 (−0.22, −0.14)	<0.001	−0.18 (−0.22, −0.14)	<0.001	−0.19 (−0.22, −0.15)	<0.001
0.5–1.5	7,928	−0.09 (−0.11, −0.07)	<0.001	−0.09 (−0.11, −0.07)	<0.001	−0.10 (−0.12, −0.08)	<0.001
No change	11,815	0 (Ref.)		0 (Ref.)		0 (Ref.)	
Decrease							
0.5–1.5	4,184	0.12 (0.10, 0.15)	<0.001	0.12 (0.10, 0.15)	<0.001	0.13 (0.11, 0.16)	<0.001
1.5–2.5	842	0.23 (0.18, 0.29)	<0.001	0.23 (0.18, 0.29)	<0.001	0.25 (0.20, 0.30)	<0.001
≥2.5	353	0.42 (0.34, 0.50)	<0.001	0.42 (0.34, 0.50)	<0.001	0.44 (0.36, 0.52)	<0.001

* 95% CI = 95% confidence interval. BMI variance means the difference between the BMI of 2015 and that of 2017. Serum uric acid variability means the level of change between 2015 and 2017.

Impact of BMI variance on the achievement of target serum UA levels. The influence of BMI variance as a predictor of achieving the target serum UA levels was analyzed by binary logistic regression. First, the target serum UA level was <7 mg/dl (normouricemia) as the dependent variable (Table 3). Compared with no BMI change, the multivariate odds ratios (ORs) of achieving normouricemia for a BMI increase of 0.5–1.5, 1.5–2.5, and ≥2.5 were 0.88 (95% confidence interval [95% CI] 0.83, 0.95), 0.67 (95% CI 0.60, 0.75), and 0.60 (95% CI 0.49, 0.74), respectively, whereas those for a BMI decrease of 0.5–1.5, 1.5–2.5, and ≥2.5 were 1.17 (95% CI 1.07, 1.27), 1.28 (95% CI 1.08, 1.52), and 1.46 (95% CI 1.13, 1.88), respectively. In every 1 unit of BMI loss, the multivariate OR for achieving normouricemia was 1.17 (95% CI 1.14, 1.20). Meanwhile, an alternative target serum UA level was <6 mg/dl (Table 4). Of this target achievement, the multivariate ORs for a BMI increase of 0.5–1.5, 1.5–2.5, and ≥2.5 were 0.87 (95% CI 0.82, 0.92), 0.73 (95% CI 0.66, 0.81), and 0.52 (95% CI 0.42, 0.64), respectively, whereas those for a BMI decrease of 0.5–1.5, 1.5–2.5, and ≥2.5 were 1.12 (95% CI 1.04, 0.21), 1.39 (95% CI 1.20, 1.61), and 1.74 (95% CI 1.39, 2.17), respectively. The multivariate OR for achieving <6 mg/dl was 1.18 (95% CI 1.15, 1.21) per 1 unit of BMI loss. Moreover, sensitivity analysis revealed a comparably strong impact of BMI variance on achieving target SUA levels; the multivariate OR per 1 unit of BMI loss was 1.18 (95% CI 1.13, 1.23) for achieving normouricemia and 1.20 (95% CI 1.15, 1.24) for achieving <6 mg/dl (see Supplementary Tables 3 and 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24576).

Effect modification by age, alcohol intake, physical activity, obesity, education level, and insulin resistance. In the stratified analysis, the graded association between BMI variance and the ORs to achieve the target serum UA levels seemed to be similar across prespecified subgroups. Achieving normouricemia was more strongly associated in those with an age of ≥43.5 years, daily alcohol intake of ≥20 gm, BMI of ≥25 kg/m^2, and HOMA-IR of <2.5% (*P* for interaction <0.001, = 0.004, <0.001, and <0.001, respectively) (Table 5). Education had a modifying effect on achieving the lowered target serum UA levels of 6 mg/dl (*P* for interaction = 0.043) (Table 6). Age, BMI, and HOMA-IR were effect modifiers consistently in the sensitivity analysis (see Supplementary Tables 5 and 6, available on the *Arthritis*

Table 3. Impact of body mass index (BMI) variance as a predictor for achieving the target serum uric acid level of normouricemia*

BMI variance	No.	Model 1		Model 2		Model 3	
		OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
BMI, 1 unit decreased	20,496	1.12 (1.09, 1.15)	<0.001	1.12 (1.1, 1.16)	<0.001	1.17 (1.14, 1.2)	<0.001
Increase							
≥2.5	301	0.64 (0.52, 0.77)	<0.001	0.64 (0.52, 0.77)	<0.001	0.60 (0.49, 0.74)	<0.001
1.5–2.5	1,221	0.70 (0.63, 0.78)	<0.001	0.70 (0.63, 0.78)	<0.001	0.67 (0.60, 0.75)	<0.001
0.5–1.5	5,852	0.91 (0.85, 0.97)	0.003	0.91 (0.85, 0.97)	0.004	0.88 (0.83, 0.95)	<0.001
No change	8,996	1 (Ref.)		1 (Ref.)		1 (Ref.)	
Decrease							
0.5–1.5	3,224	1.10 (1.01, 1.20)	0.027	1.10 (1.01, 1.2)	0.024	1.17 (1.07, 1.27)	0.001
1.5–2.5	639	1.11 (0.94, 1.32)	0.203	1.12 (0.95, 1.32)	0.185	1.28 (1.08, 1.52)	0.005
≥2.5	263	1.16 (0.90, 1.48)	0.243	1.16 (0.91, 1.49)	0.226	1.46 (1.13, 1.88)	0.004

* 95% CI = 95% confidence interval; OR = odds ratio. BMI variance means the difference between the BMI of 2015 and that of 2017. Serum uric acid variability means the level of change between 2015 and 2017.

Table 4. Impact of body mass index (BMI) variance as a predictor for achieving the target serum uric acid level of <6 mg/dl*

		Model 1		Model 2		Model 3	
BMI variance	No.	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
BMI 1 unit decreased	11,580	1.14 (1.11, 1.17)	<0.001	1.14 (1.11, 1.17)	<0.001	1.18 (1.15, 1.21)	<0.001
Increase							
≥2.5	128	0.55 (0.45, 0.68)	<0.001	0.55 (0.45, 0.68)	<0.001	0.52 (0.42, 0.64)	<0.001
1.5–2.5	636	0.75 (0.68, 0.84)	<0.001	0.76 (0.68, 0.84)	<0.001	0.73 (0.66, 0.81)	<0.001
0.5–1.5	3,200	0.89 (0.84, 0.94)	<0.001	0.89 (0.84, 0.94)	<0.001	0.87 (0.82, 0.92)	<0.001
No change	5,182	1 (Ref.)		1 (Ref.)		1 (Ref.)	
Decrease							
0.5–1.5	1,875	1.08 (1.00, 1.16)	0.048	1.08 (1.00, 1.16)	0.048	1.12 (1.04, 1.21)	0.002
1.5–2.5	392	1.24 (1.07, 1.42)	0.004	1.24 (1.07, 1.43)	0.003	1.39 (1.20, 1.61)	<0.001
≥2.5	167	1.41 (1.14, 1.75)	0.002	1.42 (1.14, 1.76)	0.001	1.74 (1.39, 2.17)	<0.001

* 95% CI = 95% confidence interval; OR = odds ratio. BMI variance means the difference between the BMI of 2015 and that of 2017. Serum uric acid variability means the level of change between 2015 and 2017.

Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24576). Physical activity had no modifying effect in all analyses.

DISCUSSION

The present study evaluated the association of weight change (gain or loss) with the subsequent changes in serum UA levels (rise or fall) according to BMI variance in 27,422 clinically healthy men. The BMI change over 2 years could be positively associated with the subsequent changes in serum UA levels in terms of both direction and magnitude. BMI gain was associated with rises in serum UA levels, and BMI loss was associated with falls in serum UA levels. Small changes in BMI were associated with low variability of serum UA levels, while large changes in BMI were associated with high variability of serum UA levels. The odds of achieving the target serum UA levels showed a dose-response relationship; large gain of BMI was associated with being more difficult to attain the target, while small gain of BMI was associated with being less difficult. Large loss of BMI was associated with being easier to attain the target, while small loss of BMI was associated with being less easy. This pattern of graded association seemed to be similar across prespecified subgroups, and the modifying effects were consistently shown in older ages (age ≥43.5 years), obesity (BMI ≥25 kg/m^2), and lower insulin resistance (HOMA-IR <2.5 %) in the main and sensitivity analysis.

The mean serum UA level was 6.23 ± 1.22 mg/dl, and the proportion of subjects with hyperuricemia was one-quarter (25.3%) at baseline. These values were slightly higher than those of our previous report on male health examinees from 2011 to 2014 (mean ± SD serum UA level 5.92 ± 1.20 mg/dl; frequency of hyperuricemia [17.9%]) (21). Over 2 years, more subjects experienced a gain in BMI rather than a loss (37.3% versus 19.6%). These results are consistent with the national trends of overweight in Korean adults (26) as well as the worldwide data indicating the rising prevalence of overweight and obesity (27). Furthermore, the association between hyperuricemia and obesity has been constantly suggested (28) despite the relatively limited data for losing weight and decreasing serum UA. The strength of the present study is greater in this regard.

Our observation corroborated prior studies that demonstrated the impact of weight loss on serum UA level. In a prospective cohort of 12,379 men investigated for over 7 years, a graded relationship has been observed between weight loss and normouricemia achievement in which the effect of weight loss ≥10 kg increased the odds by nearly 4-fold in a given individual (17). In comparison, the odds of our study for achieving the target serum UA levels of normouricemia were relatively small; when BMI loss was >2.5 units, the OR increased by ~0.5 times only (OR 1.46 [95% CI 1.13–1.88]). This discrepancy would stem from the characteristics of the study population, the period of observation, and the unit of measure for weight change. These 12,379 men had a high cardiovascular risk profile, their mean serum UA level was 407 μmoles/liter at baseline, and they were followed up for 7 years. In our study, we recruited apparently healthy men from a health check-up program, excluding those with high-risk characteristics (e.g., heart disease, hypertension, diabetes mellitus, and stroke) and assessed their BMI and serum UA level changes in a relatively short period (2 years). Furthermore, BMI is a formula that considers both the height and weight and is thereby more appropriate to gauge health-related risks. Thus, despite the small odds, the present study obtained more practical results regarding the relationship between weight and serum UA level in the general population.

A southwestern Chinese study recently investigated a general population focusing on sex- and age-specific impact for the odds of developing hyperuricemia along with weight gain (19). Despite the fact that the study design was different from ours, this cross-sectional study suggested that active weight gain would increase the likelihood of developing hyperuricemia. This impact was modified by age; younger subjects had a higher OR than

Table 5. Effect modification of body mass index (BMI) change on achieving the target serum uric acid levels of normouricemia by clinically relevant subgroups*

Subgroups	No.	Increase in BMI			No change	Decrease in BMI			*P* for interaction
		≥2.5	1.5–2.5	0.5–1.5		0.5–1.5	1.5–2.5	≥2.5	
Age, years									<0.001
<43.5	13,720	0.61 (0.48, 0.77)	0.65 (0.57, 0.75)	0.91 (0.83, 1.00)	1 (Ref.)	1.22 (1.08, 1.38)	1.23 (0.97, 1.55)	1.32 (0.95, 1.83)	
≥43.5	13,702	0.47 (0.31, 0.69)	0.62 (0.50, 0.75)	0.82 (0.74, 0.91)	1 (Ref.)	1.11 (0.98, 1.25)	1.35 (1.04, 1.75)	1.66 (1.08, 2.54)	
Alcohol, gm/day									0.004
<20	19,321	0.60 (0.47, 0.76)	0.66 (0.57, 0.75)	0.86 (0.79, 0.94)	1 (Ref.)	1.17 (1.06, 1.31)	1.28 (1.03, 1.58)	1.28 (0.94, 1.75)	
≥20	8,101	0.60 (0.42, 0.86)	0.70 (0.57, 0.86)	0.94 (0.83, 1.06)	1 (Ref.)	1.15 (0.98, 1.34)	1.30 (0.96, 1.75)	1.88 (1.20, 2.95)	
Physical activity level									0.835
<1 times/week	14,551	0.61 (0.45, 0.83)	0.65 (0.55, 0.76)	0.89 (0.8, 0.97)	1 (Ref.)	1.15 (1.02, 1.29)	1.38 (1.08, 1.75)	1.58 (1.12, 2.22)	
≥1 times/week but insufficient to HEPA	8,670	0.53 (0.37, 0.75)	0.74 (0.60, 0.90)	0.86 (0.76, 0.97)	1 (Ref.)	1.13 (0.97, 1.32)	1.14 (0.84, 1.53)	1.09 (0.69, 1.72)	
HEPA	4,183	0.71 (0.47, 1.07)	0.61 (0.47, 0.79)	0.92 (0.77, 1.1)	1 (Ref.)	1.33 (1.04, 1.69)	1.31 (0.84, 2.06)	2.08 (0.97, 4.48)	
BMI, kg/m^2									<0.001
<25	15,829	0.65 (0.49, 0.87)	0.70 (0.60, 0.83)	0.87 (0.79, 0.96)	1 (Ref.)	1.05 (0.92, 1.2)	1.64 (1.14, 2.37)	1.31 (0.57, 3.01)	
≥25	11,593	0.59 (0.44, 0.79)	0.65 (0.56, 0.77)	0.88 (0.79, 0.97)	1 (Ref.)	1.33 (1.18, 1.49)	1.36 (1.12, 1.66)	1.72 (1.31, 2.24)	
Education									0.113
≤ high school	2,934	0.81 (0.39, 1.67)	0.74 (0.52, 1.06)	0.91 (0.74, 1.13)	1 (Ref.)	1.09 (0.83, 1.44)	1.67 (0.92, 3.01)	1.85 (0.73, 4.72)	
≥ college graduate	24,331	0.59 (0.48, 0.73)	0.67 (0.59, 0.75)	0.88 (0.82, 0.95)	1 (Ref.)	1.17 (1.07, 1.28)	1.27 (1.06, 1.52)	1.43 (1.09, 1.86)	
HOMA-IR, %									<0.001
<2.5	23,086	0.60 (0.48, 0.74)	0.68 (0.6, 0.77)	0.90 (0.84, 0.97)	1 (Ref.)	1.14 (1.03, 1.26)	1.15 (0.93, 1.42)	1.34 (0.96, 1.88)	
≥2.5	4,336	0.71 (0.42, 1.18)	0.71 (0.54, 0.94)	0.84 (0.71, 0.99)	1 (Ref.)	1.14 (0.96, 1.36)	1.37 (1.02, 1.85)	1.33 (0.91, 1.95)	

* The final multivariable model was applied. HEPA = health-enhancing physical activity; HOMA-IR = homeostasis model assessment–estimated insulin resistance.

older subjects. Such a finding was explained by the distribution of overweight and obesity, which were more prevalent in younger ages. The present study identified age, BMI, alcohol, and HOMA-IR as the effect modifiers for the association between BMI change and achieving the target serum UA levels. There was a difference in the insulin sensitivity between the 2 subgroups of age, and this may bring out the modifying effect. It may be supposed that the subgroups of daily alcohol ≥20 gm and of BMI ≥25 had quickly reacted to BMI change in particular. In the sensitivity analysis, however, the ORs for modifying effect were slightly different but still statistically significant. Nevertheless, the tendency of dynamic effect modification was nearly the same as that of the main stratified analysis. Further investigations are necessary for the precise mechanism of effect modification in the results of this study.

The relationship between BMI and serum UA could be multifactorial by several feasible mechanisms. The first possible mechanism is the change of urinary uric acid clearance along with weight change. Urinary uric acid excretion is seemingly inversely proportional to insulin resistance (29). The trajectories of BMI change strongly affect glucose and insulin metabolism (30). Accordingly, BMI variance may lead to insulin sensitivity change, which possibly modifies the urinary uric acid clearance, thereby altering serum UA level. Furthermore, the response in the fractional excretion of uric acid via the kidneys is different between healthy subjects and patients with gout (31). Thus, the impact of BMI change on serum UA level could differ among individuals with or without gout. The second mechanism to explain the relationship between BMI and serum UA could be body fat distribution. Visceral fat accumulation was elucidated for uric acid overproduction in men with obesity; in addition, visceral fat is more closely linked to uric acid overproduction than subcutaneous fat (32). Moreover, reduced fat mass along with BMI loss induces truncal fat reduction, demonstrating a positive association with serum UA improvement in men (33). This finding deserves special emphasis, considering that excessive android fat deposition is associated with cardiovascular and metabolic risks, such as an increasing serum UA level (34). Xanthine oxidoreductase (XOR)

Table 6. Effect modification of body mass index (BMI) change on achieving the target serum uric acid levels of 6 mg/dl by clinically relevant subgroups*

Subgroups	No.	Increase in BMI			No change	Decrease in BMI			*P* for interaction
		≥2.5	1.5–2.5	0.5–1.5		0.5–1.5	1.5–2.5	≥2.5	
Age, years									<0.001
<43.5	13,720	0.57 (0.45, 0.73)	0.72 (0.63, 0.82)	0.85 (0.78, 0.92)	1 (Ref.)	1.13 (1.01, 1.26)	1.39 (1.13, 1.71)	1.72 (1.28, 2.32)	
≥43.5	13,702	0.31 (0.20, 0.48)	0.65 (0.54, 0.78)	0.86 (0.79, 0.93)	1 (Ref.)	1.12 (1.02, 1.24)	1.40 (1.14, 1.72)	1.70 (1.22, 2.38)	
Alcohol, gm/day									0.018
<20	19,321	0.53 (0.41, 0.68)	0.70 (0.61, 0.79)	0.86 (0.80, 0.93)	1 (Ref.)	1.12 (1.03, 1.23)	1.39 (1.17, 1.66)	1.68 (1.28, 2.21)	
≥20	8,101	0.48 (0.32, 0.72)	0.83 (0.68, 1.01)	0.88 (0.79, 0.99)	1 (Ref.)	1.12 (0.98, 1.28)	1.38 (1.06, 1.79)	1.83 (1.25, 2.68)	
Physical activity level									0.298
<1 times/week	14,551	0.56 (0.41, 0.76)	0.76 (0.65, 0.88)	0.86 (0.79, 0.94)	1 (Ref.)	1.13 (1.03, 1.25)	1.44 (1.18, 1.76)	1.97 (1.47, 2.63)	
≥1 times/week but insufficient to HEPA	8,670	0.47 (0.31, 0.69)	0.77 (0.63, 0.94)	0.9 (0.81, 1.00)	1 (Ref.)	1.15 (1.01, 1.32)	1.62 (1.25, 2.11)	1.39 (0.91, 2.11)	
HEPA	4,183	0.5 (0.32, 0.76)	0.58 (0.45, 0.75)	0.82 (0.71, 0.96)	1 (Ref.)	1.00 (0.82, 1.22)	0.86 (0.58, 1.26)	1.62 (0.87, 3.05)	
BMI, kg/m^2									<0.001
<25	15,829	0.50 (0.38, 0.66)	0.77 (0.67, 0.88)	0.85 (0.79, 0.92)	1 (Ref.)	1.16 (1.05, 1.28)	1.44 (1.12, 1.85)	2.23 (1.17, 4.24)	
≥25	11,593	0.58 (0.41, 0.81)	0.68 (0.57, 0.82)	0.87 (0.79, 0.97)	1 (Ref.)	1.15 (1.03, 1.28)	1.56 (1.30, 1.88)	1.97 (1.55, 2.50)	
Education									0.043
≤ high school	2,934	0.40 (0.20, 0.82)	0.71 (0.52, 0.97)	0.89 (0.74, 1.06)	1 (Ref.)	1.05 (0.84, 1.32)	1.20 (0.76, 1.90)	2.22 (0.99, 4.97)	
≥ college graduate	24,331	0.53 (0.42, 0.66)	0.73 (0.65, 0.82)	0.86 (0.81, 0.92)	1 (Ref.)	1.13 (1.04, 1.22)	1.42 (1.22, 1.66)	1.69 (1.34, 2.13)	
HOMA-IR, %									<0.001
<2.5	23,086	0.52 (0.42, 0.65)	0.76 (0.68, 0.85)	0.89 (0.84, 0.95)	1 (Ref.)	1.14 (1.05, 1.24)	1.28 (1.08, 1.51)	1.64 (1.25, 2.16)	
≥2.5	4,336	0.60 (0.31, 1.17)	0.63 (0.45, 0.88)	0.77 (0.64, 0.92)	1 (Ref.)	0.91 (0.76, 1.09)	1.44 (1.08, 1.93)	1.56 (1.07, 2.26)	

* The final multivariable model was applied. HEPA = health-enhancing physical activity; HOMA-IR = homeostasis model assessment–estimated insulin resistance.

activity may participate in the process of BMI change and serum UA alteration. In the purine metabolism pathway, XOR works in the 2 final steps, catalyzing the oxidation of hypoxanthine to xanthine, and xanthine to uric acid (35). The plasma activity of this enzyme is elevated in children with obesity (36) and positively correlates with BMI and uric acid (37,38). Furthermore, human adipose tissue is reportedly the source of hypoxanthine secretion (39). Taken together, the alteration of serum UA level could be induced by body fat change according to BMI variance. Hence, we would like to concentrate on the association of body fat with serum UA in men in the follow-up study.

In terms of serum UA variability, the impact of BMI loss in lowering serum UA level seemed considerably smaller than the effect of uric acid–lowering therapy (ULT). However, ULT is not indicated for asymptomatic subjects without gout attack according to the current guidelines despite its correlation with various complicated diseases. Thus, evident measures other than ULT should be recommended for those subjects. The present study may provide the basis for health providers to comment on the relationship between BMI and serum UA on that account, which is another strength of this study. In addition, the consecutive 2-year BMI change could be more influential than the 1-year BMI change to the subsequent serum UA change.

However, this observational study has some limitations. The first is its cross-sectional nature, which imposes limits on the temporal relationship. Second, weight change was assessed using BMI variance, which is difficult to use to differentiate the change of fat mass from that of lean mass. Third, BMI loss over 2 years was not interventional; thus, it could be a consequence of diverse means, such as diet and physical activity, which might include unintentional loss. Therefore, the measures for losing weight and all the effects of BMI loss on serum UA level are unfeasible. Fourth, we collected data using a self-administered questionnaire, which might have involved a recall bias, especially in lifestyle factors. Fifth, the study results would not be directly applicable to patients with gout because the study population was not formed for the purpose of

investigating the impact of BMI change in gout. Finally, this study included Korean men with a mean age of 38.8 years who underwent regular health check-ups; thus, our findings might not reflect the situation of women, other age groups, or other ethnic groups.

In conclusion, the present study suggests a possible basis for BMI change in serum UA levels among apparently healthy men. Although the evidence had a small effect size, the health risks and benefits of BMI change would be emphasized for serum UA alteration. Further research is necessary, with more focus on elucidating the underlying mechanisms and specifying the preferential intervention of BMI change, thus identifying its optimal degree and quantity regarding serum UA level modification.

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. Ahn and Cha 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. Hwang, Ahn, Cha.
Acquisition of data. Hwang, Lee, Ahn, Cha.
Analysis and interpretation of data. Hwang, Lee, Ahn, Cha.

REFERENCES

1. Richette P, Bardin T. Gout. Lancet 2010;375:318–28.
2. Alvarez-Lario B, Macarron-Vicente J. Uric acid and evolution. Rheumatology (Oxford) 2010;49:2010–5.
3. Kanbay M, Jensen T, Solak Y, Le M, Roncal-Jimenez C, Rivard C, et al. Uric acid in metabolic syndrome: from an innocent bystander to a central player. Eur J Intern Med 2016;29:3–8.
4. Li M, Hou W, Zhang X, Hu L, Tang Z. Hyperuricemia and risk of stroke: a systematic review and meta-analysis of prospective studies. Atherosclerosis 2014;232:265–70.
5. Kuo CF, Grainge MJ, Zhang W, Doherty M. Global epidemiology of gout: prevalence, incidence and risk factors. Nat Rev Rheumtol 2015;11:649–62.
6. Kim JW, Kwak SG, Lee H, Kim SK, Choe JY, Park SH. Prevalence and incidence of gout in Korea: data from the national health claims database 2007–2015. Rheumatol Int 2017;37:1499–506.
7. Kuo CF, Grainge MJ, Mallen C, Zhang W, Doherty M. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. Ann Rheum Dis 2015;74:661–7.
8. Lee CH, Sung NY. The prevalence and features of Korean gout patients using the National Health Insurance Corporation Database. J Rheum Dis 2011;18:94–100.
9. Zhu Y, Pandya BJ, Choi HK. Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutrition Examination Survey 2007–2008. Arthritis Rheum 2011;63:3136–41.
10. Campion EW, Glynn RJ, DeLabry LO. Asymptomatic hyperuricemia: risks and consequences in the Normative Aging Study. Am J Med 1987;82:421–6.
11. Friedrich MJ. Global obesity epidemic worsening. JAMA 2017; 318:603.
12. Richette P, Doherty M, Pascual E, Barskova V, Becce F, Castaneda-Sanabria J, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis 2017;76:29–42.
13. FitzGerald JD, Dalbeth N, Mikuls T, Brignardello-Petersen R, Guyatt G, Abeles AM, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Care Res (Hoboken) 2020;72:744–60.
14. Dalbeth N, Chen P, White M, Gamble GD, Barratt-Boyes C, Gow PJ, et al. Impact of bariatric surgery on serum urate targets in people with morbid obesity and diabetes: a prospective longitudinal study. Ann Rheum Dis 2014;73:797–802.
15. Williams PT. Effects of diet, physical activity and performance, and body weight on incident gout in ostensibly healthy, vigorously active men. Am J Clin Nutr 2008;87:1480–7.
16. Richette P, Poitou C, Manivet P, Denis J, Bouillot JL, Clément K, et al. Weight loss, xanthine oxidase, and serum urate levels: a prospective longitudinal study of obese patients. Arthritis Care Res (Hoboken) 2016;68:1036–42.
17. Zhu Y, Zhang Y, Choi HK. The serum urate-lowering impact of weight loss among men with a high cardiovascular risk profile: the Multiple Risk Factor Intervention Trial. Rheumatology (Oxford) 2010;49: 2391–9.
18. Nielsen SM, Bartels EM, Henriksen M, Waehrens EE, Gudbergsen H, Bliddal H, et al. Weight loss for overweight and obese individuals with gout: a systematic review of longitudinal studies. Ann Rheum Diseases 2017;76:1870–82.
19. Liu DM, Jiang LD, Gan L, Su Y, Li F. Association between serum uric acid level and body mass index in sex- and age-specific groups in Southwestern China. Endocr Pract 2019;25:438–45.
20. Song JU, Hwang J, Ahn JK. Serum uric acid is positively associated with pulmonary function in Korean health screening examinees. Mod Rheumatol 2017;27:1057–65.
21. Hwang J, Hwang JH, Ryu S, Ahn JK. Higher serum uric acid is associated with higher lumbar spine bone mineral density in male health-screening examinees: a cross-sectional study. J Bone Miner Metab 2019;37:142–51.
22. Ahn Y, Kwon E, Shim JE, Park MK, Joo Y, Kimm K, et al. Validation and reproducibility of food frequency questionnaire for Korean genome epidemiologic study. Eur J Clin Nutr 2007;61:1435–41.
23. Kim MK, Lee WY, Kang JH, Kang JH, Kim BT, Kim SM, et al. 2014 clinical practice guidelines for overweight and obesity in Korea. Endocrinol Metab (Seoul) 2014;29:405–9.
24. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003;35: 1381–95.
25. Levey AS, Coresh J, Greene T, Marsh J, Stevens LA, Kusek JW, et al. Expressing the Modification of Diet in Renal Disease Study equation for estimating glomerular filtration rate with standardized serum creatinine values. Clin Chem 2007;53:766–72.
26. Shin HY, Kang HT. Recent trends in the prevalence of underweight, overweight, and obesity in Korean adults: the Korean National Health and Nutrition Examination Survey from 1998 to 2014. J Epidemiol 2017;27:413–9.
27. Fryar CD, Kruszon-Moran D, Gu Q, Ogden CL. Mean body weight, height, waist circumference, and body mass index among adults: United States, 1999–2000 through 2015–2016. Natl Health Stat Report 2018:1–16.
28. Gong M, Wen S, Nguyen T, Wang C, Jin J, Zhou L. Converging relationships of obesity and hyperuricemia with special reference to metabolic disorders and plausible therapeutic implications. Diabetes Metab Syndr Obes 2020;13:943–62.
29. Facchini F, Chen YD, Hollenbeck CB, Reaven GM. Relationship between resistance to insulin-mediated glucose uptake, urinary uric acid clearance, and plasma uric acid concentration. JAMA 1991; 266:3008–11.

30. Walsh EI, Shaw J, Cherbuin N. Trajectories of BMI change impact glucose and insulin metabolism. Nutr Metab Cardiovasc Dis 2018;28: 243–51.
31. Liu S, Perez-Ruiz F, Miner JN. Patients with gout differ from healthy subjects in renal response to changes in serum uric acid. Joint Bone Spine 2017;84:183–8.
32. Matsuura F, Yamashita S, Nakamura T, Nishida M, Nozaki S, Funahashi T, et al. Effect of visceral fat accumulation on uric acid metabolism in male obese subjects: visceral fat obesity is linked more closely to overproduction of uric acid than subcutaneous fat obesity. Metabolism 1998;47:929–33.
33. Zhang X, Zhu C, Gao J, Mei F, Yin J, Bu L, et al. Gender difference in the relationship between serum uric acid reduction and improvement in body fat distribution after laparoscopic sleeve gastrectomy in Chinese obese patients: a 6-month follow-up. Lipids Health Dis 2018;17:288.
34. Sari CI, Eikelis N, Head GA, Schlaich M, Meikle P, Lambert G, et al. Android fat deposition and its association with cardiovascular risk factors in overweight young males. Front Physiol 2019;10:1162.
35. Glantzounis GK, Tsimoyiannis EC, Kappas AM, Galaris DA. Uric acid and oxidative stress. Curr Pharm Des 2005;11:4145–51.
36. Chiney MS, Schwarzenberg SJ, Johnson LA. Altered xanthine oxidase and N-acetyltransferase activity in obese children. Br J Clin Pharmacol 2011;72:109–15.
37. Goicoechea M, Garcia de Vinuesa S, Verdalles U, Verde E, Macias N, Santos A, et al. Allopurinol and progression of CKD and cardiovascular events: long-term follow-up of a randomized clinical trial. Am J Kidney Dis 2015;65:543–9.
38. Furuhashi M, Matsumoto M, Tanaka M, Moniwa N, Murase T, Nakamura T, et al. Plasma xanthine oxidoreductase activity as a novel biomarker of metabolic disorders in a general population. Circ J 2018; 82:1892–9.
39. Nagao H, Nishizawa H, Tanaka Y, Fukata T, Mizushima T, Furuno M, et al. Hypoxanthine secretion from human adipose tissue and its increase in hypoxia. Obesity (Silver Spring) 2018;26: 1168–78.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1287–1293
DOI 10.1002/acr.24565

American College of Rheumatology
Empowering Rheumatology Professionals

Relationship Between Body Mass Index, Disease Activity, and Exercise in Ankylosing Spondylitis

Jean W. Liew,[1] Milena A. Gianfrancesco,[2] Susan R. Heckbert,[3] and Lianne S. Gensler[2]

Objective. Ankylosing spondylitis (AS) is associated with elevated cardiovascular risk, and obesity is a common, modifiable risk factor. Our aims were to assess the relationship of body mass index (BMI) with disease activity in AS patients and to assess the extent to which the effect is mediated through exercise.

Methods. We used data from a prospective AS cohort with a median follow-up of 7 years. To determine the association of BMI (kg/m^2) with disease activity as measured by the Ankylosing Spondylitis Disease Activity Score (ASDAS), we used generalized estimating equations with inverse probability weighting to account for repeated measures per subject and time-varying confounding. To estimate the direct effect of overweight/obese BMI on disease activity and the indirect effect through exercise, we performed a mediation analysis.

Results. There were 183 subjects with available BMI and disease activity data (77% male, 70% White, mean ± SD age 40.8 ± 13.3 years). Higher BMI was significantly associated with higher disease activity over time; on average, for a 1 kg/m^2 higher BMI, the ASDAS was 0.06 units higher (95% confidence interval 0.04–0.08) after adjustment for important confounders. The direct effect of an overweight/obese BMI accounted for most of the total effect on disease activity, with a smaller indirect effect mediated by exercise (7%).

Conclusion. Higher BMI was associated with higher disease activity in a prospective AS cohort. We found that being overweight/obese largely influenced disease activity directly rather than indirectly through exercise. Other mechanisms, such as increased inflammation, may better explain the obesity–disease activity association.

INTRODUCTION

Axial spondyloarthritis (SpA) is a chronic inflammatory arthritis that affects the spine and sacroiliac joints and can be separated into ankylosing spondylitis (AS), which is also known as radiographic axial SpA, and nonradiographic axial SpA. Epidemiologic studies of axial SpA have shown that elevated acute-phase reactants, smoking, and baseline radiographic damage are risk factors for spinal progression (1–3). In addition, multiple population-based studies have demonstrated increased cardiovascular (CV) events and CV-related mortality in axial SpA (4–8).

Traditional modifiable CV risk factors in the general population include diabetes mellitus, hypertension, dyslipidemia, tobacco smoking, and obesity (9). Although obesity is an important and modifiable CV risk factor, there have been limited studies addressing the impact of higher body mass index (BMI), or of being overweight or obese, on clinical outcomes in axial SpA. In a US-based registry of patients with AS and psoriatic arthritis (PsA), obesity was a significant predictor of tumor necrosis factor inhibitor (TNFi) switching or discontinuation (10). Two systematic reviews and meta-analyses assessing the effect of BMI on TNFi response in multiple inflammatory diseases found that higher BMI was associated with increased odds of an inadequate response to TNFi treatment in individuals with axial SpA (11,12). Higher BMI was associated with higher disease activity scores in observational studies of axial SpA in another systematic review and meta-analysis; however, this analysis was limited to cross-sectional data (13).

Sedentary behavior is another modifiable CV risk factor in the general population (14,15). Cross-sectional studies of AS have

Dr. Liew's work was supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases training grant T32-AR-007108), the Assessment of SpondyloArthritis international Society, and the Spondyloarthritis Research and Treatment Network. Dr. Gianfrancesco's work was supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant K01-AR-070585). Dr. Gensler's work was supported by the Russell/Engleman Rheumatology Research Center.

[1]Jean W. Liew, MD, MS: Boston University School of Medicine, Boston, Massachusetts; [2]Milena A. Gianfrancesco, PhD, MPH, Lianne S. Gensler, MD: University of California, San Francisco; [3]Susan R. Heckbert, MD, PhD: University of Washington, Seattle.

Dr. Liew has received research support from Pfizer (less than $10,000). Dr. Gensler has received consulting fees from AbbVie, Eli Lilly and Company, GSK, Gilead, Pfizer (less than $10,000 each) and research support from UCB and Novartis. No other disclosures relevant to this article were reported.

Address correspondence to Jean W. Liew, MD, MS, Section of Rheumatology, 650 Albany Street, X200, Boston, MA 02118. Email: jwliew@bu.edu.

Submitted for publication September 4, 2020; accepted in revised form January 21, 2021.

SIGNIFICANCE & INNOVATIONS

- A higher body mass index (BMI) is associated with a poorer response to tumor necrosis factor inhibitor therapy in axial spondyloarthritis (SpA), and a higher body mass index (BMI) is cross-sectionally associated with higher disease activity overall in prior observational studies.
- In this study, we found that among individuals with ankylosing spondylitis (AS), there is an association of higher BMI with higher disease activity over time, and being overweight/obese largely influences disease activity directly, rather than indirectly through exercise.
- Future studies should focus on intervenable aspects of the association between BMI and disease activity, as targeted interventions among overweight/obese individuals with AS may improve both disease activity and cardiovascular risk.

demonstrated that higher disease activity was associated with lower physical activity (16,17); however, in such studies, reverse causation is possible. In a randomized controlled trial (RCT) conducted by Sveaas et al, high intensity exercise for 3 months resulted in improved disease activity in participants with AS (18), providing further support for the recommendations for exercise in axial SpA treatment guidelines (19).

The longitudinal association of BMI and disease activity in axial SpA has not been investigated outside of cohort studies of TNFi treatment response. The causal pathway that links BMI and disease activity also requires further elucidation, including the role of exercise within this pathway. The aim of this study was to assess whether higher BMI is associated with higher disease activity in AS patients, and to what extent this association is mediated through exercise. Results from this study will support a more causal interpretation for whether interventions to reduce BMI can potentially improve disease activity in this patient population.

MATERIALS AND METHODS

Study population. The Prospective Study of Outcomes in Ankylosing Spondylitis (PSOAS) cohort has been described in detail elsewhere (20,21). Briefly, subjects were enrolled if they were at least 18 years old and met the 1984 modified New York criteria for AS (22). Individuals were recruited from the investigators' clinics, patient support groups, and community rheumatologists. Enrollment for the PSOAS cohort began in 2002 and continued through 2018. For this study, the subset of patients enrolled and followed at the University of California, San Francisco were utilized because complete longitudinal BMI data were available for this subset only.

Clinical evaluation was performed by a study site investigator using standardized protocols at study entry and at subsequent study visits every 6 months. At baseline, patient demographic data and characteristics of AS disease status, date of symptom onset, patient-reported outcomes, extra-musculoskeletal manifestations, comorbidities, and medication history were recorded. At follow-up study visits every 6 months, patient-reported outcomes and information on medications were collected. All medications used in the preceding 6 months were recorded, per patient report and investigator confirmation. The number of missed doses in the past week, month, and 6 months was also documented, along with whether the patient was still taking the medication. C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) levels were also determined at each study visit. Comorbid conditions, including CV disease and risk factors, were ascertained at baseline and every 2 years.

Variables. *Exposure of interest.* The exposure of interest was BMI (kg/m^2), which is a continuous variable in the data set. For the mediation analysis, BMI was dichotomized as overweight/obese versus normal/underweight using the WHO classification for overweight as ≥23 kg/m^2 for Asian patients and ≥25 kg/m^2 for all other racial categories.

Outcome of interest. The primary outcome of interest was disease activity as measured by the validated Ankylosing Spondylitis Disease Activity Score (ASDAS), which was a continuous variable (possible range 0.6–7.0) (23). We also explored Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) score (range 0–10) as a secondary outcome in sensitivity analyses. The BASDAI comprises 6 questions addressing 5 major symptoms in AS: fatigue, spinal pain, peripheral joint pain and swelling, localized tenderness, and morning stiffness (24). The ASDAS includes some questions from the BASDAI, as well as patient and physician global assessments, and laboratory measures (either the CRP level or ESR) (25).

Mediator of interest. Exercise was a continuous variable, measured in minutes per week and based on patient self-report (26). This was derived from 2 questions asked of patients at each study visit ("In a typical week, how many times do you exercise?" and "How long does each exercise session typically last?").

Other covariates. Age at each study visit was derived from self-reported age at the baseline visit. Race (White, Black, Asian, other) and smoking status (current smoker, yes/no) were self-reported. Nonsteroidal antiinflammatory drug (NSAID) use and TNFi use were binary variables for use versus non-use as reported for the 6 months prior to each study visit. Total follow-up time in the cohort was derived using the dates for the first and last recorded study visits.

Longitudinal analysis. To determine the association of BMI with disease activity, we used generalized estimating equations (GEEs) to account for repeated measures per subject. Potential confounders were determined a priori based on prior literature review, content knowledge, and the use of directed

acyclic graphs: age, sex, race, follow-up time in the study, current smoking, NSAID use, and TNFi use. To account for time-varying confounding, we used stabilized inverse probability weights (IPW). For the denominator of the IPW, we predicted the probability of having a certain BMI using multivariable linear regression conditioning on age, sex, and race as baseline confounders, and NSAID use and TNFi use as time-varying confounders. For the numerator of the IPW, we predicted the probability of having a certain BMI using multivariable linear regression conditioning only on the baseline confounders (age, sex, and race). We then used linear regression with GEE with an exchangeable correlation structure to fit a response model that included only the outcome and exposure of interest, weighted by the stabilized IPW. We used complete cases for this primary analysis.

In addition, we performed a series of sensitivity analyses to assess the robustness of the results. First, we excluded individuals from the cohort who did not have any measures of the exposure (BMI, 39% missing from all observations) or outcome (ASDAS, 25% missing) of interest in any of their study visits, leaving 183 subjects. For the remainder of the prespecified variables, multiple imputation with chained equations (MICE) was performed with 10 iterations (27–29). The analyses were as follows: 1) single imputation with the last observation carried forward on the subset with available BMI and disease activity measures (n = 183), using IPW-GEE; 2) MICE on the subset with available BMI and disease activity measures (n = 183), using conventional GEE; 3) MICE on the full cohort (n = 283), using conventional GEE. All analyses were performed using ASDAS and BASDAI score as the primary and secondary outcomes of interest, respectively.

Mediation analysis. We performed a mediation analysis to estimate the direct effect of overweight/obese status on disease activity, and the indirect effect through exercise (causal diagram shown in Figure 1). The goal of mediation analysis is to understand if, and to what extent, the effect of the exposure on the outcome is mediated through the intermediate variable. Mediation analysis can be useful for disentangling which modifiable factors could serve as future interventions and to further elucidate underlying causal pathways.

Older mediation methods such as the Baron and Kenny approach, while less complex, are limited due the underlying assumption of no interaction between the exposure and the mediator (30). That is, one would assume that a higher BMI would not impact the level of exercise that is undertaken, an assumption that is likely violated. Causal mediation analysis is based on the potential outcomes framework, which compares the observed data to what might have been observed (i.e., the counterfactual), with all else being equal (31). This approach to mediation analysis is additionally based upon assumptions of no unmeasured confounding between the exposure and mediator, and no unmeasured confounding between the mediator and outcome. In contrast to the approach of Baron and Kenny, causal mediation analysis allows for potential interaction between the exposure and the mediator (e.g., an interaction between BMI and exercise frequency). Under the assumption of sequential ignorability, the exposure is considered statistically independent of potential outcomes and potential mediators after adjusting for confounders; and the observed mediator is considered to be random after accounting for the exposure and baseline confounders.

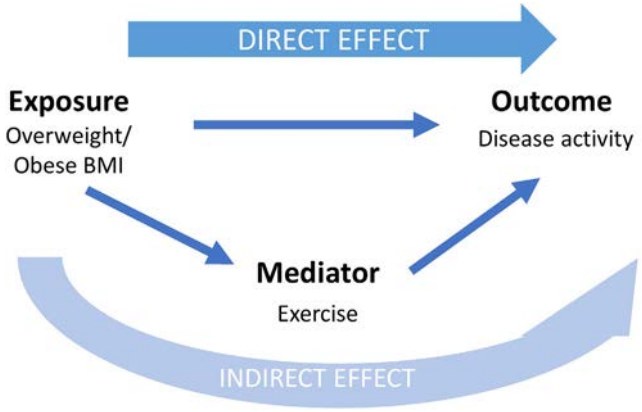


Figure 1. Simplified directed acyclic graph demonstrating the relationship between the exposure, mediator, and outcome of interest. Using mediation analysis, the total effect can be decomposed into direct and indirect effects. The direct effect is the unmediated effect of the exposure on the outcome. The indirect effect is the effect of the exposure on the outcome through the mediator. BMI = body mass index.

With causal mediation analysis, the total effect can be decomposed into natural direct and indirect effects (Figure 1). The natural direct effect is the effect of the exposure on the outcome independent of the mediator. The natural direct effect compares potential outcomes for disease activity between those with overweight/obese BMI versus those with normal BMI, while the level of the mediator (exercise) is held constant. The natural indirect effect is the effect of the exposure on the outcome through the mediator. The natural indirect effect compares potential outcomes for disease activity if exercise (mediator) is at the level it would have been in the setting of an overweight/obese BMI versus exercise at the level it would have been in the setting of a normal BMI. We also determined the proportion of the effect of an overweight/obese BMI that is mediated through exercise (proportion mediated).

We used a 2-stage approach in which the outcome was regressed on the exposure, mediator, exposure–mediator interaction term, and confounders; and the mediator was regressed on the exposure and the confounders (31,32) (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24565). In this approach, the natural direct and indirect effects were estimated using Monte Carlo simulations. For the association of BMI and disease activity, we did not identify a clinically relevant confounder of the association between the mediator (exercise) and the outcome

(disease activity), which is also affected by the exposure (BMI); therefore, marginal structural models were not determined to be necessary to produce less biased estimates. The mediation analysis was performed on the imputed data using MICE on 183 subjects (as per the main analysis) and on the complete cases.

We conducted a sensitivity analysis to determine the robustness of the mediation analysis against violations of the sequential ignorability assumption (33). That is, if the sequential ignorability assumption is true, then the correlation between the 2 error terms for the regression models used for mediation analysis would equal zero ($\rho = 0$). A larger value of ρ would indicate that a very strong correlation between error terms would be needed for the indirect effect to be fully explained by unmeasured confounders. Analyses were performed in Stata, version 15, and R version 3.6.0 (34) using the MICE, IPW, and mediation (32,35,36) packages.

RESULTS

Baseline characteristics of 183 subjects with available exposure (BMI) and outcome (disease activity) information are shown in Table 1 and stratified by BMI. Among these 183 subjects, 77% were male, 70% were White, and the mean ± SD age was 40.8 ± 13.3 years. The mean first-available BMI was 25.7 ± 4.8 kg/m^2. Those with higher BMI were older, had longer symptom duration and a higher proportion of abnormally elevated CRP, and fewer minutes of exercise per week compared to those with lower BMI. Characteristics of patients with missing exposure and/or outcome values (n = 103) are shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24565. These patients had fewer study visits and were older with a longer AS symptom duration compared to those who were included in the main analysis.

Longitudinal analysis. In the main analysis, higher BMI was significantly associated with a higher ASDAS over time (β 0.06 per 1 kg/m^2 [95% confidence interval (95% CI) 0.04–0.08], $P < 0.01$). Findings were similar for the association of BMI and BASDAI score (β 0.16 per 1 kg/m^2 [95% CI 0.10–0.22], $P < 0.01$). The results of the sensitivity analyses were consistent with the main analyses (see Supplementary Tables 2–3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24565).

Table 1. Baseline characteristics of ankylosing spondylitis cohort with available exposure and outcome measures (n = 183), stratified by overweight/obese body mass index (BMI)*

Variables	All patients (n = 183)	Normal/ underweight BMI (n = 84)	Overweight/ obese BMI (n = 99)
Demographics data			
Age, years	40.8 ± 13.3	35.2 ± 10.2	45.5 ± 13.8
Male sex, %	77	77	84
Race, %			
White	70	64	74
Black	1	1	0
Asian	16	21	11
Other	13	14	15
Number of study visits	7.8 ± 4.5	7.6 ± 4.1	8.0 ± 4.8
Disease characteristics			
Symptom duration, years	17.9 ± 12.5	14.8 ± 9.7	20.4 ± 13.8
Abnormal CRP, %†	36	27	44
BASDAI score (range 0–10)	3.6 ± 2.4	3.1 ± 2.3	3.9 ± 2.4
ASDAS (0.6 to >6.0)	1.6 ± 1.0	1.4 ± 0.9	1.7 ± 1.0
NSAID use, %	72	73	71
Glucocorticoid use, %	6	5	7
TNFi use, %	46	49	44
Cardiovascular disease and risk factors			
BMI, units	25.7 ± 4.8	20.1 ± 2.2	28.9 ± 4.0
Cardiovascular disease, %‡	2	3	1
Diabetes mellitus, %	2	1	3
Current smoker, %	4	4	5
Exercise, minutes/week	173.3 ± 179.8	200.8 ± 209.9	149.7 ± 146.3

* Values are the mean ± SD unless indicated otherwise. Baseline characteristics are first available. Data were missing for abnormal C-reactive protein (CRP; n = 1), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI; n = 2) score, Ankylosing Spondylitis Disease Activity Score (ASDAS; n = 3), diabetes mellitus (n = 3), cardiovascular disease (n = 3), and exercise (n = 1). NSAID = nonsteroidal antiinflammatory drug; TNFi = tumor necrosis factor inhibitor.
† CRP level was abnormal if above the upper limit of the reference range associated with the value.
‡ Cardiovascular disease was the composite of patient-reported coronary artery disease, coronary bypass surgery, coronary angioplasty, heart attack, heart valve problems, and angina.

Mediation analysis. In the causal mediation analysis, which examined exercise as a mediator between BMI and disease activity, the direct effect for an overweight/obese BMI compared to a normal BMI was a 0.36 (95% CI 0.23–0.51) unit difference in the ASDAS. The indirect effect for an overweight/obese BMI compared to a normal BMI was a 0.02 (95% CI 0.01–0.05) unit difference in ASDAS. The total effect of overweight/obese BMI, on average, was of a 0.39-point increase in the ASDAS. Exercise mediated 7% of the effect of an overweight or obese BMI on disease activity. Results were similar using BASDAI score as the outcome (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24565).

In the sensitivity analysis performed to assess the robustness of our mediation effect estimates against violations of the sequential ignorability assumption, a negative correlation of 0.1 between the error terms of the 2 regression models would be required for the indirect effect to be fully explained by unmeasured confounders.

DISCUSSION

To our knowledge, this was the first study to assess the association of BMI and disease activity longitudinally among individuals with axial SpA. We found that on average, higher BMI was significantly associated with higher disease activity (as measured by the ASDAS) in this prospective AS cohort. We found that being overweight/obese largely influences disease activity directly rather than indirectly through exercise. This suggests that mechanisms other than exercise, such as increased inflammation, may better explain this association. Interventions targeting obesity may improve both disease activity and CV risk in this population.

Our results overall are in concordance with our previously published systematic review and meta-analysis demonstrating an association between BMI and disease activity as reported by BASDAI score and ASDAS in axial SpA (13). Our findings are also similar to prior systematic reviews and meta-analyses in axial SpA evaluating the impact of obesity or higher BMI on various clinical outcome measures (11,12,37,38). Singh et al performed a meta-analysis of the impact of BMI on TNFi response in observational cohort studies and RCTs and found that in axial SpA, higher BMI was associated with increased odds of an inadequate response to TNFi treatment (6 studies; pooled odds ratio 3.36 [95% CI 1.33–8.51], I^2 = 81%) (11). Similar results were seen in a meta-analysis conducted by Shan and Zhang (12). The impact of obesity on imaging measures has also been explored. Bakirci et al performed a systematic review on the association of BMI and imaging-defined inflammation and damage in SpA including PsA (38). In 4 studies, higher BMI was associated with new syndesmophyte formation. In 1 study, higher BMI was also associated with a higher structural damage score by the modified Stoke Ankylosing Spondylitis Spine Score.

The association of high BMI with disease activity has been studied in PsA, with similar findings as in axial SpA. In prospective cohorts of PsA patients receiving TNFi therapy, Di Minno et al have shown that obesity was associated with a lower probability of achieving or maintaining minimal disease activity, and that a weight loss of ≥5% in overweight or obese individuals predicts better treatment response (39,40). Obesity was also associated with worse response to TNFi in large Scandinavian PsA registries (41). However, these studies were limited in their generalizability, as only individuals receiving TNFi treatment were included.

The biologic mechanisms that are believed to underlie obesity as a chronic inflammatory state involve the production of pro-inflammatory cytokines by adipose tissue (42). Higher BMI and higher fat mass have been associated with chronic pain in multiple populations (43–45). Obesity may be related to disease activity independent of inflammation, such as through mechanical loading and stress (46).

As prior studies were cross-sectional in nature, our study had the strength of using a long-standing prospective cohort of AS patients with multiple study visits (median total follow-up of 7 years). As an advantage over conventional GEE, we used IPW to account for potential time-varying confounding, such as with TNFi treatment. We also applied a causal mediation analysis to assess the extent to which the BMI impacts disease activity through exercise. This approach allows us to start disentangling the causal pathway between increased adiposity and high disease activity in AS. This has benefits over a more conventional mediation analysis because the latter requires the underlying assumption that the mediator does not interact with the exposure of interest (31). Causal mediation analysis also allows for the decomposition of the total effect into direct and indirect effects and the quantification of these effect estimates.

Certain limitations of this study must be acknowledged. We used observational data from a longitudinal cohort that had >20% missingness for the exposure and outcome of interest. With this degree of missingness, both complete case analysis and imputation strategies may produce biased results. With longitudinal data sets with repeated measures, incorrect specification of the multiple imputation model may additionally lead to bias in the results (29,47). However, the concordant results across sensitivity analyses using alternative data imputation approaches is reassuring in regard to the robustness of our results, although some of these estimates did not reach statistical significance. The mediation analysis relies on a set of strong assumptions, including no unmeasured confounders, no measurement error, no confounders of the mediator–outcome association, and the absence of postexposure confounding. Violations to these assumptions could also bias our results. We used a patient-reported measure of the frequency and duration of exercise, which is subject to both reporting bias and measurement error.

We were unable to capture intensity, type, and consistency of exercise for our study, nor were we able to provide quantification of exercise in METs. Although we accounted for NSAIDs and TNFi for each study interval, there may still be residual confounding. Finally, the exposure of interest, BMI, serves as a surrogate for adiposity, and individuals may reach a certain BMI through multiple possible mechanisms. Further studies should aim to target ≥1 of these mechanisms that fulfill the "well-defined intervention" assumption for causal inference (48).

In conclusion, in this longitudinal observational study, higher BMI was associated with higher disease activity among individuals with AS. We found that being overweight/obese had a primarily direct effect on disease activity that operates only weakly through exercise, which suggests that other mechanisms such as increased inflammation may explain this association. Future work should focus on the components of BMI and disease activity that are important in this relationship, particularly those factors that are modifiable and can be targeted by specific interventions. The measurement of physical activity and exercise should be taken into account in these studies as well, including the consideration of intensity, type, and consistency.

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. Liew 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. Liew, Gianfrancesco, Heckbert, Gensler.

Acquisition of data. Gensler.

Analysis and interpretation of data. Liew, Gianfrancesco, Heckbert, Gensler.

REFERENCES

1. Poddubnyy D, Haibel H, Listing J, Märker-Hermann E, Zeidler H, Braun J, et al. Baseline radiographic damage, elevated acute-phase reactant levels, and cigarette smoking status predict spinal radiographic progression in early axial spondylarthritis. Arthritis Rheum 2012;64:1388–98.
2. Dougados M, Sepriano A, Molto A, van Lunteren M, Ramiro S, de Hooge M, et al. Sacroiliac radiographic progression in recent onset axial spondyloarthritis: the 5-year data of the DESIR cohort. Ann Rheum Dis 2017;76:1823–8.
3. Haroon N, Inman RD, Learch TJ, Weisman MH, Lee M, Rahbar MH, et al. The impact of tumor necrosis factor α inhibitors on radiographic progression in ankylosing spondylitis. Arthritis Rheum 2013;65:2645–54.
4. Bakland G, Gran JT, Nossent JC. Increased mortality in ankylosing spondylitis is related to disease activity. Ann Rheum Dis 2011;70:1921–5.
5. Bremander A, Petersson IF, Bergman S, Englund M. Population-based estimates of common comorbidities and cardiovascular disease in ankylosing spondylitis. Arthritis Care Res (Hoboken) 2011;63:550–6.
6. Haroon NN, Paterson JM, Li P, Inman RD, Haroon N. Patients with ankylosing spondylitis have increased cardiovascular and cerebrovascular mortality: a population-based study. Ann Intern Med 2015;163:409–16.
7. Bengtsson K, Forsblad-d'Elia H, Lie E, Klingberg E, Dehlin M, Exarchou S, et al. Are ankylosing spondylitis, psoriatic arthritis and undifferentiated spondyloarthritis associated with an increased risk of cardiovascular events? A prospective nationwide population-based cohort study. Arthritis Res Ther 2017;19:102.
8. Mathieu S, Soubrier M. Cardiovascular events in ankylosing spondylitis: a 2018 meta-analysis. Ann Rheum Dis 2019;78:e57.
9. Whelton PK, Carey RM, Aronow WS, Ovbiagele B, Casey DE, Smith SC, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice. Hypertension 2018;71:1269–324.
10. Mease P, van der Heijde D, Karki C, Liu M, Park Y, Greenberg J. Tumor necrosis factor inhibition discontinuation in patients with ankylosing spondylitis: an observational study from the US-based Corrona Registry. Rheumatol Ther 2018;5:537–50.
11. Singh S, Facciorusso A, Singh AG, Casteele NV, Zarrinpar A, Prokop LJ, et al. Obesity and response to anti-tumor necrosis factor-α agents in patients with select immune-mediated inflammatory diseases: a systematic review and meta-analysis. PLoS One 2018;13:1–26.
12. Shan J, Zhang J. Impact of obesity on the efficacy of different biologic agents in inflammatory diseases: a systematic review and meta-analysis. Joint Bone Spine 2019;86:173–83.
13. Liew JW, Huang IJ, Louden DN, Singh N, Gensler LS. Association of body mass index on disease activity in axial spondyloarthritis: systematic review and meta-analysis. RMD Open 2020;6:1–12.
14. Garcia JM, Duran AT, Schwartz JE, Booth JN, Hooker SP, Willey JZ, et al. Types of sedentary behavior and risk of cardiovascular events and mortality in blacks: the Jackson Heart Study. J Am Heart Assoc 2019;8:e010406.
15. Bellettiere J, Lamonte MJ, Evenson KR, Rillamas-Sun E, Kerr J, Lee IM, et al. Sedentary behavior and cardiovascular disease in older women: the OPACH study. Circulation 2019;139:1036–46.
16. Fongen C, Halvorsen S, Dagfinrud H. High disease activity is related to low levels of physical activity in patients with ankylosing spondylitis. Clin Rheumatol 2013;32:1719–25.
17. Coulter EH, McDonald MT, Cameron S, Siebert S, Paul L. Physical activity and sedentary behaviour and their associations with clinical measures in axial spondyloarthritis. Rheumatol Int 2020;40:375–81.
18. Sveaas SH, Bilberg A, Berg IJ, Provan SA, Rollefstad S, Semb AG, et al. High intensity exercise for 3 months reduces disease activity in axial spondyloarthritis (axSpA): a multicentre randomised trial of 100 patients. Br J Sports Med 2020;54:292–7.
19. 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.
20. Lee W, Reveille JD, Davis JC, Learch TJ, Ward MM, Weisman MH. Are there gender differences in severity of ankylosing spondylitis? Results from the PSOAS cohort. Ann Rheum Dis 2007;66:633–8.
21. Reveille JD, Lee MJ, Gensler LS, Ward MM, Hwang MC, Learch TJ, et al. The changing profile of ankylosing spondylitis in the biologic era. Clin Rheumatol 2020;39:2641–51.
22. 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.

23. Machado PM, Landewé R, van der Heijde D. Ankylosing Spondylitis Disease Activity Score (ASDAS): 2018 update of the nomenclature for disease activity states. Ann Rheum Dis 2018;77:1539–40.
24. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
25. Lukas C, Landewé R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68: 18–24.
26. Ward MM. Predictors of the progression of functional disability in patients with ankylosing spondylitis. J Rheumatol 2002;29:1420–5.
27. Van Buuren S. Flexible imputation of missing data. Boca Raton: Chapman & Hall/CRC; 2012.
28. Aloisio KM, Swanson SA, Micali N, Field A, Horton NJ. Analysis of partially observed clustered data using generalized estimating equations and multiple imputation. Stata J 2014;14:863–83.
29. De Silva AP, Moreno-Betancur M, de Livera AM, Lee KJ, Simpson JA. A comparison of multiple imputation methods for handling missing values in longitudinal data in the presence of a time-varying covariate with a non-linear association with time: a simulation study. BMC Med Res Methodol 2017;17:1–11.
30. Hayes AF. Beyond Baron and Kenny: statistical mediation analysis in the new millennium. Commun Monogr 2009;76:408–20.
31. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 2013;18:137–50.
32. Hicks R, Tingley D. Causal mediation analysis. Stata J 2011;11: 605–19.
33. Imai K, Keele L, Yamamoto T. Identification, inference and sensitivity analysis for causal mediation effects. Stat Sci 2010;25:51–71.
34. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
35. Tingley D. R and Stata for causal mechanisms. Harvard University; 2012. p. 1–12.
36. Geskus R, van der Wal W. ipw: an R package for inverse probability weighting. J Stat Softw 2011;43:1–23.
37. Lee YX, Kwan YH, Lim KK, Tan CS, Lui NL, Phang JK, et al. A systematic review of the association of obesity with the outcomes of inflammatory rheumatic diseases. Singapore Med J 2019;60:270–80.
38. Bakirci S, Dabague J, Eder L, McGonagle D, Aydin SZ. The role of obesity on inflammation and damage in spondyloarthritis: a systematic literature review on body mass index and imaging. Clin Exp Rheumatol 2019;1–6.
39. Di Minno MN, Peluso R, Iervolino S, Russolillo A, Lupoli R, Scarpa R. Weight loss and achievement of minimal disease activity in patients with psoriatic arthritis starting treatment with tumour necrosis factor α blockers. Ann Rheum Dis 2014;73:1157–62.
40. Di Minno MN, Peluso R, Iervolino S, Lupoli R, Russolillo A, Scarpa R, et al. Obesity and the prediction of minimal disease activity: a prospective study in psoriatic arthritis. Arthritis Care Res (Hoboken) 2013;65:141–7.
41. Højgaard P, Glintborg B, Kristensen LE, Gudbjornsson B, Love TJ, Dreyer L. The influence of obesity on response to tumour necrosis factor- a inhibitors in psoriatic arthritis: results from the DANBIO and ICEBIO registries. Rheumatology (Oxford) 2018;55:2191–9.
42. Xu H, Barnes G, Yang Q, Tan G, Yang D, Chou C, et al. Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. J Clin Invest 2003;112:1821–30.
43. Stone AA, Broderick JE. Obesity and pain are associated in the United States. Obesity 2012;20:1491–5.
44. Heuch I, Heuch I, Hagen K, Zwart JA. Body mass index as a risk factor for developing chronic low back pain: a follow-up in the Nord-Trøndelag Health Study. Spine 2013;38:133–9.
45. Yoo JJ, Cho NH, Lim SH, Kim HA. Relationships between body mass index, fat mass, muscle mass, and musculoskeletal pain in community residents. Arthritis Rheumatol 2014;66:3511–20.
46. Berenbaum F, Eymard F, Houard X. Osteoarthritis, inflammation and obesity. Curr Opin Rheumatol 2013;25:114–8.
47. Young R, Johnson DR. Handling missing values in longitudinal panel data with multiple imputation. J Marriage Fam 2015;77:277–94.
48. Hernan M, Taubman S. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. Int J Obes 2008;32:S8–14.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1294–1299
DOI 10.1002/acr.24573

BRIEF REPORT

Increased Mortality for Individuals With Giant Cell Arteritis: A Population-Based Study

Lillian Barra,[1] Janet E. Pope,[1] Priscila Pequeno,[2] Jodi M. Gatley,[2] and Jessica Widdifield[3]

Objective. Reports of mortality risks among individuals with giant cell arteritis (GCA) have been mixed. Our aim was to evaluate all-cause mortality among individuals with GCA relative to the general population over time.

Methods. We performed a population-based study in Ontario, Canada using health administrative data. We studied a cohort of 22,677 GCA patients ages ≥50 years that was identified using a validated case definition (with 81% positive predictive value, 100% specificity). General population comparators were residents ages ≥50 years without GCA. Deaths were ascertained from vital statistics. Annual crude, age- and sex-standardized, and age- and sex-specific all-cause mortality rates were determined for individuals with and without GCA between 2000 and 2018. Standardized mortality ratios (SMRs) were estimated.

Results. Age- and sex-standardized mortality rates were significantly higher for GCA patients than comparators, and trending to increase over time with 50.0 deaths per 1,000 GCA patients in 2000 (95% confidence interval [95% CI] 34.0–71.1) and 57.6 deaths per 1,000 GCA patients in 2018 (95% CI 50.8–65.2), whereas mortality rates in the general population significantly declined over time. The annual SMRs for GCA patients generally increased over time, with the lowest SMR occurring in 2002 (1.22 [95% CI 1.03–1.40]) and the highest in 2018 (1.92 [95% CI 1.81–2.03]). GCA mortality rates were more elevated for male patients than female patients.

Conclusion. Over a 19-year period, mortality rates were increased among GCA patients relative to the general population, and more premature deaths were occurring in younger age groups. The relative excess mortality for GCA patients did not improve over time.

INTRODUCTION

Giant cell arteritis (GCA) is the most common vasculitis affecting individuals over the age of 50 years, with incidence rates of 20–35 per 100,000 individuals reported in North American and Northern European populations (1–3). We have previously shown that the yearly prevalence of GCA is increasing in Canada. The reasons for this increase remain unclear, but include an aging population, population growth (and an increase in the number of individuals at risk), improved diagnosis, or increased survival (1).

GCA is characterized by inflammation of large and medium vessels with a predilection for extradural cranial arteries and is associated with an increased risk of morbidity, including blindness and stroke (4). The disease can cause aortic aneurysms and dissection with high mortality (5). In addition, the mainstay of treatment is long-term glucocorticoids which may increase risk of adverse events, such as severe infections, osteoporotic fractures, hypertension, diabetes mellitus, weight gain, and other metabolic changes, leading to more cardiovascular events (4).

Despite GCA and its treatment being associated with complications known to increase the risk of death, prior research on

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the Canadian Initiative for Outcomes in Rheumatology Care, ICES, the Ontario Ministry of Health and Long-Term Care, or the Canadian Institutes of Health Information.

Supported by the Canadian Initiative for Outcomes in Rheumatology Care (grant 2018-002), who played no role in the design or conduct of the study other than providing peer review of the study proposal, and ICES, formerly known as the Institute for Clinical Evaluative Sciences (annual grant from the Ontario Ministry of Health and Long-Term Care). Dr. Widdifield's work was supported by the Arthritis Society (Stars Career Development award [STAR-19-0610]). The data reported herein have been supplied in part by the Ontario Ministry of Health and Long-Term Care and Canadian Institutes of Health Information.

[1]Lillian Barra, MD, FRCPC, MPH, Janet E. Pope, MD, FRCPC, MPH: Western University and St. Joseph's Health Care, London, Ontario, Canada; [2]Priscila Pequeno, MSc, Jodi M. Gatley, MPH: Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada; [3]Jessica Widdifield, PhD: Institute for Clinical Evaluative Sciences, Sunnybrook Research Institute and University of Toronto, Toronto, Ontario, Canada.

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

Address correspondence to Lillian Barra, MD, FRCPC, MPH, 268 Grosvenor Street, London, Ontario N6A 4V2, Canada. Email: lillian.barra@sjhc.london.on.ca.

Submitted for publication August 24, 2020; accepted in revised form February 2, 2021.

SIGNIFICANCE & INNOVATIONS

- Over 19 years, mortality has remained increased among giant cell arteritis (GCA) patients relative to the general population.
- GCA mortality rates were highest among male patients, and standardized mortality ratios were highest for younger age groups.
- The relative excess mortality for GCA patients (mortality gap) did not improve over time.

mortality among individuals with GCA has demonstrated conflicting findings. A recent meta-analysis reported an increased risk of mortality in studies of GCA patients in hospitalized settings but conflicting reports for population-based studies (6). The population studies in the meta-analysis by Hill et al included diverse methodologies, with studies often confined to inception cohorts of various follow-up and including variable use of general population comparators. Moreover, some studies involving more contemporary cohorts have detected increased mortality risk in GCA (7, 8) and conflicting reports on sex-related differences in mortality (6,7,9,10). The aim of our study was to evaluate all-cause, age- and sex-specific mortality among individuals with GCA in Ontario, Canada compared to the general population over time and to determine whether there have been changes to the relative excess mortality among GCA patients over time.

PATIENTS AND METHODS

Study design and setting. We conducted a population-based retrospective cohort study in Ontario, Canada. There are ~13 million residents (as of 2018) who are covered under a publicly funded health care system within Canada's most populous province. The use of data in this project was authorized under section 45 of Ontario's Personal Health Information Protection Act, which does not require review by a research ethics board.

Data sources. Population-level health administrative databases capture details of the health care of Ontarians. The Ontario Health Insurance Plan (OHIP) database contains physician claims for services provided, dates of services, and associated diagnosis and procedure codes. The Canadian Institute for Health Information discharge abstract database contains patient-level data for acute, chronic, rehab and day surgery institutions in Ontario, including hospitalizations for GCA (International Classification of Diseases, Ninth Revision [ICD-9] code 446.5 and ICD-10 codes M31.5 or M31.6). All hospital data prior to April 1, 2002 have diagnoses coded in the ICD-9. Hospitalizations after April 1, 2002 are coded using ICD-10-CA (a version of the ICD-10 developed by the Canadian Institute for Health Information). The Ontario drug benefit database contains prescription drug claims dispensed for those individuals ages 65 years and older, including dispensations for glucocorticoids. The ICES (formerly known as the Institute for Clinical Evaluative Sciences) physician database contains information on physician specialties. The OHIP registered persons database was used to obtain information about health insurance status, age, sex, and vital status. General population comparators were derived from the Registered Persons Database. These data sets were linked using unique encoded identifiers and analyzed at ICES in Toronto, Canada (https://www.ices.on.ca/).

Patient population. Patients with GCA ages 50 years and older were identified between April 1, 2000 and March 31, 2019. GCA patients were identified from health administrative data using a previously validated case definition (1). This case definition identified GCA based on the following criteria: ≥1 hospitalization with a primary or secondary diagnosis of GCA ever; or ≥2 physician diagnoses of GCA (with ≥1 diagnosis by a specialist [rheumatologist, internist, or ophthalmologist]) and ≥1 prescription for glucocorticoids; or ≥1 temporal artery biopsy within 3 years. This case definition also excluded individuals with any kidney, lung, skin, or nasal biopsies within the 1-year period of a physician diagnosis code in order to exclude other vasculitis types (1). This case definition performed with an 81% positive predictive value (PPV), 60% sensitivity, 100% specificity, and 99% negative predictive value (NPV), using a validation of randomly selected charts from a large primary care population in Ontario (comprising 143 patients with a physician-documented GCA diagnosis in their medical record and 7,534 non-cases) (1). The GCA cases from the validation cohort had a mean ± SD age of 79 ± 9 years, and 78% were female. The age of onset and female sex predilection supports previous epidemiologic data (2).

Cohort entry was defined as the date of the first hospital admission with a GCA diagnosis or the second physician outpatient billing claim for GCA, whichever came first. The general population comparators included all Ontario residents ages 50 years and older with valid Ontario health insurance and excluded those with GCA.

Statistical analysis. For both the GCA and general population cohorts (ages ≥50 years), annual crude observed mortality rates with corresponding 95% confidence intervals (95% CIs) were calculated as the number of deaths per 1,000 individuals in the population each fiscal year. Annual observed GCA mortality rates were determined by dividing the number of deaths among individuals with GCA each year by the number of individuals with GCA in each year. The denominator included all GCA patients (prevalent and new cases identified each year). Individuals who died during follow-up were excluded from the denominator in the subsequent calendar years after their death. Similar analyses were done to obtain observed mortality rates in the general population without GCA. To adjust for differences in population

distribution over time, direct age- and sex-standardization was undertaken using 2001 Ontario census population estimates.

Age-specific mortality rates were also determined using 10-year age bins, except for the age groups including patients ages 50–64 years, which were combined due to fewer deaths occurring within each year, resulting in the following age bins: 50–64, 65–74, 75–84, and 85 years of age and older. Age was determined on April 1 of each year from 2000 to 2018. To assess differences in age- and sex-specific mortality rates, we divided the study period into smaller periods and reported the results for 2000, 2009, and 2018.

Annual standardized mortality ratios (SMRs) were calculated as the ratio of the observed number of deaths in GCA patients to the expected number of deaths determined from our general population comparator cohort within each age-specific stratum each year. All analyses were performed using SAS, version 9.4.

RESULTS

In the period between April 1, 2000 and March 31, 2019, there were a total of 22,677 GCA patients (across all years), of whom 14,195 (63%) were female. The mean ± SD age of GCA patients was 73.6 ± 9.3 years at cohort entry.

During the study period, the annual number of individuals ages ≥50 years without GCA (the general population denominator) increased from 3,274,834 in 2000 to 5,454,966 in 2018, and the GCA population increased from 1,518 to 12,792, comprising both incident and prevalent cases (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24573). There was a median of 1,014 new patients (range 861–1662) diagnosed annually and being added to the GCA population denominator each year. The crude GCA mortality rates increased over time from 63 deaths (95% CI 51.8–78.0) per 1,000 GCA patients in the early 2000s to 88.2 deaths (95% CI 83.1–93.5) per 1,000 GCA patients by 2018, whereas the crude mortality rates for the general population significantly decreased over time from 22.0 deaths (95% CI 21.9–22.2) in 2000 to 17.7 deaths (95% CI 17.5–17.8) per 1,000 individuals in the population ≥50 years of age in 2018 (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24573).

Age- and sex-standardized mortality rates were significantly higher for GCA patients than comparators (Figure 1) and trending to increase over time with 50.0 deaths (95% CI 34.0–71.1) per 1,000 GCA patients in 2000 and 57.6 deaths (95% CI 50.8–65.2) per 1,000 GCA patients in 2018. Among the general population without GCA, age- and sex-standardized mortality rates significantly declined over time, with 22.1 deaths (95% CI 21.9–22.2) per 1,000 individuals in the population in 2000 and 16.4 deaths (95% CI 16.3–16.5) per 1,000 individuals in the population in 2018 (Figure 1; see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24573).

Age-specific mortality rates increased with advancing age for both cohorts but were significantly higher for GCA patients than comparators within each age strata (Table 1). For all ages across 2000, 2009, and 2018, the excess mortality rates were 41.8 (95% CI 29.9–55.8), 57.1 (95% CI 51.3–63.3), and 70.5 (95% CI 65.6–75.7) deaths per 1,000 individuals in the population, respectively. Mortality rates were also higher for male patients than female patients within both cohorts. Male patients with GCA had a higher all-cause mortality rate ranging from 50.8 (95% CI 32.9–75.0) among those ages 50–64 years to 192.5 (95% CI 165.4–222.8)

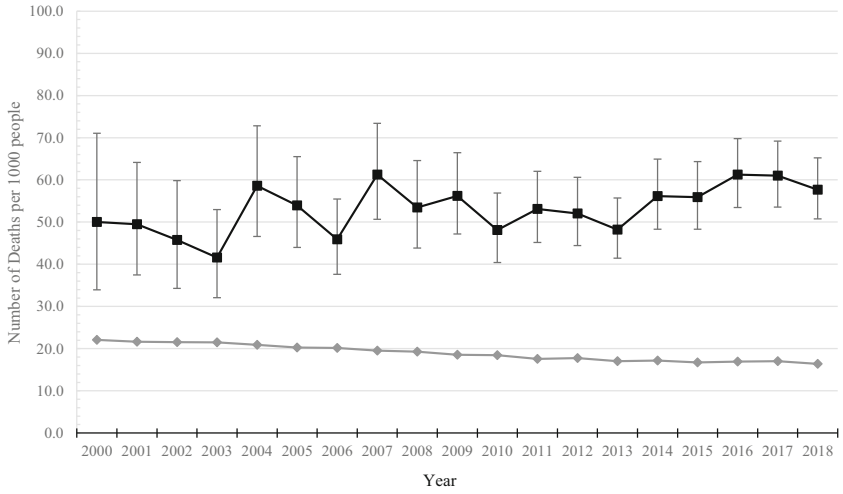


Figure 1. Age- and sex-standardized all-cause mortality rates per 1,000 individuals in the population for individuals with and without giant cell arteritis (GCA). Data are shown as box plots. Squares represent the mortality rate among GCA patients. Diamonds represent the age- and sex-standardized all-cause mortality rate among individuals in the general population. Whiskers represent the 95% confidence interval (95% CI) among the mortality rate in the GCA population (CI around rates in the general population too narrow to visually depict).

Table 1. Age- and sex-specific mortality rates by years for individuals with and without giant cell arteritis (GCA)*

Age group, years	Year 2000		Year 2009		Year 2018	
	General population death rate per 1,000 individuals in the population	GCA death rate per 1,000 GCA patients	General population death rate per 1,000 individuals in the population	GCA death rate per 1,000 GCA patients	General population death rate per 1,000 individuals in the population	GCA death rate per 1,000 GCA patients
All ages	22.1 (21.9–22.2)	63.9 (51.8–78.0)	18.8 (18.7–19.0)	75.9 (70.0–82.3)	17.7 (17.5–17.8)	88.2 (83.1–93.5)
50–64	5.8 (5.7–5.9)	30.7 (12.3–63.3)	5.0 (4.9–5.1)	40.8 (28.4–56.7)	4.6 (4.6–4.7)	41.4 (30.9–54.3)
65–74	20.1 (19.8–20.4)	37.2 (23.0–56.8)	15.6 (15.3–15.8)	45.1 (36.9–54.7)	13.2 (13.0–13.4)	53.6 (46.3–61.8)
75–84	50.6 (50.0–51.3)	83.2 (61.9–109.4)	43.4 (42.9–44.0)	70.6 (61.9–80.1)	36.1 (35.6–36.5)	82.3 (74.5–90.8)
≥85	137.9 (136.0–139.8)	160.7 (95.2–254.0)	120.3 (118.8–121.7)	164.8 (143.8–188.0)	116.0 (114.9–117.2)	156.2 (142.6–170.8)
Male sex						
All ages	23.4 (23.1–23.6)	77.5 (70.0–82.3)	19.7 (19.5–19.9)	94.9 (84.1–107.1)	18.6 (18.4–18.8)	107.0 (98.3–116.4)
50–64	7.1 (6.9–7.2)	47.6 (9.8–139.2)	6.2 (6.0–6.3)	58.4 (33.4–94.8)	5.7 (5.6–5.9)	50.8 (32.9–75.0)
65–74	25.4 (24.9–25.9)	44.9 (19.4–88.6)	19.0 (18.6–19.4)	72.1 (54.6–93.4)	15.9 (15.6–16.2)	71.1 (57.4–87.1)
75–84	64.0 (62.9–65.1)	93.6 (53.5–151.9)	52.5 (51.6–53.4)	89.9 (73.0–109.6)	42.9 (42.2–43.6)	103.8 (89.2–120.2)
≥85	157.7 (154.1–161.5)	259.3 (104.2–534.2)	135.2 (132.5–137.9)	187.3 (145.5–237.5)	126.4 (124.4–128.4)	192.5 (165.4–222.8)
Female sex						
All ages	20.9 (20.7–21.1)	58.4 (46.0–74.2)	18.1 (17.9–18.3)	67.0 (60.6–74.0)	16.8 (16.7–17.0)	78.1 (72.6–84.1)
50–64	4.5 (4.4–4.7)	24.2 (6.6–62.1)	3.9 (3.8–4.0)	32.5 (19.6–50.7)	3.6 (3.5–3.7)	35.4 (23.3–51.5)
65–74	15.4 (15.0–15.8)	33.6 (17.9–57.4)	12.4 (12.1–12.8)	31.1 (22.8–41.3)	10.7 (10.4–10.9)	43.6 (35.4–53.0)
75–84	41.7 (41.0–42.5)	79.2 (55.2–110.1)	36.5 (35.9–37.2)	61.3 (51.6–72.4)	30.4 (29.9–31.0)	70.5 (61.5–80.5)
≥85	129.4 (127.2–131.6)	129.4 (64.6–231.6)	113.2 (111.5–114.9)	156.4 (132.6 – 183.3)	110.1 (108.7–111.5)	140.5 (125.2–157.3)

* Values are the age- and sex-specific mortality rates (95% confidence interval).

among those ages >85 years. Comparatively, the all-cause mortality rate in female patients with GCA ranged from 35.4 (95% CI 23.3–51.5) among those ages 50–64 years to 140.5 (95% CI 125.2–157.3) among those ages >85 years per 1,000 patients in the GCA population (Table 1). Similar patterns were observed in the general population; however, the age-specific mortality rates, overall and by sex, were lower.

The annual SMRs for GCA patients increased over time, with the lowest SMR occurring in 2002 (1.22 [95% CI 1.03–1.40]) and the highest in 2018 (1.92 [95% CI 1.81–2.03]) (Figure 2). The age-specific SMRs were most elevated in the youngest age group (ages 50–64 years) and declined by the time patients were ages ≥85 years (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24573). Sex-specific SMRs for both male and female patients showed similar patterns across age groups (data not shown).

DISCUSSION

In this large population-based study, we found that GCA patients had a significantly higher annual risk of mortality compared to the general population over the 19-year study period. The increased mortality risk for GCA patients was seen in male and female patients and was most pronounced in individuals ages 50–65 years. Although mortality rates decreased over time for the general population, they trended upwards for GCA patients, resulting in higher SMRs for GCA patients in recent years.

The SMRs for GCA patients in our study were higher than what has been reported in prior studies (6,7). A recent meta-analysis with a total of 4,733 GCA patients reported an overall pooled mortality ratio of 1.17 (95% CI 1.02–1.35). The included studies were highly heterogeneous (6), and higher SMRs have been reported in several other studies (10,11). The relative excess mortality for GCA remains controversial, with regional differences

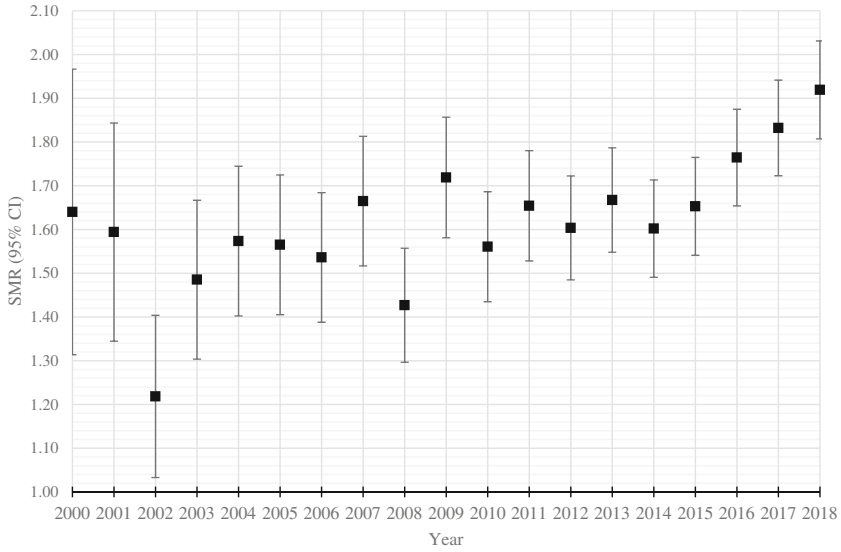


Figure 2. Annual standardized mortality ratios (SMRs) of patients with giant cell arteritis. Data are shown as box plots. Squares represent the annual SMR. Whiskers represent the 95% confidence interval (95% CI).

also being a possible contributing factor (12). Large population-based studies from Northern Europe (n = 840 to 9,778) have reported an increased mortality risk for GCA patients compared to the general population of ~20–50% in the first 1–2 years after the diagnosis, and then the risk decreased over time (7,13). Our study showed that the annual SMRs for GCA in Ontario, Canada were consistently elevated and increasing (by 1.2 to 2 times compared to non-GCA individuals) over a 19-year time period; whereas, a study from the UK based on an inception cohort did not find an increased mortality risk for GCA by calendar year (7).

Similar to the findings in our study, the increased mortality risk for GCA has been shown to be highest in younger individuals; the background mortality rates for such individuals in the general population are low (7,10). In the general population, it is known that men have higher mortality rates than women. Mohammad et al reported that women, but not men, with GCA in Sweden had an increased SMR (10), whereas Andersen et al reported men potentially having worse survival rates than women (9), and other reports did not find any excess mortality risk based on sex (6,7).

The variability in the reported GCA mortality rate can be due to several reasons, including the study design and the characteristics of the study population. GCA is not a uniform disease, and it transitions from acute into chronic vasculitis with long-term systemic and vascular inflammation (14). Some inception-based cohorts may be less likely to detect excess mortality, and hospital-based cohorts may have patients with more severe GCA or multiple comorbidities known to be associated with GCA (such as cardiovascular disease and diabetes mellitus) that can contribute to excess mortality (7). Some studies had higher proportions of biopsy-proven GCA, which may have a different prognosis than biopsy-negative disease (6,15). In addition, the geographic location of the studies likely contributed through population differences in genetic backgrounds, environmental exposures, socioeconomic status, and health care services.

A major strength of our study is that we captured mortality data over a long period of time for the entire population of the most populous province of Canada, Ontario, with more than 5 million individuals ages ≥50 years. Due to legal requirements, all deaths are accurately captured, and <1% of individuals die out of the province. Inherent limitations of using health administrative databases to ascertain cases of GCA need to be recognized. These include the potential for misclassification bias because of errors in diagnosis codes and incompleteness of administrative databases. Although our cohort is limited in that we do not have access to information regarding temporal artery biopsy results, the GCA definition we utilized was validated and has a high PPV, NPV, and specificity. Our administrative data case definition does not have 100% sensitivity in capturing GCA, and thus may miss some GCA patients, such as those patients whose symptoms have resolved and who (for these or other reasons) are no longer seeking care for their GCA. While our previous validation study yielded a case definition with high PPVs (especially for a rare disease), the potential for misclassification within our GCA cohort still exists. The OHIP diagnosis code for GCA is not limited to GCA but may include polyarteritis nodosa, isolated aortitis, or other vasculitides. We presume that the impact from such misclassification on our estimates is minor given the rarity of these other vasculitis conditions.

The purpose of this study was to determine all-cause mortality for patients with GCA relative to the general population. We did not report causes of death, and thus it is unknown if the excess mortality observed in our study was attributable to complications of GCA or treatment effects. Other studies have found that GCA

patients had a higher risk of death due to cardiovascular disease compared to the general population (5,8). Risk factors associated with an increased mortality in GCA populations included vascular, respiratory, thromboembolic, renal, and inflammatory musculoskeletal disease, cancer, diabetes mellitus, infections, and smoking (5,6). To perform analyses of predictors of mortality fell outside the aim of this study; we had limited access to data on lifestyle factors and did not include data on comorbidities or different clinical phenotypes of GCA. In future work, we will explore causes of death, predictors for mortality, and causes for the observed temporal increase in GCA mortality compared to the general population.

In our study conducted over a 19-year period, we observed mortality rates to be increased among GCA patients relative to the general population, and more premature deaths were occurring in younger age groups. The relative excess mortality for GCA patients did not improve over time.

ACKNOWLEDGMENTS

The authors thank IQVIA Solutions Canada, Inc. for use of their Drug Information File.

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. Barra 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. Barra, Pope, Widdifield.
Acquisition of data. Pequeno, Gatley, Widdifield.
Analysis and interpretation of data. Barra, Pope, Pequeno, Gatley, Widdifield.

REFERENCES

1. Barra L, Pope JE, Pequeno P, Saxena FE, Bell M, Haaland D, et al. Incidence and prevalence of giant cell arteritis in Ontario, Canada. Rheumatology (Oxford) 2020;59:3250–8.
2. Lee JL, Naguwa SM, Cheema GS, Gershwin ME. The geoepidemiology of temporal (giant cell) arteritis. Clin Rev Allergy Immunol 2008;35:88–95.
3. Sharma A, Mohammad AJ, Turesson C. Incidence and prevalence of giant cell arteritis and polymyalgia rheumatica: a systematic literature review. Semin Arthritis Rheum 2020;50:1040–8.
4. Borchers AT, Gershwin ME. Giant cell arteritis: a review of classification, pathophysiology, geoepidemiology and treatment. Autoimmun Rev 2012;11:A544–54.
5. Nuenninghoff DM, Hunder GG, Christianson TJ, McClelland RL, Matteson EL. Mortality of large-artery complication (aortic aneurysm, aortic dissection, and/or large-artery stenosis) in patients with giant cell arteritis: a population-based study over 50 years. Arthritis Rheum 2003;48:3532–7.
6. Hill CL, Black RJ, Nossent JC, Ruediger C, Nguyen L, Ninan JV, et al. Risk of mortality in patients with giant cell arteritis: a systematic review and meta-analysis. Semin Arthritis Rheum 2017;46:513–9.
7. Li L, Neogi T, Jick S. Mortality in patients with giant cell arteritis: a cohort study in UK primary care. Arthritis Care Res (Hoboken) 2018; 70:1251–6.
8. Ben-Shabat N, Tiosano S, Shovman O, Comaneshter D, Shoenfeld Y, Cohen AD, et al. Mortality among patients with giant cell arteritis: a large-scale population-based cohort study. J Rheumatol 2020;47: 1385–91.
9. Andersen JB, Myklebust G, Haugeberg G, Pripp AH, Diamantopoulos AP. Incidence trends and mortality of giant cell arteritis in southern Norway. Arthritis Care Res (Hoboken) 2021;73: 409–14.
10. Mohammad AJ, Nilsson JA, Jacobsson LT, Merkel PA, Turesson C. Incidence and mortality rates of biopsy-proven giant cell arteritis in southern Sweden. Ann Rheum Dis 2015;74:993–7.
11. Uddhammar A, Eriksson AL, Nystrom L, Stenling R, Rantapaa-Dahlqvist S. Increased mortality due to cardiovascular disease in patients with giant cell arteritis in northern Sweden. J Rheumatol 2002;29:737–42.
12. Breuer GS, Poltorak V, Nesher G. Survival of patients with giant cell arteritis: a controversial issue. Clin Exp Rheumatol 2020;38 Suppl 124:210–3.
13. Baslund B, Helleberg M, Faurschou M, Obel N. Mortality in patients with giant cell arteritis. Rheumatology (Oxford) 2015;54:139–43.
14. Maleszewski JJ, Younge BR, Fritzlen JT, Hunder GG, Goronzy JJ, Warrington KJ, et al. Clinical and pathological evolution of giant cell arteritis: a prospective study of follow-up temporal artery biopsies in 40 treated patients. Mod Pathol 2017;30:788–96.
15. Gonzalez-Gay MA, Garcia-Porrua C, Llorca J, Gonzalez-Louzao C, Rodriguez-Ledo P. Biopsy-negative giant cell arteritis: clinical spectrum and predictive factors for positive temporal artery biopsy. Semin Arthritis Rheum 2001;30:249–56.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1300–1310
DOI 10.1002/acr.24567

Barriers and Facilitators to Physical Activity for People With Scleroderma: A Scleroderma Patient-Centered Intervention Network Cohort Study

Sami Harb,[1] Sandra Peláez,[2] Marie-Eve Carrier,[3] Linda Kwakkenbos,[4] Susan J. Bartlett,[5] Marie Hudson,[1] Luc Mouthon,[6] Maureen Sauvé,[7] Joep Welling,[8] Ian Shrier,[3] and Brett D. Thombs,[1]
for the SPIN Physical Activity Enhancement Patient Advisory Team and SPIN Investigators

Objective. To support physical activity among people with systemic sclerosis (SSc [scleroderma]), we sought to determine the prevalence and importance of barriers and the likelihood of using possible facilitators.

Methods. We invited 1,707 participants from an international SSc cohort to rate the importance of 20 barriers (14 medical, 4 social or personal, 1 lifestyle, and 1 environmental) and the likelihood of using 91 corresponding barrier-specific and 12 general facilitators.

Results. Among 721 respondents, 13 barriers were experienced by ≥25% of participants, including 2 barriers (fatigue and Raynaud's phenomenon) rated "important" or "very important" by ≥50% of participants, 7 barriers (joint stiffness and contractures, shortness of breath, gastrointestinal problems, difficulty grasping, pain, muscle weakness and mobility limitations, and low motivation) by 26–50%, and 4 barriers by <26%. Overall, 23 of 103 facilitators (18 medical-related) were rated by ≥75% of participants as "likely" or "very likely" to use among those who experienced corresponding barriers. These facilitators focused on adapting exercise (e.g., using controlled, slow movement), taking care of one's body (e.g., stretching), keeping warm (e.g., wearing gloves), and protecting skin (e.g., covering ulcers). Among those participants who had previously tried the facilitator, all facilitators were rated by ≥50% as "likely" or "very likely" to use. Among those participants with the barrier who had not tried the facilitator, only 12 of 103 facilitators were rated by >50% of participants as "likely" or "very likely" to use.

Conclusion. Medical-related physical activity barriers were common and considered important. Facilitators considered as most likely to be used involved adapting exercise, taking care of one's body, keeping warm, and protecting skin.

INTRODUCTION

Systemic sclerosis (SSc [scleroderma]) is a rare, chronic, autoimmune rheumatic disease characterized by abnormal fibrotic processes and excessive collagen production that can affect the skin, musculoskeletal system, and internal organs, including the heart, lungs, and gastrointestinal tract (1,2). People with SSc experience significantly lower health-related quality of life in comparison to the general population (3). Disease onset typically occurs at ~50 years of age, and ~80% of people with SSc are women (4,5).

Although regular physical activity is important to enhance health for all people (6,7), including those with autoimmune

Supported by a Canadian Institutes of Health Research – Strategy for Patient-Oriented Research Grant (Dr. Shrier) with partner funding from the Scleroderma Society of Ontario. Dr. Thombs' work was supported by a Fonds de Recherche du Québec - Santé researcher salary award.

[1]Sami Harb, BSc, Marie Hudson, MD, MPH, Brett D. Thombs, PhD: Jewish General Hospital and McGill University, Montreal, Quebec, Canada; [2]Sandra Peláez, PhD: Jewish General Hospital and Centre Hospitalier Universitaire Sainte-Justine, Montreal, Quebec, Canada; [3]Marie-Eve Carrier, MSc, Ian Shrier, MD, PhD: Jewish General Hospital, Montreal, Quebec, Canada; [4]Linda Kwakkenbos, PhD: Radboud University, Nijmegen, The Netherlands; [5]Susan J. Bartlett, PhD: McGill University, Montreal, Quebec, Canada; [6]Luc Mouthon, MD, PhD: Hôpitaux de Paris, Université Paris Descartes and Service de Médecine Interne, Hôpital Cohin, Paris, France; [7]Maureen Sauvé: Scleroderma Canada and Scleroderma Society of Ontario, Hamilton, Ontario, Canada; [8]Joep Welling: Federation of European Scleroderma Associations, Brussels, Belgium.

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

Address correspondence to Brett D. Thombs, PhD, Jewish General Hospital, 4333 Cote Ste Catherine Road, Montréal, Québec, H3T 1E4, Canada. Email: brett.thombs@mcgill.ca.

Submitted for publication July 16, 2020; accepted in revised form January 28, 2021.

SIGNIFICANCE & INNOVATIONS

- Based on a survey of 721 people with scleroderma, barriers to physical activity that were most commonly considered important involved compromised hand dexterity or condition (e.g., Raynaud's phenomenon), general symptoms (e.g., fatigue) or localized symptoms (e.g., gastrointestinal problems), and low motivation.
- Barrier-specific physical activity facilitators that were most likely to be used addressed adapting the exercise type or setting, using health behaviors to take care of the body, and strategies to keep warm and protect the skin.
- Generally, participants who experienced the barrier and had tried the linked facilitator were likely to use it, whereas participants who experienced the barrier and had not tried the linked facilitator were not likely to use it.
- Health care providers can use facilitators identified in this study to adapt physical activity options so that people with scleroderma can overcome barriers to physical activity.

rheumatic diseases (8), people with SSc experience a wide range of barriers that may impede engagement. Data from a large international SSc cohort demonstrated that ~50% of patients were physically inactive, and patients who were active rarely engaged in activities other than walking (9). This study, by Azar at al (9), and other studies on physical activity in SSc (10–12) have not addressed barriers or facilitators to being physically active.

For health care providers to advise SSc patients on how to be physically active, they need to be able to identify possible facilitators, or strategies, to overcome specific barriers faced by individual patients. We previously conducted a nominal group technique study to identify barriers to physical activity, along with potential facilitators, experienced by people with SSc (13). That study included only 41 people, which did not allow conclusions to be drawn about the prevalence of barriers and likelihood that people with SSc would use identified facilitators. The aim of the present study was to obtain information on the prevalence of barriers and perceived utility of facilitators to help tailor physical activity recommendations to the specific needs of people with SSc. Specific objectives were to determine the prevalence and importance of different barriers experienced in SSc and the likelihood that people with SSc would use different patient-generated, barrier-specific, and general facilitators to support physical activity.

PATIENTS AND METHODS

The present study was cross-sectional, in which survey results from the Scleroderma Patient-Centered Intervention Network (SPIN) Physical Activity Survey were deterministically linked (using participant usernames [email addresses]) to participant sociodemographic, medical, and patient-reported outcome measure data from the ongoing SPIN Cohort.

Participants and procedures. Eligible SPIN Cohort participants had to be: classified as having SSc according to the 2013 American College of Rheumatology/European Alliance of Associations for Rheumatology criteria (14), ≥18 years of age, fluent in English, French, or Spanish, and able to respond to questionnaires via the internet. Eligible individuals are invited by their attending physician or supervised nurse coordinator to participate in the SPIN Cohort, and written informed consent was obtained. The local SPIN physician or supervised nurse coordinator completed a medical data form that was submitted online to initiate participant registration. After completion of online registration, an automated welcoming email was sent to participants with instructions for activating their SPIN account and completing SPIN Cohort measures online. SPIN Cohort participants completed online outcome measures upon enrollment and subsequently every 3 months.

For the present study, in July 2019 we invited active SPIN Cohort participants to complete a survey, separately from their routine cohort assessments. We sent email invitations to all 1,707 SPIN Cohort participants who had active SPIN accounts and who completed assessments in English or French. We sent follow-up emails 2, 4, and 8 weeks later to those who had not completed the survey. In addition, we advertised the survey through an announcement presented to SPIN Cohort participants when they logged into the SPIN Cohort portal to complete their routine online assessments. To promote participation, we informed participants that 1 survey respondent would be randomly selected to win a trip to the 2020 SSc World Congress in Prague, Czech Republic. The email invitation and announcements provided a link to the survey on the Qualtrics survey platform (15). In Qualtrics, participants entered their SPIN username (email address) in order to access and complete the survey questions. The survey was closed in October 2019. We excluded participants who only partially completed the survey. SPIN Cohort assessment data were obtained from the most recently completed assessments prior to completing the SPIN Physical Activity Survey for participants and prior to the initial survey invitation for nonparticipants, without time restriction.

The SPIN Cohort was approved by the Research Ethics Committee of the Centre intégré universitaire de santé et de services sociaux du Centre-Ouest-de-l'Île-de-Montréal (#MP-05-2013-150) and by the research ethics committees of each participating center. The present study was approved as an amendment to the SPIN Cohort by the Research Ethics Committee of the Centre intégré universitaire de santé et de services sociaux du Centre-Ouest-de-l'Île-de-Montréal.

Measures. *Sociodemographic and medical characteristics.* Medical data were provided by SPIN physicians upon enrollment in the SPIN Cohort, and included time since first non–Raynaud's

phenomenon symptoms, time since SSc diagnosis, SSc subtype, degree of joint contractures for small and large joints, tendon friction rubs status, interstitial lung disease status, pulmonary arterial hypertension status, Raynaud's phenomenon status, digital ulcer status (digital pulp and anywhere else on the finger), and gastrointestinal tract involvement status (esophageal, stomach, and intestinal). For each participant, we calculated the time from when sociodemographic and medical characteristics were obtained at entry into the SPIN Cohort to survey completion.

Physical activity. The SPIN Cohort assessment included the following 2 items: 1) "Compared to other people your age, how would you rate your physical activity during the past year?" (physically inactive, somewhat active, moderately active, quite active, very active); and 2) "Do you exercise at present?" (yes, no). Among participants who reported exercising at present, 2 additional items were administered: 1) "On the average, how many hours per week do you exercise?", and 2) "What type(s) of exercise(s) do you do?" (walking, jogging, aerobics, swimming, other [specify]). For the "other" option, participants could indicate more than 1 type of exercise. All exercises described by participants in the "other" option were classified based on the 2011 Compendium of Physical Activities (16).

Physical function. We used the 4-item Patient-Reported Outcomes Measurement Information System (PROMIS) physical function domain 4a (profile version 2.0) to evaluate self-reported physical activity capability. Each item is scored on a 5-point scale (1–5), where higher scores reflect better physical function over the previous 7 days. The total score is obtained by converting the sum of raw item scores into T scores standardized from the general US population (mean ± SD 50 ± 10). The PROMIS physical function domain 4a (profile version 2.0) has been validated in SSc (17–19).

Functional disability. The Health Assessment Questionnaire disability index (HAQ DI) assesses 8 disability categories over the past 7 days. Each item is rated on a 4-point scale, ranging from 0 (without any difficulty) to 3 (unable to do), where higher scores reflect greater functional disability. The highest score from each category determines the score for that category, and the total score is the mean of the 8 category scores, ranging from 0 (no disability) to 3 (severe disability). The HAQ DI is a valid measure of functional disability in SSc (20).

SPIN Physical Activity Survey. We developed the SPIN Physical Activity Survey to evaluate whether possible physical activity barriers were experienced and, if experienced, their importance and to evaluate possible facilitators for likelihood of use (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24567). An initial list of barriers and facilitators was generated via 9 nominal group technique sessions with 41 people with SSc at patient conferences in Canada, the US, and France (13). Study investigators consolidated overlapping items, reworded unclear items, and excluded vague or unrelated items. Next, the 9-member SPIN Physical Activity Patient Advisory Team (see Appendix A for advisory team members) and SPIN-affiliated health care providers made recommendations to reword, exclude, or add barrier and facilitator items. The item list included 20 barriers classified into 4 categories (21), including health and medical (n = 14); social and personal (n = 4); time, work, and lifestyle (n = 1); and environmental (n = 1). There were 91 barrier-specific facilitators and 12 general facilitators. Patient advisors pilot tested the survey and provided feedback on usability; survey instructions were revised accordingly. The survey was then translated into French using a standard forward–backward translation process (22).

In the survey, to reduce burden, participants were asked to select up to 10 of the 20 total barriers that they had experienced and believed were important for them, initially order selected barriers from most to least important by dragging them into position, and rate each selected barrier on a 4-point Likert scale based on importance to them when thinking about or actually being physically active (not important, somewhat important, important, very important). We next presented participants with all barrier-specific facilitators that corresponded to their selected barriers, and they rated the likelihood that they would use each barrier-specific facilitator to overcome the corresponding barrier (not likely, somewhat likely, likely, very likely) and indicated whether they had previously tried it. Participants similarly rated general facilitators. At the end of the survey, participants were able to provide suggestions for additional barriers and facilitators.

Data analysis. We used descriptive statistics, summarized continuous variables using medians (ranges) and categorical variables using percentages, and listed additional barriers and facilitators provided by participants. To gain further insights, we stratified the analyses related to barriers by whether participants exercised or not and by sex. In addition, because we believed that those who tried a facilitator that helped their physical activity would be likely to use it again, we stratified the analyses based on the likelihood of using facilitators separately by those who had experienced the barrier and previously tried the facilitator in comparison to those who had experienced the barrier but had not tried the facilitator.

We classified barriers using the same 4 categories used to classify them in the nominal group technique study where the list was generated (13). Also, based on consensus among investigators and the SPIN Physical Activity Patient Advisory Team, we applied descriptive labels in the text to similar barriers and facilitators in order to clearly and succinctly summarize results. All analyses were conducted with Microsoft Excel, version 16.16.

RESULTS

Participant characteristics. Of 1,707 invited SPIN Cohort participants, 721 (42%) completed the full SPIN Physical Activity Survey and were included in analyses). A total of 70 participants who partially completed the survey were excluded. The

median age of participants was 59 years (range 22–89 years), ~90% were women, and almost half were employed full- or part-time (Table 1). Median time duration since SSc diagnosis was 10.4 years, and ~40% of participants had diffuse SSc. Approximately one-third of participants were ≥1 SD below the US population mean score on the PROMIS physical function domain 4a (profile version 2.0), and half had at least mild functional impairment (median HAQ DI score 0.6). As shown in Table 2, walking was performed by 47% of participants and conditioning exercises by 26%.

Sociodemographic and medical characteristics of respondents were similar to nonrespondents; the range of differences for categorical variables was 0–7% (Table 1). However, there were some differences in physical activity characteristics between respondents and nonrespondents. There was a 15% difference in the proportion who reported currently exercising (61% of respondents versus 46% of nonrespondents) and differences in the proportion who performed specific types of exercises.

Physical activity barriers. There were 172 participants (24%) who experienced and selected 10 barriers for rating and 549 (76%) who selected fewer than 10. The median number of barriers selected was 7. There were 4 barriers, all health and medical barriers, that were experienced and selected for rating by ≥50% of the 721 total participants, including Raynaud's phenomenon, fatigue, joint stiffness and contractures, and difficulty grasping objects. Of these 4 barriers, fatigue (58%) and Raynaud's phenomenon (57%) were selected for rating and classified as important or very important by ≥50% of total participants. The joint stiffness and contractures barrier was selected and rated as important or very important by 49% of participants, shortness of breath by 38%, gastrointestinal problems by 36%, difficulty grasping objects

Table 1. Participant sociodemographic and medical characteristics*

Variable	SPIN Cohort	
	Respondents (n = 721)	Nonrespondents (n = 986)
Sociodemographic variables		
Age, median (range) years	59 (22–89)	57 (21–91)
Women	640 (89)	865 (88)
White race/ethnicity	603 (85); n = 714†	717 (79); n = 912‡
Education completed, median (range) years§	16 (3–27); n = 708†	15 (0–28); n = 900‡
Employed full- or part-time	323 (46); n = 708†	369 (41); n = 903‡
Married or living as married	455 (64); n = 708†	547 (61); n = 903‡
Geographic region		
North America	429 (60)	584 (59)
Europe	292 (40)	401 (41)
Australia	0 (0)	1 (0)
English survey language	447 (62)	649 (69); n = 935‡
Medical variables		
Time in years since baseline assessment when medical data were recorded, median (range)	3.1 (0.4–5.8)	3.1 (0.4–6.7)
Time in years since first non–Raynaud's phenomenon symptom, median (range)	12.3 (0.4–47.3); n = 666†	11.3 (1.6–58.8); n = 899‡
Time in years since systemic sclerosis diagnosis, median (range)	10.4 (0.4–43.8); n = 697†	9.8 (0.8–58.8); n = 939‡
Diffuse systemic sclerosis subtype	279 (39); n = 713†	409 (42); n = 979‡
Body mass index, median (range)	24.0 (14.7–60.7)	24.6 (13.0–64.4)
Raynaud's phenomenon	695 (98); n = 711†	963 (98); n = 979‡
Digital ulcers (distal pulp)	238 (34); n = 703†	364 (38); n = 970‡
Digital ulcers (anywhere else on the finger)	101 (15); n = 692†	184 (19); n = 944‡
Current or past tendon friction rubs	154 (25); n = 618†	210 (24); n = 865‡
Moderate or severe contractures of small joints	172 (26); n = 673†	253 (27); n = 934‡
Moderate or severe contractures of large joints	79 (12); n = 657†	136 (15); n = 918‡
Any gastrointestinal involvement	621 (87); n = 706†	873 (89); n = 983‡
Interstitial lung disease	228 (33); n = 692†	346 (36); n = 974‡
Pulmonary arterial hypertension	45 (7); n = 691†	80 (9); n = 937‡
PROMIS physical function domain score, median (range)	43.4 (22.9–56.9); n = 705†	41.8 (22.9–56.9); n = 876‡
Total HAQ DI score, median (range)	0.6 (0.0–3.0); n = 701†	0.6 (0.0–3.0); n = 862‡

* Values are the number (% of data recorded) of participants unless indicated otherwise. HAQ DI = Health Assessment Questionnaire disability index; PROMIS = Patient Reported Outcomes Measurement Information System (profile version 2.0).
† No. of Scleroderma Patient-Centered Intervention Network (SPIN) respondents due to missing data.
‡ No. of SPIN nonrespondents due to missing data.
§ Years of education completed beginning from elementary/primary school and including all levels of formal education.

Table 2. Participant physical activity characteristics (n = 721)*

Variable	SPIN Cohort respondents	SPIN Cohort nonrespondents
Participants' perception of their physical activity level in the past year compared to other people their age		
Physically inactive	85 (12)	155 (17)
Somewhat active	199 (28)	316 (34)†
Moderately active	233 (33)	270 (29)†
Quite active	148 (21)	115 (12)†
Very active	50 (7)	66 (7)†
Currently exercise	433 (61)	421 (46)‡
Hours per week of exercise, median (range)	4 (1–15)§	4 (1–15)¶
Types of exercises performed		
Walking	333 (47)	328 (35)
Jogging	24 (3)	25 (3)
Aerobics	75 (11)	64 (7)
Swimming	59 (8)	41 (4)
Other	275 (39)	209 (22)
"Other" exercises (selected examples)#		
Bicycling (biking, cycling, spinning)	42 (6)	29 (3)
Conditioning (elliptical, gym, pilates, stretching, tai chi, weight-lifting, yoga)	183 (26)	152 (16)
Lawn and garden (gardening, landscaping, yard work)	16 (2)	9 (1)
Sports (badminton, racquetball, bowling, golf)	25 (4)	26 (3)
Walking (Nordic walking)	13 (2)	9 (1)
Water activities (aquatic classes, kayaking, pool exercises)	14 (2)	7 (1)
Other categories**	52 (7)	12 (1)

* Values are the number (% of data recorded) of participants unless indicated otherwise. N = 715 for Scleroderma Patient-Centered Intervention Network (SPIN) Cohort respondents (due to missing data) and n = 933 for SPIN Cohort nonrespondents.
† N = 922 (due to missing data).
‡ N = 921 (due to missing data).
§ Participants who reported currently exercising and their average hours per week of exercise (n = 433).
¶ Participants who reported currently exercising and their average hours per week of exercise (n = 418).
# Participants could indicate >1 exercise, and each exercise was classified into 1 category.
** Other categories of activities performed by ≤2% of participants were dancing, fishing and hunting, home activity, miscellaneous, music playing, and winter activities.

by 33%, pain by 33%, muscle weakness and difficulty with mobility by 29%, and lack of motivation and difficulty committing to exercise by 26%. A summary of the initial sorted rankings of barriers by importance, rather than by ratings, is available (see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24567).

The distribution of barrier ratings separately for participants who did (n = 433) and did not (n = 282) report presently engaging in exercise is shown (see Supplementary Appendices C and D, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24567). The importance of barriers tended to be rated higher by those who did not exercise. The 3 largest differences in the percentage of participants rating barriers as important or very important were for lack of motivation (21% difference), fatigue (14% difference), and difficulty grasping objects (11% difference).

The distribution of barrier ratings for male (n = 81) and female (n = 640) participants is shown (see Supplementary Appendices E and F, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24567). Overall, the distributions of barrier ratings for male and female participants were generally similar. The 2 barriers with the largest differences were gastrointestinal problems (12%) and Raynaud's phenomenon (10%), which both had a higher percentage of female participants rating the barrier as important or very important (Figure 1).

Physical activity facilitators. Overall, of 103 facilitators rated by participants who had experienced the linked barrier, 23 (22%) were rated as likely or very likely to use by ≥75% of participants and an additional 58 (56%) facilitators were rated the same by ≥50% of participants. The full list of barriers, their facilitators, and participant ratings is available (see Supplementary Appendix G at http://onlinelibrary.wiley.com/doi/10.1002/acr.24567); it is also accessible online as an interactive spreadsheet (https://osf.io/2mxj5/) that facilitates sorting and identifying facilitators for different barriers. Table 3 shows the 12 health and medical barriers that were experienced and selected for rating by ≥25% of total participants and a selection of corresponding barrier-specific facilitators that were commonly rated as likely or very likely to use among those who tried them. The most common facilitators overall and among those presented in Table 3 involved strategies for adapting exercise type, conduct, or setting (e.g., using controlled, slow movement), changing health

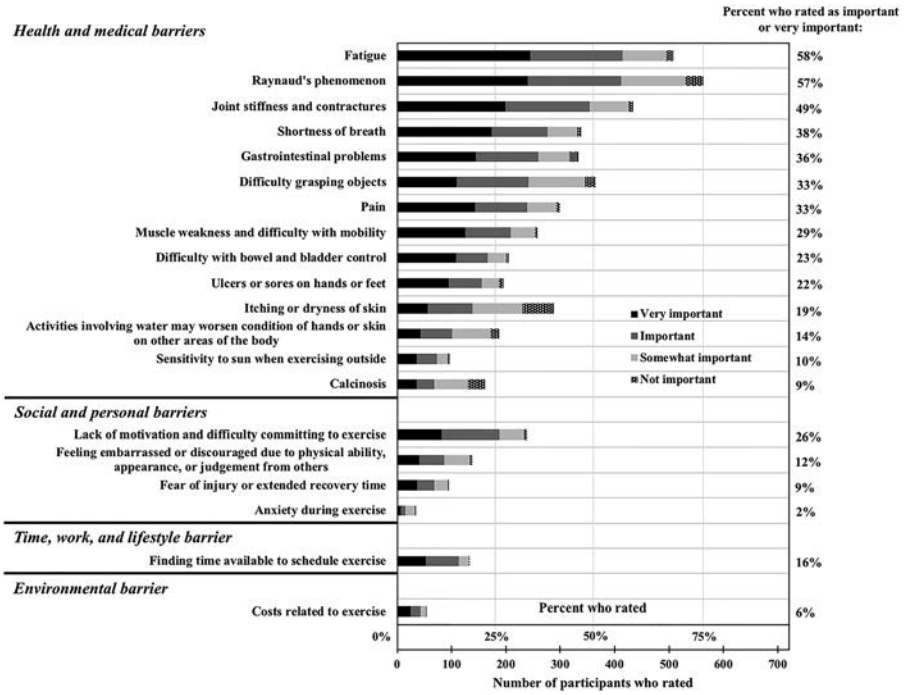


Figure 1. Distribution of ratings for barriers. Participants (n = 721 total participants) only rated up to a maximum of 10 barriers that they experienced and selected for rating. Using a 4-point Likert scale, participants rated each of their selected barriers based on how important it is to them personally when thinking about or actually being physically active (not important, somewhat important, important, very important). Because 172 participants rated the maximum of 10 barriers, it is possible that they experienced other barriers as well. Percentages refer to the percent of 721 participants who rated the adjacent barrier as important or very important.

behaviors to take care of the body (e.g., stretching), keeping warm (e.g., wearing gloves), and protecting the skin (e.g., covering ulcers). Additional barrier and facilitator suggestions to those presented in our survey, which were provided by survey respondents and were substantively different from those included in the survey, are shown (see Supplementary Appendix H, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24567).

The majority of the facilitators (62 of 103 [60%]) had been tried by ≥50% of participants who rated them. Among those participants who tried facilitators, 103 of 103 facilitators were rated by ≥50% as likely or very likely to use, and 65 of 103 facilitators were rated by ≥80% of participants as likely or very likely to use. In contrast, only 12 of 103 facilitators were rated as likely or very likely to use by ≥50% of participants who had not tried them previously.

DISCUSSION

The main results of our study include the prevalence of barriers to physical activity among more than 700 people with SSc, along with their ratings of the importance of each barrier and of the likelihood that they would use corresponding and more general facilitators of physical activity. The most common barriers to physical activity were Raynaud's phenomenon and fatigue, followed by compromised hand dexterity and challenges related to respiratory, gastrointestinal, and skin pathologies. Among the 103 barrier-specific and general facilitators in the survey, for participants who had tried each of them, at least 50% of participants said they would be likely or very likely to use them to facilitate physical activity. Health care providers can use our interactive Excel spreadsheet (https://osf.io/2mxj5/) to review physical activity barriers and identify patient-generated facilitators to address these barriers and support physical activity among individuals with SSc.

Although this was the first study to evaluate patient-generated physical activity barriers and possible facilitators to overcome such barriers in a large SSc sample, results are consistent with findings from previous studies. A previous study with the SPIN Cohort (n = 752) found that presently reported exercise was associated with fatigue, pain, degree of skin thickening, and functional disability (9), all of which were identified by participants in the present study as barriers. Facilitators rated widely as likely to be used for such barriers were often related to adapting the exercise form (e.g., use controlled, slow movements for pain), conduct (e.g., take rest breaks for fatigue, pain, and muscle weakness and difficulty with mobility), and equipment (e.g., use wrist weights for difficulty grasping objects). Consistent with the shortness of breath barrier, lung involvement (23) and pulmonary hypertension (24) have been found to be associated with reduced aerobic capacity in 2 small exercise studies (n = 46 and n = 18 participants). Two of our barrier-specific facilitators ("take rest breaks

Table 3. The 12 medical barriers experienced and selected for rating by ≥25% of participants and a subset of corresponding novel and common facilitators (n = 721 total participants)*

Barriers	Participants who experienced and selected barrier for rating†	Facilitators	Participants who tried facilitator and "likely" or "very likely" to use it‡
Raynaud's phenomenon	78 (564)	Dress to stay warm (keep your core warm and cover areas of the body that become cold – e.g., wear a warm hat, gloves, or mittens)	93 (501/539)
		Exercise in an area with a temperature that is comfortable for you	90 (451/502)
		Wear heated or nonheated warm gloves or mittens and socks	92 (452/494)
		Insert warmers (i.e., liners, or electric or chemical warmers) in gloves or mittens or socks	86 (334/387)
Fatigue	71 (508)	Take rest breaks while exercising (e.g., between activities)	83 (333/403)
		Break exercise into several short periods (e.g., three 10-minute walks) rather than a single long period (e.g., one 30-minute walk)	82 (235/286)
		Get enough sleep and plan to take a nap during the day	80 (273/342)
Joint stiffness and contractures	60 (434)	Do daily gentle stretching and exercises that move your joints through their maximum range of motion	82 (256/312)
		Use controlled, slow movements that are comfortable for you	85 (263/309)
Difficulty grasping objects	51 (365)	Use adapted exercise equipment (e.g., weights with a larger handle or wrist weights)	82 (108/132)
Shortness of breath	47 (338)	Lower the intensity of the exercise to not experience shortness of breath	86 (251/291)
Gastrointestinal problems	46 (334)	If you have acid reflux, modify exercise positions to keep your body upright (e.g., do push-ups against the wall instead of push-ups against the ground)	89 (148/166)
Pain	42 (300)	Modify exercise so it does not cause pain (e.g., use lighter weights or walk slower)	87 (223/256)
Itching or dryness of skin	40 (289)	Moisturize regularly or as needed (e.g., use lotion or wear moisturizing gloves or socks)	89 (223/251)
Muscle weakness and difficulty with mobility	36 (258)	If you have difficulty with balance, place a hand against an immovable object (e.g., wall or pole) for support or exercise while sitting on an immovable chair or seat	88 (151/172)
		If you have difficulty with balance, use assistive devices (e.g., hiking poles)	81 (77/95)
Difficulty with bowel and bladder control	28 (205)	Wear a pad or underwear designed for bowel and bladder control issues	90 (132/146)
Ulcers or sores on hands or feet	27 (195)	Apply nonadhesive bandages to cover and protect ulcers or sores	92 (140/153)
		Wear appropriate clothing to cover and protect ulcers or sores (e.g., gloves or mittens)	90 (148/165)
		If you have foot ulcers or sores, put pads in shoes or wear specialized soles or shoes (e.g., open-toe shoes)	87 (65/75)
Activities involving water may worsen condition of hands or skin on other areas of the body	26 (188)	Wear a wet suit, gloves, or socks designed for water exercises to stay warm	72 (33/46)

* Participants rated on a 4-point Likert scale the likelihood that they would use each barrier-specific facilitator to overcome the corresponding barrier to be physically active (not likely, somewhat likely, likely, very likely). See interactive Excel file (https://osf.io/2mxj5/) for the full list.
† Values are the % (number) of participants who experienced and selected the barrier for rating.
‡ Values are the % (number/total number) of participants who rated the facilitator as "likely" or "very likely" to use among those who experienced the barrier and had tried the facilitator.

while exercising" and "lower the intensity of exercise to not experience shortness of breath") directly address reduced aerobic capacity.

Barriers outside the medical category were generally less common than medical barriers. The most common nonmedical barrier was "lack of motivation," which was rated important or very important by 26% of total participants, followed by "finding time available to schedule exercise" (16%) and "feeling embarrassed or discouraged due to physical ability, appearance, or judgement from others" (12%). While motivation- and time-related barriers have been reported as important barriers to physical activity in the general population (25,26), the barrier about feeling embarrassed or discouraged seems to more directly reflect the unique experiences of people with SSc, particularly psychosocial consequences due to concerns about visible changes to one's appearance (27).

Subgroup analyses revealed that a substantially larger proportion of inactive participants had rated 2 health and medical barriers (fatigue, difficulty grasping objects) and 1 social and personal barrier (lack of motivation) as important or very important

compared to active participants. These 3 barriers could be targeted when developing general interventions to promote physical activity in SSc patients.

All facilitators were rated by at least half of participants who tried the facilitators as likely or very likely to use. Some facilitators commonly rated as likely to be used are consistent with widely recommended strategies, such as for warming in Raynaud's phenomenon (28), and identifying enjoyable activities for people who have difficulty with motivation or exercise adherence (29). On the other hand, there were novel barrier-specific facilitators widely perceived as likely to be used that, to our knowledge, have not been reported in the literature but could be helpful for health care providers promoting physical activity to individuals with SSc. Many novel facilitators addressed adapting the exercise (either by adapting the exercise conduct, type, or setting), including use of adapted exercise equipment (barriers of difficulty grasping objects and joint stiffness and contractures) and "participate in gentle exercise classes that may be intended for new exercisers or people with limitations for exercising" (barrier of fear of injury or extended recovery time). Importantly, individuals with SSc should consult a qualified clinician about how to exercise safely.

In general, participants who tried facilitators rated them favorably, as likely or very likely to use, in comparison to those who had not tried them. This finding suggests that some challenges may exist when proposing new facilitators to SSc patients. Communication skills and strategy may be very important. A widely used intervention to support physical activity in the general population, Active Living Every Day (30), uses a social modeling component when exposing individuals to new facilitators. This intervention involves sharing the personal experiences of people who describe how they overcame specific barriers to leading a more active lifestyle. We expect that such social modeling would be a potentially effective strategy to promote physical activity in SSc, especially for those patients who had not tried a proposed facilitator.

Our findings suggest barriers that could be targeted to facilitate physical activity. Strategies to treat fatigue in rheumatoid arthritis include exercise, cognitive behavioral therapy, and self-management programs (31). SPIN is currently testing a SSc self-management program (SPIN-SELF) (32). Strategies to reduce the effects of Raynaud's phenomenon include keeping a diary and identifying activities that trigger attacks, keeping the body and hands warm (e.g., layered clothing, gloves), and avoiding smoking (33). Limitations in mobility, which are common in the hands (34), may be addressed through hand stretches and exercises, and SPIN has developed the SPIN-HAND program, which is available online, free-of-charge (35). Social support is a strong predictor of exercise intention and stage of behavior change for exercise (36). Many people with SSc attend support groups (37), and the SPIN-SELF program also contains a group component.

There are limitations to take into account in interpreting results of the present study. First, the results may not be generalizable to people who do not speak English or French, reside outside of North America and Europe, or do not have access to a device with internet. Second, a higher proportion of respondents (61%) reported currently exercising in comparison to SPIN Cohort nonrespondents (46%). Third, participants were presented with 20 possible barriers, but in order to reduce respondent burden, we only allowed them to select up to 10 barriers that they had experienced. Almost 25% of participants selected 10 barriers and might have experienced and selected additional barriers, if that had been permitted, although these would have been of lesser importance to the participant than the ones they selected. Fourth, although participants were asked to select the barriers for rating that they experienced and felt were important, some participants rated at least 1 of their selections as "not important." Fifth, although participants rated the importance of barriers and likelihood of using facilitators, the survey did not elicit explanations for why they rated barriers and facilitators as they did. Such explanations might help to fine tune guidance to better address physical activity difficulties experienced by individuals with SSc. Sixth, although our measure of physical activity behavior was modeled after part of an existing validated questionnaire (38,39), we did not administer a validated measure of physical activity behavior, which would have allowed us to better characterize participants and to compare their physical activity behavior with other samples. This was an effort to reduce respondent burden because there were constraints on the number of items that we were able to add to a preexisting cohort assessment. One area of future research could include comparison of general levels of physical activity behavior in SSc patients to the published norms in the general population.

In summary, medical-related barriers to activity were most commonly experienced and considered important; Raynaud's phenomenon and fatigue were the most commonly experienced among them. Facilitators widely considered likely to be used addressed adapting exercise type or setting, using health behaviors to take care of the body, and using clothing or materials to protect the skin or to keep warm. Participants who had tried facilitators were generally more likely to use them again compared to participants who had never tried them. Our online interactive Excel file (https://osf.io/2mxj5/) allows health care providers to easily identify relevant facilitators for common barriers to physical activity experienced by individuals with SSc.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Thombs 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. Harb, Peláez, Carrier, Kwakkenbos, Bartlett, Hudson, Mouthon, Sauvé, Welling, Shrier, Thombs.

Acquisition of data. Harb, Carrier, Kwakkenbos, Thombs.

Analysis and interpretation of data. Harb, Peláez, Shrier, Thombs.

REFERENCES

1. Seibold JR. Scleroderma. In: Harris E, editor. Kelley's textbooks of rheumatology. 7th ed. Philadelphia: Elsevier; 2005. p. 1279–80.
2. Wigley FM. Clinical features of systemic sclerosis. In: Hochberg MC, editor. Rheumatology. 3rd ed. Philadelphia: Mosby; 2003. p. 1463–80.
3. Hudson M, Thombs BD, Steele R, Panopalis P, Newton E, Baron M, et al, for the Canadian Scleroderma Research Group. Health-related quality of life in systemic sclerosis: a systematic review. Arthritis Rheum 2009;61:1112–20.
4. Mayes MD. Systemic sclerosis: clinical features. In: Klippel JH, Stone JH, White PH, editors. Primer on the rheumatic diseases. New York: Springer Science & Business Media; 2008. p. 343–50.
5. Gelber AC, Manno RL, Shah AA, Woods A, Le EN, Boin F, et al. Race and association with disease manifestations and mortality in scleroderma: a 20-year experience at the Johns Hopkins Scleroderma Center and review of the literature. Medicine (Baltimore) 2013;92: 191–205.
6. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity: a systematic review of longitudinal studies. BMC Public Health 2013;13:813.
7. Warburton DE, Charlesworth S, Ivey A, Nettlefold L, Bredin SS. A systematic review of the evidence for Canada's physical activity guidelines for adults. Int J Behav Nutr Phys Act 2010;7:39.
8. Perandini LA, de Sá-Pinto AL, Roschel H, Benatti FB, Lima FR, Bonfá E, et al. Exercise as a therapeutic tool to counteract inflammation and clinical symptoms in autoimmune rheumatic diseases. Autoimmun Rev 2012;12:218–24.
9. Azar M, Rice DB, Kwakkenbos L, Carrier M, Shrier I, Bartlett SJ, et al. Exercise habits and factors associated with exercise in systemic sclerosis: a Scleroderma Patient-centered Intervention Network (SPIN) cohort study. Disabil Rehabil 2018;40:1997–2003.
10. Battaglia S, Bellia M, Serafino-Agrusa L, Giardina A, Messina M, Cannizzaro F, et al. Physical capacity in performing daily activities is reduced in scleroderma patients with early lung involvement. Clin Respir J 2017;11:36–42.
11. Pettersson H, Åkerström A, Nordin A, Svenungsson E, Alexanderson H, Boström C. Self-reported physical capacity and activity in patients with systemic sclerosis and matched controls. Scand J Rheumatol 2017;46:490–5.
12. Liem S, Meessen J, Wolterbeek R, Marsan NA, Ninaber M, Vlieland TV, et al. Physical activity in patients with systemic sclerosis. Rheumatol Int 2018;38:443–53.
13. Harb S, Cumin J, Rice DB, Peláez S, Hudson M, Bartlett SJ, et al. Identifying barriers and facilitators to physical activity for people with scleroderma: a nominal group technique study. Disabil Rehabil 2020:1–8.
14. 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.
15. Qualtrics. Provo U. Qualtrics survey platform. 2002.
16. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 compendium of physical activities: a second update of codes and MET values. Med Sci Sports Exerc 2011; 43:1575–81.
17. Kwakkenbos L, Thombs BD, Khanna D, Carrier M, Baron M, Furst DE, et al. Performance of the patient-reported outcomes measurement information system-29 in scleroderma: a Scleroderma Patient-centered Intervention Network Cohort study. Rheumatology (Oxford) 2017;56:1302–11.
18. Hinchcliff M, Beaumont JL, Thavarajah K, Varga J, Chung A, Podlusky S, et al. Validity of two new patient-reported outcome measures in systemic sclerosis: patient-reported outcomes measurement information system 29-item health profile and functional assessment of chronic illness therapy–dyspnea short form. Arthritis Care Res (Hoboken) 2011;63:1620–8.
19. Hinchcliff ME, Beaumont JL, Carns MA, Podlusky S, Thavarajah K, Varga J, et al. Longitudinal evaluation of PROMIS-29 and FACIT-dyspnea short forms in systemic sclerosis. J Rheumatol 2015;42: 64–72.
20. Clements PJ, Wong WK, Hurwitz EL, Furst DE, Mayes M, White B, et al. Correlates of the disability index of the health assessment questionnaire: a measure of functional impairment in systemic sclerosis. Arthritis Rheum 1999;42:2372–80.
21. Lascar N, Kennedy A, Hancock B, Jenkins D, Andrews RC, Greenfield S, et al. Attitudes and barriers to exercise in adults with type 1 diabetes (T1DM) and how best to address them: a qualitative study. PLoS One 2014;9:e108019.
22. Process of translation and adaptation of instruments. 2020. URL: https://www.who.int/substance_abuse/research_tools/translation/en/.
23. Cuomo G, Santoriello C, Polverino F, Ruocco L, Valentini G, Polverino M. Impaired exercise performance in systemic sclerosis and its clinical correlations. Scand J Rheumatol 2010;39:330–5.
24. Morelli S, Ferrante L, Sgreccia A, Eleuteri ML, Perrone C, De Marzio P, et al. Pulmonary hypertension is associated with impaired exercise performance in patients with systemic sclerosis. Scand J Rheumatol 2000;29:236–42.
25. Salmon J, Owen N, Crawford D, Bauman A, Sallis JF. Physical activity and sedentary behavior: a population-based study of barriers, enjoyment, and preference. Health Psychol 2003;22:178.
26. Booth ML, Bauman A, Owen N, Gore CJ. Physical activity preferences, preferred sources of assistance, and perceived barriers to increased activity among physically inactive Australians. Prev Med 1997;26:131–7.
27. Jewett LR, Hudson M, Malcarne VL, Baron M, Thombs BD, Canadian Scleroderma Research Group. Sociodemographic and disease correlates of body image distress among patients with systemic sclerosis. PloS One 2012;7:e33281.
28. Wigley FM, Flavahan NA. Raynaud's phenomenon. N Engl J Med 2016;375:556–65.
29. Richard M, Christina MF, Deborah LS, Rubio N, Kennon MS. Intrinsic motivation and exercise adherence. Int J Sport Psychol 1997;28: 335–54.
30. Blaire SN, Dunn AL, Marcus BH, Carpenter RA, Jaret PE. Active Living Every Day, 3rd edition. Human Kinetics; 2021.
31. Pope JE. Management of fatigue in rheumatoid arthritis. RMD Open 2020;6:e001084.
32. Carrier ME, Kwakkenbos L, Nielson WR, Fedoruk C, Nielsen K, Milette K, et al. The Scleroderma Patient-centered Intervention Network self-management program: protocol for a randomized feasibility trial. JMIR Res Protoc 2020;9:e16799.
33. Kwakkenbos L, Thombs BD. Non-drug approaches to treating Raynaud's phenomenon. In: Raynaud's Phenomenon. New York: Springer; 2015. p. 299–313.
34. Bassel M, Hudson M, Taillefer SS, Schieir O, Baron M, Thombs BD. Frequency and impact of symptoms experienced by patients with systemic sclerosis: results from a Canadian national survey. Rheumatology 2011;50:762–7.
35. Kwakkenbos L, Carrier ME, Boutron I, Welling J, Sauvé M, van den Ende CH, et al. Randomized feasibility trial of the Scleroderma Patient-centered Intervention Network Hand Exercise Program (SPIN-HAND). J Rheumatol 2019;46:816.
36. Courneya KS, Plotnikoff RC, Hotz SB, Birkett NJ. Social support and the theory of planned behavior in the exercise domain. Am J Health Behav 2000;24:300–8.

37. Gumuchian ST, Delisle VC, Kwakkenbos L, Pépin M, Carrier ME, Malcarne VL, et al. Reasons for attending support groups and organizational preferences: the European scleroderma support group members survey. Disabil Rehabil 2019;41:974–82.
38. Godin G. The Godin-Shephard leisure-time physical activity questionnaire. Health & Fitness Journal of Canada 2011;4:18–22.
39. Amireault S, Godin G. The Godin-Shephard leisure-time physical activity questionnaire: validity evidence supporting its use for classifying healthy adults into active and insufficiently active categories. Percept Mot Skills 2015;120:604–22.

APPENDIX A: SPIN PHYSICAL ACTIVITY ENHANCEMENT PATIENT ADVISORY TEAM AND SPIN INVESTIGATORS

Members of the SPIN Physical Activity Enhancement Patient Advisory Team and SPIN Investigators are as follows: Lindsay Cronin (Southwestern Pennsylvania Scleroderma Support Group, Pittsburgh, PA), Catherine Fortuné (Ottawa Scleroderma Support Group, Ottawa, Canada), Amy Gietzen (Scleroderma Foundation, Danvers, Massachusetts), Geneviève Guillot (Sclérodermie Québec, Longueuil, Canada), Shirley Haslam (Scleroderma Society of Ontario, Hamilton, Canada), Karen Nielsen (Scleroderma Society of Ontario, Hamilton, Canada), Michelle Richard (Scleroderma Atlantic, Halifax, Canada), Ken Rozee (Scleroderma Atlantic, Halifax, Canada), Joep Welling (Federation of European Scleroderma Associations, Brussels, Belgium), Murray Baron (McGill University, Montreal, Quebec, Canada), Daniel E. Furst (Division of Rheumatology, Geffen School of Medicine, University of California, Los Angeles, CA) Karen Gottesman (Scleroderma Foundation, Los Angeles, CA), Vanessa Malcarne (San Diego State University, San Diego, CA), Maureen D. Mayes (University of Texas McGovern School of Medicine, Houston, TX), Warren R. Nielson (St. Joseph's Health Care, London, Ontario, Canada), Robert Riggs (Scleroderma Foundation, Danvers, Massachusetts), Fredrick Wigley (Johns Hopkins University School of Medicine, Baltimore, MD), Shervin Assassi (University of Texas McGovern School of Medicine, Houston, TX), Andrea Benedetti (McGill University, Montreal, Quebec, Canada), Isabelle Boutron (Université Paris Descartes, and Assistance Publique - Hôpitaux de Paris, Paris, France), Ghassan El-Baalbaki (Université du Québec à Montréal, Montreal, Quebec, Canada), Carolyn Ells (McGill University, Montreal, Quebec, Canada), Cornelia van den Ende (Sint Maartenskliniek, Nijmegen, The Netherlands), Kim Fligelstone (Scleroderma & Raynaud's UK, London, UK), Tracy Frech (University of Utah, Salt Lake City, UT), Dominique Godard (Association des Sclérodermiques de France, Sorel-Moussel, France), Daphna Harel (New York University, New York, NY), Monique Hinchcliff (Yale School of Medicine, New Haven, CT), Sindhu R. Johnson (Toronto Scleroderma Program, Mount Sinai Hospital, Toronto Western Hospital, and University of Toronto, Toronto, Ontario, Canada), Maggie Larche (McMaster University, Hamilton, Ontario, Canada), Catarina Leite (University of Minho, Braga, Portugal), Christelle Nguyen (Université Paris Descartes, and Assistance Publique - Hôpitaux de Paris, Paris, France), Elisabet Pérez (Asociación Española de Esclerodermia, Madrid, Spain), Janet Pope (University of Western Ontario, London, Ontario, Canada), François Rannou (Université Paris Descartes, and Assistance Publique - Hôpitaux de Paris, Paris, France), Tatiana Sofia Rodriguez Reyna (Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico), Anne A. Schouffoer (Leiden University Medical Center, Leiden, The Netherlands), Maria E. Suarez-Almazor (University of Texas MD Anderson Cancer Center, Houston, TX), Christian Agard (Centre Hospitalier Universitaire - Hôtel-Dieu de Nantes, Nantes, France), Nassim Ait Abdallah (Assistance Publique - Hôpitaux de Paris, Hôpital St-Louis, Paris, France), Alexandra Albert (CHU de Québec - Université Laval, Quebec, Quebec, Canada), Marc André (Centre Hospitalier Universitaire Gabriel-Montpied, Clermont-Ferrand, France), Guylaine Arsenault (Université de Sherbrooke, Sherbrooke, Quebec, Canada), Elana J. Bernstein (Columbia University, New York, NY), Sabine Berthier (Centre Hospitalier Universitaire Dijon Bourgogne, Dijon, France), Lyne Bissonnette (Université de Sherbrooke, Sherbrooke, Quebec, Canada), Gilles Boire (Université de Sherbrooke, Sherbrooke, Quebec, Canada), Alessandra Bruns (Université de Sherbrooke, Sherbrooke, Quebec, Canada), Patricia Carreira (Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain), Marion Casadevall (Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France), Benjamin Chaigne (Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France), Pascal Cohen (Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France), Chase Correia (Northwestern University, Chicago, Illinois), Pierre Dagenais (Université de Sherbrooke, Sherbrooke, Quebec, Canada), Christopher Denton (Royal Free London Hospital, London, UK), Robyn Domsic (University of Pittsburgh, Pittsburgh, PA), Sandrine Dubois (Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France), James V. Dunne (St. Paul's Hospital and University of British Columbia, Vancouver, British Columbia, Canada), Bertrand Dunogue (Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France), Regina Fare (Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain), Dominique Farge-Bancel (Assistance Publique - Hôpitaux de Paris, Hôpital St-Louis, Paris, France), Paul R. Fortin (CHU de Québec - Université Laval, Quebec, Quebec, Canada), Anna Gill (Royal Free London Hospital, London, UK), Jessica Gordon (Hospital for Special Surgery, New York, NY), Brigitte Granel-Rey (Aix Marseille Université, and Assistance Publique - Hôpitaux de Marseille, Hôpital Nord, Marseille, France), Genevieve Gyger (Jewish General Hospital and McGill University, Montreal, Quebec, Canada), Eric Hachulla (Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France), Pierre-Yves Hatron (Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France), Ariane L. Herrick (University of Manchester, Salford Royal NHS Foundation Trust, Manchester, UK), Sabrina Hoa (Centre hospitalier de l'Université de Montréal – CHUM, Montreal, Quebec, Canada), Alena Ikic (Université Laval, Quebec, Quebec, Canada), Niall Jones (University of Alberta, Edmonton, Alberta, Canada), Artur Jose de B. Fernandes (Université de Sherbrooke, Sherbrooke, Quebec, Canada), Suzanne Kafaja (University of California, Los Angeles, CA), Nader Khalidi (McMaster University, Hamilton, Ontario, Canada), Marc Lambert (Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France), David Launay (Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France), Patrick Liang (Université de Sherbrooke, Sherbrooke, Quebec, Canada), Hélène Maillard (Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France), Nancy Maltez (University of Ottawa, Ottawa, Ontario, Canada), Joanne Manning (Salford Royal NHS Foundation Trust, Salford, UK), Isabelle Marie (CHU Rouen, Hôpital de Bois-Guillaume, Rouen, France), Maria Martin (Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain), Thierry Martin (Les Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France), Ariel Masetto (Université de Sherbrooke, Sherbrooke, Quebec, Canada), François Maurier (Hôpitaux Privés de Metz, Hôpital Belle-Isle, Metz, France), Arsene Mekinian (Assistance Publique - Hôpitaux de Paris, Hôpital St-Antoine, Paris, France), Sheila Melchor (Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain), Mandana Nikpour (St Vincent's Hospital and University of Melbourne, Melbourne, Victoria, Australia), Louis Olagne (Centre Hospitalier Universitaire Gabriel-Montpied, Clermont-Ferrand, France), Vincent Poindron (Les Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France), Susanna Proudman (Royal Adelaide Hospital and University of Adelaide, Adelaide, South Australia, Australia), Alexis Régent (Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France), Sébastien Rivière (Assistance Publique - Hôpitaux de Paris, Hôpital St-Antoine, Paris, France), David Robinson (University of Manitoba, Winnipeg, Manitoba, Canada), Esther Rodriguez (Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain), Sophie Roux (Université de Sherbrooke, Sherbrooke, Quebec, Canada), Perrine Smets (Centre Hospitalier Universitaire Gabriel-Montpied, Clermont-Ferrand, France), Vincent Sobanski (Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France), Robert Spiera (Hospital for Special Surgery, New York, NY), Virginia Steen (Georgetown University, Washington, DC), Wendy Stevens (St Vincent's Hospital and University of Melbourne,

Melbourne, Victoria, Australia), Evelyn Sutton (Dalhousie University, Halifax, Nova Scotia, Canada), Benjamin Terrier (Assistance Publique -Hôpitaux de Paris, Hôpital Cochin, Paris, France), Carter Thorne (Southlake Regional Health Centre, Newmarket, Ontario, Canada), John Varga (Northwestern University, Chicago, Illinois), Pearce Wilcox (St. Paul's Hospital and University of British Columbia, Vancouver, British Columbia, Canada), Angelica Bourgeault (Jewish General Hospital, Montreal, Quebec, Canada), Mara Cañedo-Ayala (Jewish General Hospital, Montreal, Quebec, Canada), Andrea Carboni-Jiménez (Jewish General Hospital, Montreal, Quebec, Canada), Maria Gagarine (Jewish General Hospital, Montreal, Quebec, Canada), Richard S. Henry (Jewish General Hospital, Montreal, Quebec, Canada), Nora Østbø (Jewish General Hospital, Montreal, Quebec, Canada), Lydia Tao (Jewish General Hospital, Montreal, Quebec, Canada).

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1311–1320
DOI 10.1002/acr.24551

American College of Rheumatology
Empowering Rheumatology Professionals

Comprehensive Assessment of Quality of Life, Functioning, and Mental Health in Children With Juvenile Idiopathic Arthritis and Noninfectious Uveitis

Joseph McDonald,[1] 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,[9] Najima Mwase,[1] Sampath Prahalad,[10] Jessica Shantha,[8] Erin Stahl,[2] Virginia Miraldi Utz,[1] A. Adrienne Walker,[11] Steven Yeh,[8] and Sheila T. Angeles-Han[12]

Objective. Pediatric uveitis can lead to sight-threatening complications and can impact quality of life (QoL) and functioning. We aimed to examine health-related QoL, mental health, physical disability, vision-related functioning (VRF), and vision-related QoL in children with juvenile idiopathic arthritis (JIA), JIA-associated uveitis (JIA-U), and other noninfectious uveitis. We hypothesized that there will be differences based on the presence of eye disease.

Methods. A multicenter cross-sectional study was conducted at four sites. Patients with JIA, JIA-U, or noninfectious uveitis were enrolled. Patients and parents completed the Pediatric Quality of Life Inventory (PedsQL; health-related QoL), the Revised Childhood Anxiety and Depression Scale (RCADS; anxiety/depression), the Childhood Health Assessment Questionnaire (C-HAQ; physical disability), and the Effects of Youngsters' Eyesight on Quality of Life (EYE-Q) (VRF/vision-related QoL). Clinical characteristics and patient-reported outcome measures were compared by diagnosis.

Results. Of 549 patients, 332 had JIA, 124 had JIA-U, and 93 had other uveitis diagnoses. Children with JIA-U had worse EYE-Q scores compared to those with JIA only. In children with uveitis, those with anterior uveitis (JIA-U and uveitis only) had less ocular complications, better EYE-Q scores, and worse C-HAQ and PedsQL physical summary scores compared to those with nonanterior disease. In children with anterior uveitis, those with JIA-U had worse PedsQL physical summary and C-HAQ scores than anterior uveitis only. Further, EYE-Q scores were worse in children with bilateral uveitis and more visual impairment. There were no differences in RCADS scores among groups.

Conclusion. We provide a comprehensive outcome assessment of children with JIA, JIA-U, and other uveitis diagnoses. Differences in QoL and function were noted based on underlying disease. Our results support the addition of a vision-specific measure to better understand the impact of uveitis.

INTRODUCTION

Pediatric uveitis is an inflammatory ocular condition that can lead to sight-threatening complications and blindness. Most forms of uveitis in the US are not attributable to infection (1). Juvenile idiopathic arthritis (JIA) is the most common systemic disease association (2,3). Uveitis without concomitant systemic disease is known as idiopathic uveitis and has a similar incidence as JIA-associated uveitis (JIA-U) (4,5). Uveitis can be found in

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

This study and Dr. Angeles-Han's work were supported by the NIH/National Eye Institute (grants K-23-EY-021760 and R01EY030521), the Rheumatology Research Foundation, the Childhood Arthritis and Rheumatology Research Alliance, the Arthritis Foundation, the Weill Cornell Medical College Clinical and Translational Science Center (grant UL1-RR-024996), Emory Egleston Children's Research Center, and the Knights Templar Eye Foundation. Also supported by the NIH/National Center for Advancing Translational Sciences (grant UL-1TR-000077).

[1]Joseph McDonald, MD, Amy Cassedy, PhD, Mekibib Altaye, PhD, Jennifer Andringa, BS, Theresa Hennard, BS, Jessi Lipscomb, BS, Najima Mwase, MD, Virginia Miraldi Utz, MD: Cincinnati Children's Hospital, University of Cincinnati, Cincinnati, Ohio; [2]Ashley M. Cooper, MD, Erin Stahl, MD: Children's Mercy Hospital and University of Missouri-Kansas City, Kansas City; [3]Carolyn Drews-Botsch, PhD: Emory University School of Medicine and Emory University Rollins School of Public Health, Atlanta, Georgia; [4]George Engelhard Jr., PhD: The University of Georgia, Athens; [5]Gary N. Holland, MD: Stein Eye Institute and David Geffen School of Medicine, 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, Jessica Shantha, MD, Steven Yeh, MD: Emory University School of Medicine, Atlanta, Georgia; [9]Deborah K. McCurdy, MD: David Geffen School of Medicine, University of California, Los Angeles; [10]Sampath Prahalad, MD, MS: David Geffen School of Medicine, University of California, Los Angeles, and Emory University School of Medicine, Atlanta, Georgia;

SIGNIFICANCE & INNOVATIONS

- Studies assessing outcomes in children with uveitis rarely examine health-related quality of life, mental health, physical disability, and vision-related function and quality of life.
- We provide a comprehensive assessment of these measures beyond ophthalmic examination findings.
- Use of a uveitis-focused instrument shows differences in outcomes based on the presence of eye disease and may help us better understand the impact of uveitis.

other systemic diseases such as HLA–B27 disease, Behçet's disease, sarcoidosis, and vasculitides. Inflammation typically occurs in the anterior portion of the eye in children with JIA-U. Uveitis may also be present in other locations, as in intermediate, posterior, and panuveitis.

In pediatric uveitis, the ophthalmic examination remains the primary method to assess disease status and to guide therapy. Intraocular inflammation (presence of inflammatory cells and flare) represent disease activity, visual acuity (VA) indicates level of vision impairment, and presence of ocular complications denotes ocular damage (6). These measures, however, likely underestimate the true impact of disease on the child, since they do not account for the effect of vision impairment and the burden of chronic disease management on a child's quality of life (QoL) and functioning (7). Few studies have examined QoL in children with pediatric uveitis (8,9). Most of these studies focused on general health-related QoL (HRQoL) and have not included uveitis-specific measures or assessments of mental health (10–12).

Studies in adults with uveitis using adult-based vision and mental health questionnaires have demonstrated that those with uveitis and vision impairment have worse QoL and visual functioning, and higher rates of anxiety and depression (13–16). In addition, the impact of uveitis diagnosed in childhood has longstanding effects into adulthood, with data showing increased rates of anxiety and depression even in those adult patients with normal VA (17). Characterizing QoL outcomes in children with uveitis could complement clinical assessment and improve disease management. The aim of this study is to examine HRQoL, mental health, physical disability, vision-related functioning (VRF), and vision-related QoL (VRQoL) in children with JIA, JIA-U, and with other noninfectious pediatric uveitis diagnoses.

PATIENTS AND METHODS

Study population. A multicenter cross-sectional study was performed at 4 sites, including Cincinnati Children's Hospital Medical Center (CCHMC) in Cincinnati, Ohio; Emory University (EU) in Atlanta, Georgia; Children's Mercy Hospital (CMH) in Kansas City, Missouri; and University of California Los Angeles (UCLA) Stein Eye Institute and UCLA Mattel Children's Hospital in Los Angeles, California. All study sites had local institutional review board approval. Children with a diagnosis of JIA, JIA-U, or noninfectious uveitis of any etiology (other uveitis) were eligible to participate. All age-eligible patients seen in clinic were recruited during their regularly scheduled rheumatology or ophthalmology clinic visits from November 30, 2011 to September 27, 2019. Patients at EU and CCHMC were enrolled in a prospective epidemiology study. For this analysis, only baseline data were used. Patients at CMH and UCLA were enrolled in a cross-sectional study and completed one study visit. Inclusion criteria were 1) a diagnosis of JIA per the International League of Associations for Rheumatology classification criteria and/or uveitis (18); 2) age of 5–18 years at time of the study visit; and 3) English speaking. Exclusion criteria were 1) comorbidity unrelated to uveitis or arthritis that substantially affects QoL and functioning; 2) major developmental disorders; 3) non-English speaking; and 4) inability to complete questionnaires for any reason. Informed consent/assent was obtained for every patient.

Data collection. *Disease characteristics.* At the baseline study visit parents and patients completed disease and demographic information questionnaires, which included the following variables: age, sex, self-identified race and ethnicity, JIA subtype, JIA diagnosis, and/or uveitis diagnosis. Medical chart review was performed to obtain disease duration for JIA and/or uveitis, current medication use, history of ocular complications, and anterior uveitis disease activity (anterior chamber [AC] cells) per the Standardization of Uveitis Nomenclature criteria and best-corrected VA (BCVA) from the ophthalmic examination (6). Uveitis location was defined as anterior and nonanterior (intermediate, posterior, and panuveitis). Anterior uveitis included children with JIA-U and anterior uveitis only.

Patient-reported outcome measures. Patients (if age appropriate) and parents also completed patient-reported outcome measures at the time of the study visit, prior to their ophthalmology evaluation. Information regarding the number of patients and parents completing each patient-reported outcome measure can be found in Supplementary Table 1, available on the *Arthritis*

[11]A. Adrienne Walker, PhD: Georgia Department of Education, Atlanta, Georgia; [12]Sheila T. Angeles-Han, MD, MSc: Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio, and Children's Mercy Hospital, University of Missouri-Kansas City, Kansas City.

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

Address correspondence to Joseph McDonald, MD, Division of Rheumatology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, OH 45229. Email: joemcdonald93@gmail.com.

Submitted for publication August 17, 2020; accepted in revised form January 5, 2021.

Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24551.

VRF and VRQoL were assessed using the Effects of Youngsters' Eyesight on Quality of Life (EYE-Q) questionnaire. The EYE-Q is a questionnaire for children and adolescents ages 5–18 years that contains 25 questions assessing VRF and ophthalmic symptoms (i.e., near and far vision, color vision, night vision, photosensitivity) and VRQoL (i.e., feelings about use of medication, participating in activities related to vision, and having a uveitis diagnosis) (19–21). Each question is answered on a 3-point Likert scale. The response format measures the level of difficulty in performing a task for VRF and how true a QoL statement is for VRQoL. Answers are converted to scores with a range from 0 to 100. Higher scores indicate better VRF and VRQoL. Patients age 8 years and older were eligible to complete the child forms.

General HRQoL was assessed using the Pediatric Quality of Life Inventory, version 4.0 (PedsQL) (22). It is a 23-item measure of general HRQoL in children ages 2–18 years of age and includes 4 core scales: 1) physical, 2) emotional, 3) social, and 4) school functioning. We focused on total, physical, and psychosocial domains. Scores range from 0 to 100, with higher scores indicating better overall HRQoL. Patients age 5 years and older were eligible to complete the form for children.

Anxiety and depression were assessed using the Revised Childhood Anxiety and Depression Scale (RCADS) (23). It is one of the more widely distributed, brief screening tools used in children for evaluating symptoms of anxiety and depression. Scores of 70 or higher on the RCADS indicate clinically significant anxiety and depression. Patients age 8 and older were eligible to complete the form for children.

Physical functioning was assessed using the Childhood Health Assessment Questionnaire (C-HAQ) (24). It is a valid arthritis-specific measure that evaluates physical functioning and disability in 8 domains: 1) dressing and grooming, 2) arising, 3) eating, 4) walking, 5) hygiene, 6) reach, 7) grip, and 8) activities. Scores range from 0 to 3, with higher scores indicating worse physical functioning. Patients age 8 years and older were eligible to complete the form for children.

Statistical analysis. Measures of central tendency, variability, and association were calculated for all variables in the study. Frequencies (percentages), means (SDs), and medians (interquartile ranges [IQRs]) were used to describe the distribution of the data, overall and by disease group. The associations between the disease group and demographic, clinical, and psychosocial outcome variables were tested using chi-square test (χ^2), Student's *t*-test (correcting for unequal variances using the Satterthwaite methods, when necessary), Kruskal-Wallis test when comparing differences between group medians, and general linear models for multivariable models controlling for study site. In order to examine the specific group differences when all three disease groups were in the model, a Bonferroni corrected alpha level of *P* less than 0.017 for statistical significance was used. When only two groups were examined, an alpha level of *P* less than 0.05 was used to determine significance. Probability values were reported in this study. Data was analyzed using SAS software, version 9.3.

RESULTS

Overall patient characteristics. Of 549 patients enrolled in the study, 332 (60.5%) had JIA only, 124 (22.6%) had JIA-U, and 93 (16.9%) had other uveitis (Table 1). The majority of patients were White (81.2%), non-Hispanic (92.5%), and female (70.5%), with a mean ± SD age of 10.5 ± 4.0 years.

Patients with arthritis (JIA and JIA-U). Overall, the median age at JIA diagnosis was 5.4 years (IQR 2.8–10.4 years), and the median JIA disease duration was 2.6 years (IQR 0.8–6.1 years). Oligoarticular-persistent JIA was the most common subtype, in 39.7% of all JIA patients. There were group differences at age of JIA diagnosis and duration of disease. Patients with JIA-U were younger at diagnosis of their arthritis than those with JIA only (3.0 versus 6.4 years; $P < 0.001$) and had a longer JIA disease duration (5.5 versus 1.9 years; $P < 0.001$).

Patients with uveitis (JIA-U and other uveitis patients). Overall, the median age at uveitis diagnosis among all uveitis patients was 6.1 years (IQR 3.5–10.2 years) and the median disease duration was 2.8 years (IQR 0.6–6.3 years). A majority of uveitis patients had bilateral disease (68.2%). All JIA-U patients with known location of disease had anterior involvement (n = 117). Of the other uveitis patients, 51.6% had anterior disease only, and 48.4% had disease extending beyond the AC cells. Of 141 uveitis patients having an eye examination within 60 days of the study visit, 120 (85.1%) had anterior disease only, with 48 (40%) having active uveitis (≥0.5 positive AC cells) and 72 (60%) experiencing no activity at the time. Of 182 uveitis patients with VA recorded, 29.1% had bilateral and 13.7% had unilateral visual impairment equal to or worse than 20/50, and 11.0% had bilateral and 3.8% had unilateral visual impairment equal to or worse than 20/200 during the course of their disease.

There were group differences for demographic and clinical characteristics. Patients with JIA-U were younger at uveitis diagnosis (median age of 4.0 versus 9.1 years; $P < 0.001$) and were more likely to have anterior disease (93.6% versus 65.6%; $P < 0.001$) compared to patients with other uveitis. There were more male (46.2% versus 20.2%; $P < 0.001$) and Black patients (38% versus 8.9%; $P < 0.001$) with other uveitis compared to JIA-U patients. They were also more likely to have intermediate disease (28.0% versus 2.4%; $P < 0.001$). Patients with other uveitis had increased visual impairment (56.2% versus 33.7%; $P = 0.002$), synechiae (41.3% versus 24.2%; $P = 0.007$), and cystoid macular edema (20.7% versus 2.4%; $P < 0.001$). There were no significant differences in other ocular complications. When examining uveitis based on location, those with non-anterior

uveitis (intermediate, posterior, and panuveitis) had increased rates of cystoid macular edema compared to those with anterior uveitis (JIA-U and other uveitis).

QoL measures. *VRQoL (EYE-Q).* Children with JIA-U reported significantly worse VRF and VRQoL compared to those with JIA only, as seen in the total scores (mean ± SD 82.1 ± 1.4

Table 1. Characteristics of children with JIA and various forms of uveitis*

Characteristic	Overall (n = 549)	JIA only (n = 332)	JIA-U (n = 124)	Other uveitis† (n = 93)	*P*
Study site					<0.001‡
CCHMC	180 (32.8)	102 (30.7)	43 (34.7)	35 (37.6)	–
Emory	337 (61.4)	230 (69.3)	50 (40.3)	57 (61.3)	–
CMH	16 (2.9)	0 (0.0)	16 (12.9)	0 (0.0)	–
UCLA	16 (2.9)	0 (0.0)	15 (12.1)	1 (1.1)	–
Demographic information					
Age, mean ± SD years	10.5 ± 4.0	10.4 ± 4.0	10.5 ± 3.9	11.1 ± 3.9	0.319
Male sex	164 (30.2)	97 (29.5)	25 (20.2)	42 (46.2)	<0.001‡
Race (not mutually exclusive)					
White	445 (81.2)	289 (87.1)	100 (80.7)	56 (60.9)	<0.001‡
African American	74 (13.5)	28 (8.4)	11 (8.9)	35 (38.0)	<0.001‡
Asian	19 (3.5)	11 (3.3)	6 (4.8)	2 (2.2)	0.554
Multi-racial	14 (2.6)	7 (2.1)	2 (1.7)	5 (5.6)	0.15
American Indian/Alaskan Native	10 (1.8)	7 (2.1)	2 (1.6)	1 (1.1)	0.795
Unknown/declined	10 (1.8)	3 (0.9)	6 (4.8)	1 (1.1)	<0.017‡
Ethnicity, Hispanic	40 (7.5)	18 (5.5)	16 (13.2)	6 (6.7)	0.022‡
JIA characteristics					
Age at diagnosis, median (IQR) years	5.4 (2.8–10.4)	6.4 (3.6–10.9)	3.0 (1.4–6.2)	–	<0.001‡
Duration of disease, median (IQR) years	2.6 (0.8–6.1)	1.9 (0.5–5.2)	5.5 (2.0–9.3)	–	<0.001‡
JIA subtype					
Oligoarticular persistent	181 (39.7)	125 (37.7)	56 (45.2)	–	0.145
Oligoarticular extended	37 (8.1)	24 (7.2)	13 (10.5)	–	0.257
Polyarticular RF negative	71 (15.6)	64 (19.3)	7 (5.7)	–	<0.001‡
Uveitis characteristics					
Age at diagnosis, median (IQR) years	6.1 (3.5–10.2)	–	4.0 (2.8–6.4)	9.1 (6.1–11.9)	<0.001‡
Duration of disease, median (IQR) years	2.8 (0.6–6.3)	–	5.1 (1.4–8.4)	1.4 (0.4–3.3)	<0.001‡
Bilateral disease	148 (68.2)	–	84 (67.7)	64 (68.8)	0.202
AC cells§					0.519
0 cells	84 (59.6)	–	44 (57.1)	40 (62.5)	
≥0.5 cells	57 (40.4)	–	33 (42.9)	24 (37.5)	
Visual acuity 20/50 or worse	79 (43.4)	–	35 (33.7)	44 (56.2)	0.002‡
Visual acuity rank ordered					0.047‡
20/20 to 20/40	104 (57.1)	–	69 (66.4)	35 (44.3)	–
20/50 to 20/190	51 (28.0)	–	23 (22.1)	28 (35.0)	–
20/200 or worse	27 (14.8)	–	10 (9.8)	17 (21.2)	–
Location					
Anterior	177 (81.6)	–	117 (94.4)	61 (65.6)	<0.001‡
Intermediate	29 (13.4)	–	2 (1.6)	26 (28.0)	<0.001‡
Posterior	5 (2.3)	–	0 (0.0)	5 (5.4)	0.009‡
Panuveitis	10 (4.6)	–	0 (0.0)	10 (10.8)	<0.001‡
Ocular complications, ever					
Glaucoma/glaucoma suspect	30 (14.0)	–	16 (12.9)	14 (15.4)	0.604
Cataracts	58 (26.9)	–	27 (21.8)	31 (33.7)	0.051
Synechiae	68 (31.5)	–	30 (24.2)	38 (41.3)	0.007‡
Band keratopathy	40 (18.5)	–	21 (16.9)	19 (20.6)	0.487
Amblyopia	13 (6.0)	–	9 (7.3)	4 (4.3)	0.364
Cystoid macular edema	22 (10.2)	–	3 (2.4)	19 (20.7)	<0.001‡
Other¶	63 (29.2)	–	29 (24.0)	34 (37.0)	0.03‡
Treatment, at time of visit					
Steroid drops	109 (50.2)	–	52 (41.9)	57 (61.3)	0.005‡
Dilating drops	10 (4.6)	–	5 (4.0)	5 (5.4)	0.64
IOP-lowering drops§	35 (16.1)	–	18 (14.5)	17 (18.3)	0.456
Methotrexate oral	101 (18.4)	64 (19.3)	27 (21.8)	10 (10.8)	0.094
Methotrexate SQ	119 (21.7)	72 (21.7)	28 (22.6)	19 (20.4)	0.93
Mycophenolate	11 (2.0)	0 (0.0)	8 (6.5)	3 (3.2)	<0.001‡
Etanercept	35 (6.4)	34 (10.2)	1 (0.8)	0 (0.0)	<0.001‡

(Continued)

Table 1. (*Cont'd*)

Characteristic	Overall (n = 549)	JIA only (n = 332)	JIA-U (n = 124)	Other uveitis† (n = 93)	*P*
Infliximab	35 (6.4)	6 (1.8)	21 (16.9)	8 (8.6)	<0.001‡
Adalimumab	56 (10.2)	29 (8.7)	25 (20.2)	2 (2.2)	<0.001‡
Abatacept	10 (1.8)	8 (2.4)	2 (1.6)	0 (0.0)	0.302
Tocilizumab	20 (3.6)	12 (3.6)	8 (6.5)	0 (0.0)	0.043‡

* Values are the number (%) unless otherwise indicated. AC = anterior chamber; CCHMC = Cincinnati Children's Hospital Medical Center; CMH = Children's Mercy Hospital; IOP = intraocular pressure; IQR = interquartile range; JIA = juvenile idiopathic arthritis; JIA-U = JIA-associated uveitis; RF = rheumatoid factor; SQ = subcutaneous injection; UCLA = University of California, Los Angeles.
† Other uveitis: all other uveitis not associated with JIA.
‡ $P < 0.05$.
§ Cells taken at eye examination closest to study visit ± a 60-day period; 47 (38%) of JIA-U patients and 29 (31%) of uveitis only patients either did not have an eye examination within a 60-day period of the study visit or had missing data.
¶ Other complications include vitreous hemorrhage, optic disc edema, aphakia, choroidal neovascular membranes, choriorential scare, retinal neovascularization, retinal detachment, keratic precipitates, peri retinal fibrosis, floaters, blindness, ocular hypertension.

versus 89.7 ± 0.9; $P < 0.001$), VRF (mean ± SD 86.0 ± 1.5 versus 91.4 ± 1.0; $P = 0.006$), and VRQoL (mean ± SD 70.4 ± 2.2 versus 82.0 ± 1.6; $P < 0.001$) for the parent report (Table 2). Results were similar in the child reports, with the exception of VRF scores.

When comparing according to disease location, all patients with anterior disease (JIA-U and anterior uveitis only) had better total EYE-Q scores (mean ± SD 81.8 ± 1.2 versus 76.7 ± 2.2; $P = 0.027$) and VF scores (mean ± SD 85.2 ± 1.4 versus 79.6 ± 2.5; $P = 0.037$) by parent report compared to uveitis patients with disease in other locations (intermediate, posterior, and panuveitis) (Table 3). These differences were not seen in child reports.

Children with JIA-U and those with anterior uveitis only (no arthritis) did not have significant differences in EYE-Q scores by parent or child report (Table 4). When excluding patients with JIA, there were also no differences appreciated between other uveitis patients with anterior disease compared to other uveitis patients with non-anterior disease (Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24551).

All patients with uveitis had VRF and VRQoL assessed based on laterality (unilateral versus bilateral disease), vision impairment (VA), and presence of ocular complications. EYE-Q total and VRF scores were worse by parent report among uveitis patients with bilateral disease compared to those with unilateral disease (data not shown). These differences were not appreciated by child report. Additionally, EYE-Q scores were worse by parent and child report based on degree of vision impairment (Table 5).

Table 2. Comparison of patient-reported outcome measures in children with JIA with and without uveitis*

Outcome measure	JIA only (n = 332)	JIA-U, anterior only (n = 112)	*P*
EYE-Q parent total	89.7 ± 0.9	82.1 ± 1.4	<0.001†
Vision-related function	91.4 ± 1.0	86.0 ± 1.5	0.011†
Vision-related QoL	82.0 ± 1.6	70.4 ± 2.2	<0.001†
EYE-Q child total	86.8 ± 1.1	82.0 ± 1.6	0.016†
Vision-related function	87.8 ± 1.1	84.2 ± 1.7	0.094
Vision-related QoL	81.1 ± 1.9	74.7 ± 2.6	0.007†
RCADS parent total	43.1 ± 1.5	43.0 ± 1.8	0.920
RCADS child total	34.4 ± 1.8	32.9 ± 2.0	0.460
PedsQL parent total	78.3 ± 1.2	79.7 ± 2.0	0.866
Physical	76.7 ± 2.7	80.3 ± 2.5	0.371
Psychosocial	80.2 ± 2.2	79.2 ± 2.0	0.391
PedsQL child total	78.8 ± 1.2	79.6 ± 1.9	0.822
Physical	80.8 ± 2.4	82.3 ± 2.2	0.689
Psychosocial	77.9 ± 2.3	78.2 ± 2.1	0.935
C-HAQ parent total	0.4 ± 0.06	0.26 ± 0.05	0.051
C-HAQ child total	0.4 ± 0.07	0.32 ± 0.06	0.285

* Values are the mean ± SD except where indicated otherwise. All models controlled for study site, sex, race, and age. C-HAQ = Childhood Health Assessment Questionnaire, scores range 0–3, higher scores indicate higher difficulty with activities of daily living; EYE-Q = Effects of Youngsters' Eyesight on Quality of Life, scores range 0–100, lower scores indicate worse visual function and vision-related quality of life (QoL); JIA = juvenile idiopathic arthritis; JIA-U = JIA-associated uveitis; PedsQL = Pediatric Quality of Life Inventory, scores range 0–100, higher scores indicate better QoL; RCADS = Revised Children's Anxiety and Depression Scale, scores ≥70 indicate clinically significant anxiety and depression.
† $P < 0.05$.

Table 3. Comparison of patient-reported outcome measures in children with anterior uveitis and non-anterior uveitis*

Outcome measure	JIA-U and anterior other uveitis (n = 160)	Non-anterior other uveitis (n = 45)	*P*
EYE-Q parent total	81.8 ± 1.2	76.7 ± 2.2	0.027†
Vision-related function	85.2 ± 1.4	79.6 ± 2.5	0.037†
Vision-related QoL	71.3 ± 2.0	66.7 ± 3.5	0.128
EYE-Q child total	81.9 ± 1.5	81.8 ± 2.4	0.956
Vision-related function	84.4 ± 1.5	84.9 ± 2.5	0.887
Vision-related QoL	74.0 ± 2.4	72.8 ± 3.9	0.877
RCADS parent total	43.2 ± 1.5	40.6 ± 2.6	0.375
RCADS child total	32.0 ± 1.7	37.7 ± 2.9	0.079
PedsQL parent total	81.2 ± 1.9	87.7 ± 3.3	0.105
Physical	82.7 ± 2.3	95.7 ± 4.1	0.005†
Psychosocial	80.4 ± 1.9	83.5 ± 3.3	0.502
PedsQL child total	80.7 ± 1.8	87.1 ± 3.3	0.085
Physical	84.4 ± 2.1	96.7 ± 3.7	0.004†
Psychosocial	78.7 ± 1.9	82.1 ± 3.6	0.374
C-HAQ parent total	0.21 ± 0.05	0.05 ± 0.08	0.125
C-HAQ child total	0.27 ± 0.06	0.02 ± 0.10	0.024†

* Values are the mean ± SD except where indicated otherwise. All models controlled for study site, sex, race, and age. JIA-U = juvenile idiopathic arthritis–associated uveitis (see Table 2 for other definitions).
† $P < 0.05$.

EYE-Q scores were not significantly different for uveitis patients based on presence or absence of ocular complications. A comparison of EYE-Q scores based on disease activity was not performed because of lack of variability in disease severity.

Anxiety and depression (RCADS). There were no significant differences in RCADS scores among JIA, JIA-U, and other uveitis patients for both child and parent reports (Tables 2–4). All RCADS scores were below 70, suggesting that no patients were experiencing clinically significant anxiety or depression.

HRQoL (PedsQL). PedsQL total scores were similar in patients with JIA-U and JIA by parent report (mean ± SD 79.7 ± 2.0 versus 78.3 ± 1.2; $P = 0.866$) and child report (mean ± SD 79.6 ± 1.9 versus 78.8 ± 1.2; $P = 0.822$) (Table 2). Patients with JIA-U, however, had worse PedsQL physical summary scores compared to those with anterior other uveitis (no arthritis) by parent report (mean ± SD 80.3 ± 2.5 versus 91.0 ± 4.0; $P = 0.013$) and child report (mean ± SD 82.3 ± 2.2 versus 92.6 ± 3.7; $P = 0.008$) (Table 3). Comparisons were not performed between JIA-U patients and those with non-anterior uveitis since it is uncommon for children with JIA to have uveitis in other locations. There were no differences in PedsQL scores between patients with other uveitis based on uveitis location (Supplementary Table 2, available on

Table 4. Comparison of patient-reported outcome measures in children with anterior uveitis by presence of JIA*

Outcome measure	JIA-U, anterior (n = 112)	Anterior, other uveitis (n = 48)	*P*
EYE-Q parent total	82.1 ± 1.4	80.6 ± 2.2	0.516
Vision-related function	86.0 ± 1.5	82.4 ± 2.4	0.159
Vision-related QoL	70.4 ± 2.2	74.3 ± 3.4	0.274
EYE-Q child total	82.0 ± 1.6	81.7 ± 2.6	0.892
Vision-related function	84.2 ± 1.7	85.0 ± 2.6	0.765
Vision-related QoL	74.7 ± 2.6	71.4 ± 4.1	0.442
RCADS parent total	43.0 ± 1.8	43.5 ± 2.3	0.857
RCADS child total	32.9 ± 2.0	30.4 ± 2.5	0.408
PedsQL parent total	79.7 ± 2.0	86.8 ± 3.2	0.040†
Physical	80.3 ± 2.5	91.0 ± 4.0	0.013†
Psychosocial	79.2 ± 2.0	84.7 ± 3.3	0.118
PedsQL child total	79.6 ± 1.9	84.5 ± 3.2	0.152
Physical	82.3 ± 2.2	92.6 ± 3.7	0.008†
Psychosocial	78.2 ± 2.1	80.3 ± 3.5	0.584
C-HAQ parent total	0.26 ± 0.05	0.05 ± 0.08	0.017†
C-HAQ child total	0.32 ± 0.06	0.09 ± 0.11	0.046†

* Values are the mean ± SD except where otherwise indicated. All models controlled for study site, sex, race, and age. JIA-U = juvenile idiopathic arthritis–associated uveitis (see Table 2 for other definitions).
† $P < 0.05$.

Table 5. Vision-specific outcome measures by visual acuity in children with uveitis*

	Group 1, 20/20–20/30	Group 2, 20/40–20/60	Group 3, ≥70	*P*, group comparison†		
EYE-Q parent						
Total	90.7 (1.9)	77.1 (3.0)	65.2 (3.2)	<0.001	<0.001	0.011
VRF	94.8 (2.4)	81.7 (3.4)	66.3 (3.7)	<0.001	<0.001	0.004
VRQoL	80.3 (2.7)	64.1 (4.3)	58.8 (4.6)	<0.001	<0.001	1
EYE-Q child						
Total	89.4 (2.2.)	78.9 (3.3)	75.6 (3.7)	0.008	0.004	1
VRF	92.4 (2.2)	81.2 (3.3)	77.1 (3.7)	0.005	0.002	1
VRQoL‡	80.8 (3.7)	70.6 (5.5)	68.6 (6.4)	0.245	0.291	1

* Values are the least mean square (SE), except where otherwise indicated. All models controlled for site, race, age, and sex. EYE-Q = Effects of Youngsters' Eyesight on Quality of Life; scores range 0–100, lower scores indicate worse visual function (VRF) and vision-related quality of life (VRQoL).
† Bonferroni adjusted for multiple comparison.
‡ Nonsignficant findings due to the large SE for both groups 2 and 3.

the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24551). Notably, there were no significant differences seen in the psychosocial summary scores of the PedsQL for any of the group comparisons.

Physical functioning (C-HAQ). In patients with arthritis, C-HAQ scores by parent and child report did not differ between JIA-U versus JIA only patients, suggesting similar difficulty with activities of daily living, as was also seen with the PedsQL physical summary scores (Table 2). Patients with other uveitis with anterior disease (no arthritis) had better physical functioning compared to those with JIA-U by both parent report (mean ± SD 0.05 ± 0.08 versus 0.26 ± 0.05; $P = 0.017$) and child report (mean ± SD 0.09 ± 0.11 versus 0.32 ± 0.06; $P = 0.046$) (Table 4). There were no differences among uveitis only patients with anterior disease compared to those patients with uveitis in all other locations (Supplementary Table 2).

Therapy regimen and VRQoL at time of study visit. Topical medications. Among all uveitis patients, 109 (50.2%) were using topical glucocorticoids, 35 (16.1%) were using intraocular pressure (IOP)-lowering drops, and 10 (4.6%) were using dilating drops (Table 1). When comparing JIA-U to other uveitis patients, other uveitis patients were more likely to be using glucocorticoid drops at the time of the visit (61.3% versus 41.9%; $P = 0.027$). There were no differences with frequency of use of dilating drops or IOP-lowering drops.

Systemic medications. Methotrexate was the most common systemic medication used. Of the 220 (40.1%) total patients taking methotrexate, 119 (54%) were receiving subcutaneous injections. Biologics were the next most common class of medication: adalimumab 56 (10.2%), infliximab 35 (6.3%), etanercept 35 (6.3%), tocilizumab 20 (3.6%), and abatacept 10 (1.8%).

The impact of treatment on VRF and VRQoL in uveitis patients (those with JIA-U and other uveitis) was evaluated by examining EYE-Q scores based on treatment regimen. Uveitis patients were split into groups, including patients receiving no treatment (n = 65), topical treatment only (n = 39), systemic treatment only (n = 47), and both topical and systemic treatment (n = 65) (Table 6).

There were differences in EYE-Q scores when comparing no treatment versus combined treatment. EYE-Q parent total scores (mean ± SD 76.2 ± 2.1 versus 85.5 ± 2.7; $P = 0.015$) and child total scores (mean ± SD 77.1 ± 2.1 versus 85.6 ± 2.4; $P = 0.013$) were worse when taking combined treatment versus no treatment. Children and parents in the combined treatment group reported lower VRF subscores compared to the no treatment group. In addition, children in the combined treatment

Table 6. Vision-specific outcome measures by treatment regimen in children with uveitis*

Measure	A: no topical or systemic Rx (n = 48)	B: topical Rx (n = 44)	C: systemic Rx (n = 51)	D: topical and systemic Rx (n = 73)	*P*, A vs. B	*P*, A vs. C	*P*, A vs. D
EYE-Q parent							
Total	85.5 (2.7)	82.8 (2.8)	84.8 (2.4)	76.2 (2.1)	1	1	0.015†
VRF	89.3 (3.2)	84.8 (3.2)	88.9 (2.8)	79.3 (2.5)	1	1	0.034†
VRQoL	73.0 (3.3)	76.1 (3.4)	72.6 (2.9)	65.0 (2.6)	1	1	0.197
EYE-Q child							
Total	85.6 (2.4)	84.3 (2.6)	83.8 (2.3)	77.1 (2.1)	1	1	0.013†
VRF	88.5 (2.6)	86.6 (2.9)	85.6 (2.5)	80.3 (2.3)	1	1	0.040†
VRQoL	77.3 (3.6)	76.6 (3.9)	78.8 (3.4)	66.0 (3.1)	1	1	0.037†

* Values are the least mean square (SE), except where otherwise indicated. EYE-Q = Effects of Youngsters' Eyesight on Quality of Life, scores range 0–100, lower scores indicate worse visual function (VRF) and vision-related quality of life (VRQoL).
† $P < 0.05$.

group reported lower VRQoL scores compared to the no treatment group. There were no differences observed in EYE-Q total or subscores by parent or child report when comparing no treatment versus topical treatment, no treatment versus systemic treatment, and topical treatment versus systemic treatment (Table 6).

DISCUSSION

We provide a comprehensive assessment of QoL and functioning in a large multicenter cohort of children with JIA and noninfectious uveitis that includes a patient-reported outcome measure focused on children with uveitis. Our results suggest that uveitis has a significant impact on VRQoL and VRF, and that vision-specific instruments are important in the assessment of patient outcomes.

Overall, the demographic and disease characteristics of our cohort were consistent with previous reports (25–28). Patients with JIA-U were predominantly female, White, and presented with bilateral, anterior disease. They were diagnosed with JIA earlier than those with JIA alone, which supports that young age at diagnosis is associated with a higher risk for uveitis development (29,30). Similarly, our cohort of patients with uveitis only were older at diagnosis compared to those with JIA-U (31,32).

Children with uveitis can have significant burden of ocular disease. In our cohort, 48 of 120 children had active anterior disease within 60 days of their study visit. Of the total cohort, 78% needed topical and/or systemic medication at the time of their study visit. Further, 37% experienced one or more ocular complications, and over 40% had visual impairment during their disease course. In general, children with non-anterior uveitis tend to have worse visual outcomes (33,34). In our cohort there were higher rates of visual impairment, ocular complications, and topical glucocorticoid use in patients with other uveitis compared to JIA-U. This may be secondary to delayed uveitis diagnosis given the lack of regular ophthalmic screening that would come with a diagnosis of JIA or more likely the extent of ocular involvement often seen in non-anterior disease uveitis. Additionally, patients with other uveitis may not be receiving the benefit of early initiation of systemic therapy that comes with a diagnosis of JIA and would have increased ocular complications secondary to longer courses of topical therapy. These burdens highlight the importance of patient-reported outcome measures to determine the effect uveitis-related disease complications and management regimens have on these children.

The use of the EYE-Q was able to distinguish patients by uveitis diagnosis, illustrated by significant differences in total, VRF, and VRQoL scores between those patients with JIA only compared to those with JIA-U. The EYE-Q was also able to demonstrate differences between those patients based on laterality and degree of vision impairment, similar to our earlier studies. In contrast, the PedsQL and C-HAQ did not distinguish patients by the presence of eye disease. These questionnaires, however, did differentiate patients by arthritis diagnosis. Patients with arthritis (JIA and JIA-U groups) had worse PedsQL scores compared to patients without arthritis (other uveitis group). This was specifically seen in the physical summary scores of the PedsQL, likely secondary to underlying arthritis. Additionally, patients with arthritis had C-HAQ scores suggestive of mild to moderate physical disability (35). Interestingly, patients with other uveitis had PedsQL scores that are comparable to healthy controls, which suggests that they may have similar overall QoL to healthy children or that the PedsQL is not detecting vision-specific effects from uveitis (22). As expected, C-HAQ scores were similar to patients with no physical disability since these patients did not have underlying arthritis. Thus, a vision-specific instrument such as the EYE-Q likely assesses vision-specific impairments from uveitis not assessed by PedsQL.

Notably, 2 previous studies using the EYE-Q to assess VRQoL and VRF in children with JIA-U were not able to distinguish patients based on underlying uveitis diagnosis, but those with worse BCVA had worse EYE-Q scores (7,8). The authors noted, however, that small sample sizes were a major limitation. Likewise, older versions of the EYE-Q were administered in these studies (34). The EYE-Q has undergone substantial modifications based on additional validation studies (19–21).

Topical and systemic immunosuppressant medications are often required for the management of uveitis. Patients receiving combined topical and systemic therapy had worse EYE-Q scores compared to patients receiving no treatment. This difference was seen in all domains for both parent and patient reports except for VRQoL scores in children. The EYE-Q VRQoL scores were used to identify burdens associated with differing therapy regimens, since these questions on the EYE-Q related specifically to medication use. This suggests there may be differences in how the child is experiencing the burden of medication use compared to the parent. The burden of medication regimen in pediatric uveitis warrants further study as this can affect medication adherence and patient outcomes.

This study is one of the first to report on anxiety and depression among patients with pediatric uveitis. Our cohort of patients did not report clinically significant anxiety or depression based on the RCADS, in contrast to adult studies, which have shown higher rates of depression and anxiety in patients with uveitis diagnoses (15,16).

Regarding patient–child concordance, previously published data show discordance between parent and child reports in children with visual impairment completing VRF and VRQoL measures (36). In our cohort there was fair concordance between parent and child reports for the EYE-Q. The only area where there was discordance was related to visual functioning, wherein VRF scores between child and parent reports differed. Children with uveitis did not report significant VRF differences compared to those without uveitis, although their parent reports did.

This discrepancy may indicate that a child's perception of their own visual function may differ from their parent's perspective and stresses the importance of including both reports.

Our study has several limitations. Previous studies have shown that the EYE-Q correlates with visual function tests such as BCVA and contrast sensitivity, thus can distinguish children based on vision impairment and laterality of eye involvement (19,20). We were able to analyze the EYE-Q based on VA in our cohort with the VA data we had available; however VA was not entered for over 15% of the patients in the study. In addition, the questionnaires were not administered on the day of the eye examination in all patients since the visit was performed during the rheumatology visit. Only patients with uveitis were enrolled at CMH and UCLA due to the enrollment requirement of a uveitis diagnosis. There are likely treatment or physician practice differences among all sites. The differences observed in VRF and VRQoL outcomes based on treatment regimen may be confounded by medication indication relating to disease severity and not related to the medication use itself.

In conclusion, we described a large cohort of pediatric uveitis patients and provided a comprehensive outcome assessment, including traditional outcome measures in uveitis such as VA, disease activity, and ocular complications. The results were bolstered by the inclusion of patient-reported outcome instruments measuring general HRQoL, VRF, VRQoL, physical disability, and mental health. We demonstrated the importance of implementing a vision- and uveitis-specific measure, which was able to not only distinguish those patients with uveitis but also differentiate those with bilateral disease and vision impairment. The use of patient-reported outcome measures allows us to more accurately describe the effect that uveitis had on the everyday life of a child experiencing this disease. Larger and more diverse cohorts are needed to study the impact of visual impairment and ocular complications on QoL and functioning.

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. McDonald 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, Drews-Botsch, Lambert, McCracken, Prahalad, Walker, Angeles-Han.

Acquisition of data. Cooper, Drews-Botsch, Hennard, Holland, Jenkins, McCurdy, Mwase, Prahalad, Shantha, Stahl, Utz, Yeh, Angeles-Han.

Analysis and interpretation of data. McDonald, Cassedy, Altaye, Andringa, Cooper, Drews-Botsch, Engelhard, Hennard, Holland, Jenkins, Lambert, Lipscomb, McCracken, McCurdy, Shantha, Stahl, Utz, Walker, Yeh, Angeles-Han.

REFERENCES

1. Thorne JE, Suhler E, Skup M, Tari S, Macaulay D, Chao J, et al. Prevalence of noninfectious uveitis in the United States. JAMA Ophthalmol 2016;134:1237. URL: http://archopht.jamanetwork.com/article.aspx?doi=10.1001/jamaophthalmol.2016.3229.
2. Nordal E, Rypdal V, Christoffersen T, Aalto K, Berntson L, Fasth A, et al. Incidence and predictors of uveitis in juvenile idiopathic arthritis in a Nordic long-term cohort study. Pediatr Rheumatol 2017;15:66. URL: http://ped-rheum.biomedcentral.com/articles/10.1186/s12969-017-0195-8.
3. BenEzra D, Cohen E, Behar-Cohen F. Uveitis and juvenile idiopathic arthritis: a cohort study. Clin Ophthalmol 2007;1:513–8. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2704519&tool=pmcentrez&rendertype=abstract.
4. Kump LI, Cervantes-Castañeda RA, Androudi SN, Foster CS. Analysis of pediatric uveitis cases at a tertiary referral center. Ophthalmology 2005;112:1287–92.
5. De Boer J. Visual loss in uveitis of childhood. Br J Ophthalmol 2003;87:879–84.
6. Jabs DA, Nussenblatt RB, Rosenbaum JT, Standardization of Uveitis Nomenclature (SUN) Working Group. Standardization of uveitis nomenclature for reporting clinical data: results of the first international workshop. Am J Ophthalmol 2005;140:509–16. URL: http://www.ncbi.nlm.nih.gov/pubmed/16196117.
7. 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:xvi–xxiv.
8. Taha R, Papadopoulou M, Zetterberg M, Oskarsdottir S, Grönlund MA. Visual function and quality of life in a cohort of Swedish children with juvenile idiopathic arthritis. Clin Ophthalmol 2019;13:2081–91.
9. Ezzahri M, Amine B, Rostom S, Rifay Y, Badri D, Mawani N, et al. Uveitis and its relationship with disease activity and quality of life in Moroccan children with juvenile idiopathic arthritis. Clin Rheumatol 2013;32:1387–91.
10. 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.
11. Solebo AL, Barry RJ, Keane PA, Rahi JS, Denniston AK. Underutilisation of reproducible, child appropriate or patient reported outcome measures in childhood uveitis interventional research. Orphanet J Rare Dis 2019;14. URL: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L627968620%0Ahttps://doi.org/10.1186/s13023-019-1108-3.
12. 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.
13. Schiffman RM, Jacobsen G, Whitcup SM. Visual functioning and general health status in patients with uveitis. Arch Ophthalmol 2001;119:841–9.
14. Naik RK, Rentz AM, Foster CS, Lightman S, Belfort R, Lowder C, et al. Normative comparison of patient-reported outcomes in patients with noninfectious uveitis. JAMA Ophthalmol 2013;131:219–25.
15. Silva LM, Arantes TE, Casaroli-Marano R, Vaz T, Belfort R, Muccioli C. Quality of life and psychological aspects in patients with visual impairment secondary to uveitis: a clinical study in a tertiary care hospital in Brazil. Ocul Immunol Inflamm 2019;27:99–107.
16. Onal S, Oray M, Yasa C, Akman M, Uludag G, Koc Akbay A, et al. Screening for depression and anxiety in patients with active uveitis. Ocul Immunol Inflamm 2018;26:1078–93.
17. Haasnoot AM, 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;69:1895–1902.
18. 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: 2nd rev. Edmonton, 2001. J Rheumatol 2004;31:390–2.

19. 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.
20. 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.
21. Cassedy A, Altaye M, Andringa J, Cooper AM, Drews-Botsch C, Engelhard G, et al. Assessing the validity and reliability of the Effects of Youngsters' Eyesight on Quality of Life (EYE-Q) questionnaire among children with uveitis. Arthritis Care Res (Hoboken) 2022;74: 355–63.
22. 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.
23. Piqueras JA, Martín-Vivar M, Sandin B, San Luis C, Pineda D. The revised Child Anxiety and Depression Scale: a systematic review and reliability generalization meta-analysis. J Affect Disord 2017;218: 153–69.
24. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 1994;37:1761–9.
25. Sabri K, Saurenmann RK, Silverman ED, Levin AV. Course, complications, and outcome of juvenile arthritis-related uveitis. J AAPOS 2008; 12:539–45.
26. 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.
27. 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. Pediatr Rheumatol 2019;17:66.
28. 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.
29. Kotaniemi K, Kautiainen H, Karma A, Aho K. Occurrence of uveitis in recently diagnosed juvenile chronic arthritis: a prospective study. Ophthalmology 2001;108:2071–5.
30. Heiligenhaus A, Niewerth M, Ganser G, Heinz C, Minden K, Ganser G, et al. 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.
31. Cann M, Ramanan AV, Crawford A, Dick AD, Clarke SL, Rashed F, et al. Outcomes of non-infectious paediatric uveitis in the era of biologic therapy. Pediatr Rheumatol 2018;16.
32. Heiligenhaus A, Klotsche J, Niewerth M, Horneff G, Ganser G, Haas JP, et al. Similarities in clinical course and outcome between juvenile idiopathic arthritis (JIA)-associated and ANA-positive idiopathic anterior uveitis: data from a population-based nationwide study in Germany. Arthritis Res Ther 2020;22.
33. Reiff A. Clinical presentation, management, and long-term outcome of pars planitis, panuveitis, and Vogt-Koyanagi-Harada disease in children and adolescents. Arthritis Care Res (Hoboken) 2019;72: 1589–96.
34. Rosenberg KD, Feuer WJ, Davis JL. Ocular complications of pediatric uveitis. Ophthalmology 2004;111:2299–306.
35. Dempster H, Porepa M, Young N, Feldman BM. The clinical meaning of functional outcome scores in children with juvenile arthritis. Arthritis Rheum 2001;44:1768–74.
36. Tadic V, Cumberland PM, Lewando-Hundt G, Rahi JS. Do visually impaired children and their parents agree on the child's vision-related quality of life and functional vision? Br J Ophthalmol 2017;101: 244–50.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1321–1324
DOI 10.1002/acr.24575

AMERICAN COLLEGE of RHEUMATOLOGY
Empowering Rheumatology Professionals

BRIEF REPORT

Trends in Permanent Work Disability Associated With Rheumatoid Arthritis in the United States, 1999–2015

Michael M. Ward

Objective. Advances in treatment over the past 20 years have resulted in improved control of rheumatoid arthritis (RA). The objective of our study was to investigate whether there has been a decrease in permanent work disability associated with RA in the US.

Methods. Medicare data from 1999 to 2015 were used to identify beneficiaries age 20–59 years with RA who became eligible for Medicare coverage under Social Security Disability Insurance. Diagnosis of RA was based on physician claims in the first year of enrollment. Annual rates of enrollment were sex- and age-standardized to the 2000 US population.

Results. The study included 97,787 beneficiaries with RA and Social Security Disability Insurance across all years. Medicare enrollment was 26.0 per million in 1999 and 26.0 per million in 2015. Rates increased following the Great Recession of 2008–2009 before returning to prerecession levels. There was no linear trend over time after adjusting for the annual national unemployment rate (relative risk 0.99 per year [95% confidence interval 0.99–1.00]; $P = 0.69$). Risks of work disability were much higher among workers over age 50 years.

Conclusion. Based on Medicare enrollment by recipients of Social Security Disability Insurance, there was no decrease in permanent work disability among young and middle-age workers with RA in the US between 1999 and 2015.

INTRODUCTION

The treatment of rheumatoid arthritis (RA) has changed dramatically over the past 20 years, with the emphasis on early and consistent use of disease-modifying medications and the introduction of biologics. With better treatments, the health outcomes of patients with RA have improved over time, with many studies reporting decreases in disease activity, less functional difficulty, less joint damage, and reduced need for joint surgery (1,2). Clinical remission is now an achievable goal.

Although joint inflammation and damage have decreased, whether other long-term outcomes such as permanent work disability have improved is not clear. In studies of patients observed in the late 1980s and 1990s, 30–50% of employed patients developed permanent work disability after 10 years of RA (3,4). Many studies suggest that tumor necrosis factor inhibitors are efficacious in preventing short-term work loss, which holds the prospect of lower incidences of permanent work loss (5,6). However, permanent work disability in RA depends not only on the severity of joint symptoms and impairments, but also on the type of work, workplace accommodations, psychological response to illness, and social and financial supports (4). Risks of permanent work loss are higher among older persons with RA and those with less formal education and physically demanding jobs (4).

Population-based studies from Sweden and Finland have reported decreases of >50% in rates of permanent work disability among persons with RA between 1990 and 2010 (7,8). In Norway, rates of permanent disability pensioning for RA were stable between 1968 and 1997, but lagged rising incidences in the general population between 1983 and 1997 (9). Similar national studies have not been reported in the US. Social Security Disability Insurance (SSDI) provides permanent benefits to eligible American workers who are certified as having a medical condition that makes them unable to work for at least 1 year and that is unlikely to improve. SSDI recipients are eligible for Medicare insurance after 2 years. The aim of this study was to examine rates of Medicare enrollment from 1999 to 2015 among SSDI recipients ages 20–59 years with RA.

Supported by the Intramural Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases (ZIA-AR-041153).

Michael M. Ward, MD, MPH: National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH, Bethesda, Maryland.

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

Address correspondence to Michael M. Ward, MD, MPH, NIH/NIAMS, Building 10 CRC, Room 4-1339, 10 Center Drive, Bethesda, MD 20892. Email: wardm1@mail.nih.gov.

Submitted for publication October 14, 2020; accepted in revised form February 2, 2021.

SIGNIFICANCE & INNOVATIONS

- This is the first population-based study of rates of permanent work disability associated with rheumatoid arthritis in the US.
- Rates of Medicare enrollment under Social Security Disability Insurance by persons with rheumatoid arthritis were the same in 1999 and 2015, after peaking in 2011 following the Great Recession.
- This is the first study to demonstrate the impact of national economic downturns on work disability in rheumatoid arthritis.
- The absence of a decrease in rates of permanent work disability among US workers with rheumatoid arthritis may indicate inadequate access to treatment among those at highest risk for work disability.

MATERIALS AND METHODS

The data source was 100% fee-for-service Medicare inpatient and outpatient administrative claims files from 1999 to 2015. In each year, I identified newly enrolled beneficiaries ages 20–59 years who entered Medicare via SSDI eligibility. I excluded those who entered because of end-stage renal disease. Among these beneficiaries, I identified those with RA, based on the presence of at least 2 claims on separate dates in the first year of Medicare enrollment, who had a principal diagnosis of RA based on International Classification of Diseases, Ninth Revision, code 714 (10). This criterion was based on the consideration that medical care would be sought specifically for the condition that permitted Medicare eligibility. The study protocol was approved by the National Institute of Diabetes and Digestive and Kidney Diseases Institutional Review Board, which waived the requirement for informed consent.

I examined rates by year of enrollment relative to the US population, based on census data. Rates were sex- and age-standardized (in 5-year age groups) to the US population in 2000. Applications for SSDI benefits and SSDI enrollment increase during economic downturns and job loss (11). National unemployment rates varied between 4.0% and 6.0% from 1999 to 2008, and increased to 9.3% and 9.6% in 2009 and 2010, respectively, before decreasing gradually to 5.3% in 2015 (12). To account for changes associated with the 2008–2009 Great Recession, I used Poisson regression models to determine whether there was a linear trend in SSDI enrollment, while adjusting for changes in the national unemployment rate over these years. The independent variables in the model were calendar year, the 2-year lagged national unemployment rate (to account for the time between SSDI claim and Medicare enrollment), sex, and indicator variables for each 5-year age group. This analysis provided the relative risk (RR) of the average yearly change in rate of SSDI enrollment, while partitioning out the contribution of the unemployment rate and variations in the sex and age composition of the sample over time. In sensitivity analyses, I examined beneficiaries who had 3 or more claims with RA as the principal diagnosis in the first year in Medicare, and those ages 20–39 years on enrollment in Medicare, who may have had greater potential to benefit from recent treatment advances.

RESULTS

Across all years, 97,787 beneficiaries (75.5% women; mean ± SD age 50.1 ± 8.3 years) enrolled in Medicare under SSDI with RA claims in their first year. In all, 69% had their first visit for RA within 90 days of Medicare enrollment. The rate of Medicare enrollment under SSDI was 26.0 per million in 1999 (Table 1 and Figure 1). Rates were comparable or higher in each subsequent year, and peaked at 40.0 per million in 2011, 2 years after the Great Recession.

There was no linear trend in Medicare enrollment among beneficiaries with RA over time (RR 0.99 per year [95% confidence interval (95% CI) 0.99–1.00]; $P = 0.69$) after adjustment for the rise in unemployment rates in 2009–2012. Risks increased progressively with age. Compared to those ages 20–24 years, the RR for enrollment among those ages 25–29 years was 1.75 (95% CI 1.63–1.88), while the RR among those ages 50–54 years was 20.85 (95% CI 19.68–21.98) and for those ages 55–59 years was 31.50 (95% CI 29.66–33.11). Risks were lower among men compared to women (RR 0.33 [95% CI 0.33–0.34]).

The subgroup with 3 claims for RA in the first year included 75,930 beneficiaries (76.0% women; mean ± SD age 50.2 ± 8.3 years), while the subgroup ages 20–39 years included 12,564 beneficiaries (79.2% women; mean ± SD age 33.2 ± 5.2 years). While the absolute rates were lower in both

Table 1. Rates of Medicare enrollment via Social Security Disability Insurance by beneficiaries with rheumatoid arthritis*

Year	Number	Rates
1999	4,057	26.0 (25.2–26.9)
2000	4,181	26.8 (26.0–27.7)
2001	4,685	29.3 (28.4–30.2)
2002	5,291	32.1 (31.2–33.0)
2003	5,757	34.2 (33.4–35.2)
2004	5,781	33.5 (32.7–34.5)
2005	5,721	32.3 (31.4–33.2)
2006	5,741	31.7 (30.8–32.5)
2007	5,361	29.4 (28.6–30.2)
2008	5,380	29.3 (28.5–30.1)
2009	6,027	32.5 (31.7–33.4)
2010	6,574	35.2 (34.3–36.1)
2011	7,530	40.0 (39.1–41.0)
2012	7,319	38.3 (37.4–39.3)
2013	7,013	36.5 (35.6–37.4)
2014	6,257	32.4 (31.5–33.2)
2015	5,112	26.0 (25.3–26.8)

* Rates are the age- and sex-standardized rate, per million population (95% confidence interval).

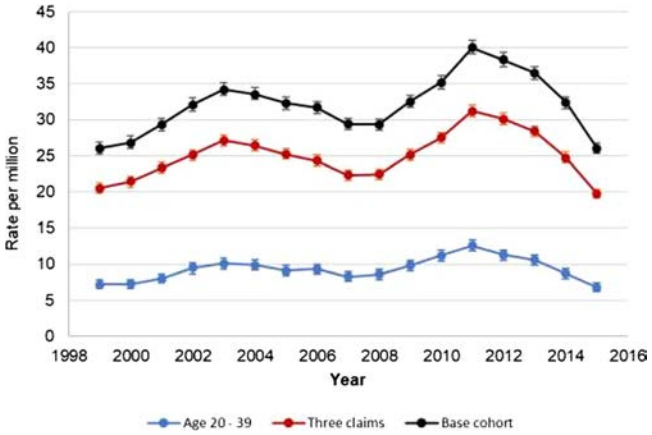


Figure 1. Age- and sex-standardized rates of Medicare enrollment under Social Security Disability Insurance by year for beneficiaries with rheumatoid arthritis. Error bars indicate 95% confidence limits.

subgroups, the incidences in these subgroups paralleled the incidence in the overall group, with no decrease over time (Figure 1). These results indicate that the use of a more stringent requirement for inclusion did not affect the conclusion, and that the results of the large proportion of middle-aged beneficiaries were not obscuring a decrease in SSDI enrollment over time among young adults.

DISCUSSION

These results do not indicate a decrease in rates of permanent work disability among American workers with RA between 1999 and 2015. Apart from temporary increases in Medicare enrollment under SSDI as a consequence of labor market changes following the Great Recession, rates of enrollment were stable over this period.

Older workers were most susceptible to work disability, as also shown in many prior studies (4). Compared to younger workers, this group likely included a higher proportion of workers with more longstanding RA, whose joint damage may have been less amenable to improvement (13). However, rates were also stable among those ages 20–39 years, who might have been expected to gain more benefit from recent treatment advances.

These findings raise questions about access to newer treatments, particularly by low-income nonprofessional workers who are most at risk for work disability related to RA (4). We did not have data on medication use and could not examine this question directly. However, lower income and less formal education have been consistently associated with less access to disease-modifying medications, including biologics, in patients with RA in the US (14,15). The absence of a national decrease in work disability over time may therefore reflect lack of access to treatment advances by workers of lower socioeconomic status, who are at highest risk of work disability. Access to new treatments by highly educated workers may not have had an impact on national rates of work disability because their baseline risk of health-related job loss was comparative low. Despite improvements in physical health over time among workers with RA, social or psychological influences possibly continued to foster work disability. Differences between these results and those from Nordic countries may relate to differences in the organization and financing of health care (7,8).

An individual's decision to pursue a work disability claim includes considerations of the severity of illness, coping resources and skills, prospects for future improvement in health, job requirements and accommodations, opportunities for retraining, age and the time horizon, social supports (including economic support from family), and personal wealth (3,4). Our results also highlight the association of disability claims with the national economy and demonstrate the major role that macrolevel factors have on what has often been considered a personal decision. In the US, SSDI claims increased by >100,000 per quarter following the Great Recession, and tracked more closely with changes in the national unemployment rate than the gross domestic product (11). These data show a similar peak among persons with RA in 2011–2012, which aligns with the lag between enrollment in SSDI and eligibility for Medicare. The smaller peak in 2003–2004 may be a consequence of the 2001 recession (11). If similar associations hold for the current economic downturn and unemployment due to the COVID-19 pandemic, we might expect sharp increases in disability claims for RA in the near future. This effect may be particularly pronounced given the disproportionate economic impact of the pandemic on female service workers.

The study is limited in that data after 2015 were not available. However, there was no suggestion of a decrease in work disability through 2015, many years after the introduction of new treatments. RA might have developed between enrollment in SSDI and enrollment in Medicare in some individuals, but this event was likely rare. The analysis also assumed that the prevalence of RA has been stable over these years. As in any study of administrative claims, there may be inaccuracies in coding, but differential inaccuracies over time that would be large enough to alter the trends are unlikely. The validity of coding is supported by the detection of the peak in rates following the Great Recession, as also found in all-cause disability claims (11).

These findings suggest that recent treatment advances have not yet had a major impact on permanent work disability associated with RA at the population level in the US. Future research should investigate the association between access to treatment and rates of work disability across the population.

AUTHOR CONTRIBUTIONS

Dr. Ward 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. Ward.

Acquisition of data. Ward.
Analysis and interpretation of data. Ward.

REFERENCES

1. Carpenter L, Barnett R, Mahendran P, Nikiphorou E, Gwinnutt J, Verstappen S, et al. Secular changes in functional disability, pain, fatigue and mental well-being in early rheumatoid arthritis: a longitudinal meta-analysis. Semin Arthritis Rheum 2020;50:209–19.
2. Cordtz RL, Hawley S, Prieto-Alhambra D, Højgaard P, Zobbe K, Overgaard S, et al. Incidence of hip and knee replacement in patients with rheumatoid arthritis following the introduction of biological DMARDs: an interrupted time-series analysis using nationwide Danish healthcare registers. Ann Rheum Dis 2018;77:684–9.
3. Burton W, Morrison A, Maclean R, Ruderman E. Systematic review of studies of productivity loss due to rheumatoid arthritis. Occup Med (Lond) 2006;56:18–27.
4. Verstappen SM, Bijlsma JW, Verkleij H, Buskens E, Blaauw AA, ter Borg EJ, et al. Overview of work disability in rheumatoid arthritis patients as observed in cross-sectional and longitudinal surveys. Arthritis Rheum 2004;51:488–97.
5. Ter Wee MM, Lems WF, Usan H, Gulpen A, Boonen A. The effect of biological agents on work participation in rheumatoid arthritis patients: a systematic review. Ann Rheum Dis 2012;71:161–71.
6. Wolfe F, Allaire S, Michaud K. The prevalence and incidence of work disability in rheumatoid arthritis, and the effect of anti-tumor necrosis factor on work disability. J Rheumatol 2007;34:2211–7.
7. Hallert E, Husberg M, Bernfort L. The incidence of permanent work disability in patients with rheumatoid arthritis in Sweden 1990–2010: before and after introduction of biologic agents. Rheumatology (Oxford) 2012;51:338–46.
8. Rantalaiho VM, Kautiainen H, Järvenpää S, Virta L, Pohjolainen T, Korpela M, et al. Decline in work disability caused by early rheumatoid arthritis: results from a nationwide Finnish register, 2000–8. Ann Rheum Dis 2013;72:672–7.
9. Holte HH, Tambs K, Bjerkedal T. Time trends in disability pensioning for rheumatoid arthritis, osteoarthritis and soft tissue rheumatism in Norway 1968-97. Scand J Public Health 2003;31:17–23.
10. Kim SY, Servi A, Polinski JM, Mogun H, Weinblatt ME, Katz JN, et al. Validation of rheumatoid arthritis diagnoses in health care utilization data. Arthritis Res Ther 2011;13:R32.
11. Maestes N, Mullen KJ, Strand A. Disability insurance and the Great Recession. Am Econ Rev 2015;105:177–82.
12. U.S. Bureau of Labor Statistics. Labor force statistics from the current beneficiary survey. URL: https://data.bls.gov/pdq/SurveyOutputServlet.
13. Allaire S, Wolfe F, Niu J, Zhang Y, Zhang B, LaValley M. Evaluation of the effect of anti–tumor necrosis factor agent use on rheumatoid arthritis work disability: the jury is still out. Arthritis Rheum 2008;59:1082–9.
14. Molina E, del Rincon I, Restrepo JF, Battafarano DF, Escalante A. Association of socioeconomic status with treatment delays, disease activity, joint damage, and disability in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2015;67:940–6.
15. Yelin E, Tonner C, Kim SC, Katz JN, Ayanian JZ, Brookhart MA, et al. Sociodemographic, disease, health system, and contextual factors affecting the initiation of biologic agents in rheumatoid arthritis: a longitudinal study. Arthritis Care Res (Hoboken) 2014;66:980–9.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1325–1331
DOI 10.1002/acr.24579

American College of Rheumatology
Empowering Rheumatology Professionals

Decision Needs and Preferred Strategies for Shared Decision-Making in Rheumatoid Arthritis: Perspectives of Canadian Urban Indigenous Women

Valerie Umaefulam, Terri-Lynn Fox, and Cheryl Barnabe

Objective. Decision-making for treatment of rheumatoid arthritis (RA) is complex, with multiple beneficial medication options available, but with the potential for treatment-related adverse effects and significant economic considerations. Indigenous patients make treatment decisions informed by an interplay of clinical, family, and societal factors. Shared decision-making may represent an approach to support treatment decisions in a culturally congruent manner. Our objective was to identify aspects of arthritis care that Indigenous participants found relevant for shared decision-making and to explore preferences for shared decision-making strategies.

Methods. A purposive sampling from rheumatology clinics that provide services to Indigenous patients in a Canadian urban center was used to recruit participants for interviews. Seven participants were recruited to reach content saturation. Interview content was coded by 2 individuals, including an Indigenous patient with RA, and the data were analyzed via thematic analysis.

Results. Participants were all women ages 37–61 years living with RA. Participants supported the idea that shared decision-making would be beneficial, primarily to support decisions around treatment plans and medication changes. Shared decision-making approaches would need to reflect Indigenous-specific content areas, such as benefits and risks of therapy informed by data from Indigenous patient populations and inclusion of traditional modes of healing. All participants were interested in having a decision coach and preferred that decision aids be in both paper and electronic formats for accessibility.

Conclusion. This study advances knowledge in the priority areas and specific content needed in the shared decision-making process and the preferences of shared decision-making strategies relevant and appropriate for urban Indigenous women living with RA in Canada.

INTRODUCTION

Rheumatoid arthritis (RA) is an autoimmune disease that causes inflammation and damage to joint tissues. Indigenous patients in Canada, including First Nations, Métis, and Inuit Peoples, experience a significant burden of RA, both in increased prevalence and in differential treatment outcomes relative to the general population (1). Colonization events and ongoing structural and interpersonal racism have created mistrust of the health care system, impacting access to care, and influencing decision-making approaches for treatment. In a prior study, we explored with Indigenous patients how they make treatment decisions. Their approach includes the use of nonpharmacologic methods, with pharmacologic decisions representing an interplay of clinical, family, and societal factors, including ease of access to medication options and fear of drug-dependency stigmatization (2). In other population groups facing arthritis inequities, shared decision-making is increasingly advocated as a mechanism to improve patient satisfaction and decision quality (3). Additionally, shared decision-making is highly valuable to employ when there are various medically reasonable options available, such as in current-day arthritis treatment (4,5). Shared decision-making, as used in this study, involves the health provider ensuring that patients are aware of available

Dr. Umaefulam's work was supported by Eyes High Postdoctoral Scholarship, University of Calgary. Dr. Barnabe's work was supported by a Canada Research Chair, Rheumatoid Arthritis and Autoimmune Diseases, Canadian Institutes of Health Research.

Valerie Umaefulam, PhD, Terri-Lynn Fox, PhD, Cheryl Barnabe, MD, MSc: Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada.

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

Address correspondence to Cheryl Barnabe, MD, MSc, Associate Professor, Department of Medicine and Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, 3330 Hospital Drive NW, Calgary, Alberta, Canada T2N 4N1. Email: ccbarnab@ucalgary.ca.

Submitted for publication November 16, 2020; accepted in revised form February 9, 2021.

SIGNIFICANCE & INNOVATIONS

- The study advances understanding of shared decision-making with urban Indigenous women living with rheumatoid arthritis in Canada.
- Shared decision-making approaches would need to reflect Indigenous-specific content areas, such as benefits and risks of therapy informed by data from Indigenous patient populations, inclusion of traditional modes of healing in available options, and medication cost coverage details.
- Urban Indigenous women with rheumatoid arthritis were interested in a shared decision-making strategy that involves having a decision coach and preferred that decision aids be in both paper and electronic formats for accessibility.

treatment options, providing detailed information about choices, discussing the patient's preferences, and making treatment decisions with patients (4).

Cultural belief systems influence perception of engagement in health care (6) and may result in varying interests in shared decision-making. Although several approaches are available to support shared decision-making in arthritis, there is limited exploration of the role of shared decision-making in arthritis care in the context of Indigenous peoples' health care. Further, few shared decision-making strategies have been specifically developed with and for Indigenous peoples (7,8). Thus, this study sought to identify whether Indigenous patients living in an urban center had an interest in participating in shared decision-making for RA care and which aspects of that care were relevant for shared decision-making, and we sought to explore preferences for shared decision-making strategies that could be employed.

PATIENTS AND METHODS

This study used a qualitative descriptive approach grounded in phenomenology to explore the perceptions of urban Indigenous participants on shared decision-making for arthritis care. To ensure that participants who had experienced the phenomenon of interest were included in the study (9), Indigenous patients with RA were recruited via a purposive sampling process from urban rheumatology clinics that provide services to Indigenous patients in Calgary, Alberta, Canada. Recruitment was facilitated via written formats such as posters and bookmarks. A physician, nurse, or medical assistant informed the patient of the study, and if the patient was interested in participating, introduced them to a research team member.

The Ottawa Decision Support Framework informed this study, which shows the interplay between addressing decision needs and providing the necessary support via the appropriate strategy (10). The framework constructs, including knowledge, values, clinical counseling, decision tools, and coaching (10) were integrated in the data collection instrument. Data were collected via semistructured interviews using an interview guide (Table 1). The interview guide was developed to explore patients' perceptions about the potential role of shared decision-making in arthritis care, priorities for application of shared decision-making, and preferences for shared decision-making strategies. The authors are all female Indigenous health researchers, VU possesses experience in qualitative approaches, TLF identifies as a First Nations person with lived experience of RA, and CB is a Métis rheumatologist. Participants

Table 1. Outline of interview guide

Scope	Questions
Personal narrative	Please tell me briefly about your experience living with inflammatory arthritis. What does shared decision-making mean to you? Prompts: physicians, influence of others. Have you experienced shared decision-making in health care? If yes, can you describe when and how this happened? Prompts: past experiences, importance of certain outcomes.
Shared decision-making priorities	What decisions in arthritis treatment do you feel you need assistance with, or health care providers should spend time discussing with you? Can you give some examples of decisions that you would like to be involved in when managing arthritis? Prompts: treatment plan, choice of medication, general information. What types of decision support do you need? Prompts: clarify decisions, provide facts, monitor progress?
Shared decision-making approach	There are various approaches that can be used in shared decision-making, such as using decision tools that can be completed online or by paper, or having a decision coach work with you in making decisions. Which approach would you prefer? Probes: Why would you prefer this approach? Would you like having a decision coach? Who would you like to be your coach? Prompts: friend, nurse, physician? If the shared decision-making strategy is developed, would you use the tool/strategy? When would you like to use this shared decision-making approach? Probes: Before or during arthritis treatment.

received an information brief containing a description of shared decision-making and a summary of the purpose of the study before the interview to allow for review and reflection. VU explained the reasons for the research and conducted the interviews. Each interview was approximately 45 minutes in duration and all were held in-person at medical clinics in Calgary. Nonparticipants were not present. Audio recordings were transcribed, and the transcripts were reviewed and coded by 2 individuals (VU and TF). The data were analyzed via inductive thematic analysis using NVivo software, based on thematic analysis phases, i.e., data immersion, initial coding, theme/category creation, and reviewing and refining themes (11). The data were organized into groups of meaning that represented the lived experiences of participants (9).

RESULTS

Seven participants were recruited from 3 rheumatology clinics to reach content saturation. They were all urban Indigenous women ages 37–61 years, living with RA. In summary and as described further in detail, shared decision-making was found to be acceptable for arthritis care by urban Indigenous women. Suggested priorities of support would be for treatment decisions informed by Indigenous population data, with inclusion of traditional and cultural treatment options in care plans, and reflecting available medication cost coverage options. Preferred shared decision-making strategies were those that included decision coaching and decision aid tools in both print and electronic formats. Figure 1 outlines the shared decision-making priorities and preferred approaches.

Role of shared decision-making for RA care. Although 2 participants indicated that they had not experienced shared decision-making before in health care, most participants noted that shared decision-making was not a new concept to their general health management. For instance, a participant indicated in regards to her family physician, "She asks me questions…like, 'What do you think…is the best decision?' and 'We can do this or we can do that,' and she asks me, if I want to do this or that" (patient 1). Another participant said, "My regular family doctor, he was very more understanding, and he showed me…and he sat down with me and we talked about my medications" (patient 7). Some participants detailed how not being involved in health decisions impacted their lives. In the words of a participant, "I trusted and blindly accepted the treatments that other doctors had given me prior to this, and then I started getting to a point where I felt like they weren't telling me what some of the drugs were. They'd just say, 'Here, take this,' and then I'd walk off and have side effects" (patient 3). A participant commented that the communication disconnect with physicians does not enhance or support shared decision-making and stated, "[The doctor] doesn't seem to have any time. I mean, he allows you maybe 15 minutes, and then he's on to the next patient. I'd like to at least know what's going on. Don't rush me out of the room. I'd like to know what I can do, what I can take" (patient 5).

Conversely, some participants indicated having some exposure to shared decision-making in their rheumatology specialty care experiences. For example, one participant said, "My doctor, she does the best she can to share with me about the different meds I'm on, some of the side effects, and she supports my decisions, and she provides her expertise in her area. That this could work, but this one seems to be working well, so she always leaves it up to me…I'm also active in that decision, and that's what I appreciate" (patient 2).

When reflecting on whether shared decision-making would have a role in arthritis care, participants voiced valuing active engagement in their care and were open to tools and opportunities to facilitate informed arthritis care decisions with their physicians. The underlying trust, respect, and relationship between the physician and patient would impact shared decision-making, and more importantly, empowerment, self-determination, and self-efficacy. A participant said, "We're a person, and you're a part of our health care team, and so treating someone with respect goes a long way. We feel good, and that helps. That helps in anyone's well-being… and the health care professionals with their gentle guidance and their respect and sharing what they do, I felt empowered" (patient 2). Overall, all participants disclosed that if a shared decision-making strategy was available for arthritis care, they would use it.

Shared decision-making priorities in RA care. Given that interest in shared decision-making was supported by the participants, they were then asked what priority areas would be for its application. Two areas emerged: the identification of specific content needed in the shared decision-making process and the type of decision support desired.

Decision support priorities. Participants indicated that shared decision-making would help to support decisions around RA treatment plans and medication changes. To illustrate, a participant was interested in knowing, "if I can come off biologics and just try methotrexate. Like, how long can I be on it [biologics] before it is like, harming me, or if I stay on it for the rest of my life, is it going to…damage me more?" (patient 1). Further, another participant stated, "Well, the treatment plan, I think that they should all discuss a survey that was done…why they feel that this treatment plan would be good for you" (patient 3).

All participants expressed the need for extensive knowledge of available treatments for arthritis and the side effects of drugs prescribed. They were all concerned about the impact of the drugs on the body, whether the drugs were hurting them, and when is the best time to change medications. A participant noted, "Someone to clarify what the medications do to the body and I feel like everybody should have a right to know what they ingest or they put in their body, especially through doctors and therapists and stuff like that" (patient 3).

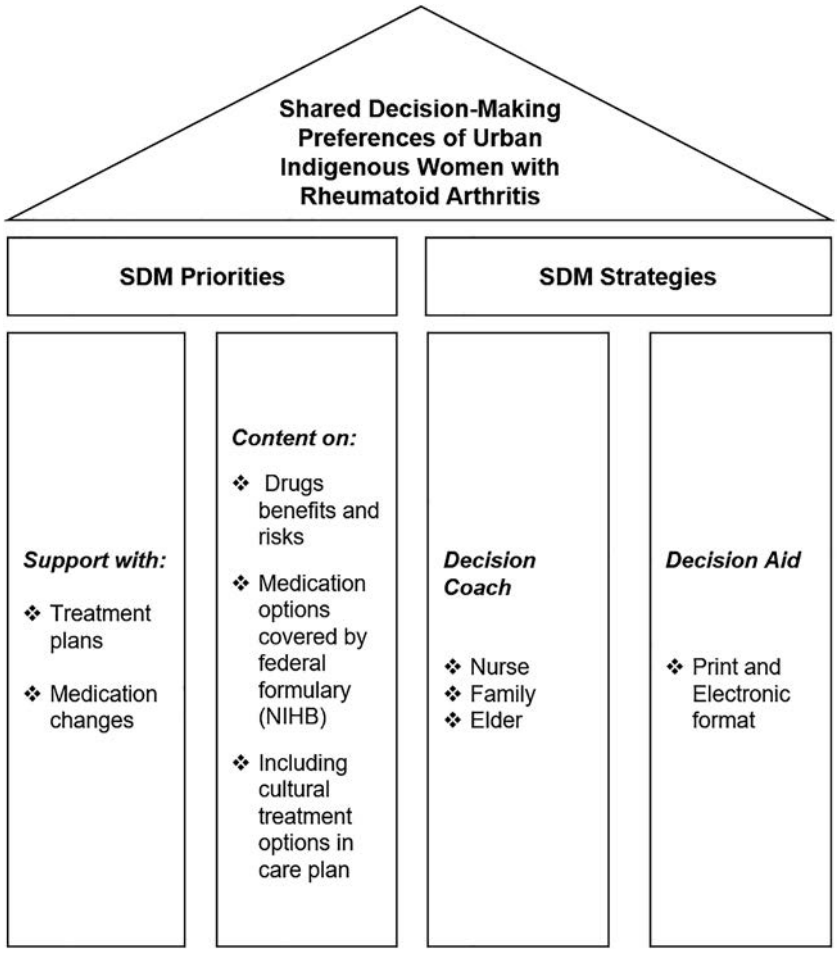


Figure 1. Priorities and preferred strategies for shared decision-making. NIHB = Non-Insured Health Benefits; SDM = shared decision-making.

Specific shared decision-making content areas for Indigenous patients. Some participants said that shared decision-making strategies should incorporate evidence from Indigenous population or patient studies, particularly regarding medications and their effects, which would assist in motivating patients in adhering to treatment plans. One participant noted, "I think there are specific data, statistics…not just general population, but maybe specific Indigenous-population data…because there's so many of us that are affected in such a short span of time" (patient 2). Another participant emphasized drawing shared decision-making information from Indigenous communities, stating, "There's just a lot of information here" (patient 5).

Another aspect of specific shared decision-making content was to include options related to traditional modes of healing in care plans. These modes of healing support not only physical wellness, but also spiritual, emotional, and mental wellness. Two participants mentioned the benefits of traditional practices such as prayers and sweats on health care, and suggested including traditional medicine practices and integrating Elders in care. A participant stated, "There's an Elder here. Perhaps it could be beneficial to meet with him…that might have a holistic approach to some kind of salve or herbs or something. You know, a lot of that wisdom and knowledge is lost today, just not been documented, and I think it would be very beneficial, instead of using all the pills or whatever the doctor wants to prescribe" (patient 5). First Nations participants with Treaty Status suggested that decision tools provide information on which medication options would be covered by the federal formulary (Non-Insured Health Benefits) and which medications would not be, as this knowledge would impact the ability to obtain medication offered in the options.

Preferred shared decision-making strategies. Participants discussed various types of shared decision-making approaches that could be used for arthritis care. They were open to using an electronic format of shared decision-making decision aids due to its convenience and the ease of obtaining information. As one person said, "I always Google stuff if I'm unsure. Yeah, so I'm always on the computer" (patient 4). Nevertheless, patients equally acknowledged the difficulties of accessing information via devices such as computers and cell phones among individuals with limited or no accessibility. Given this fact, most participants

preferred a shared decision-making decision aid that used a combination of paper and electronic resources. One participant stated, "So, online support. And I'm saying that just for me, because I find myself…I'm privileged to be able to have a cell phone, to be able to have Internet access at home. However, there's a majority of people on reserve that use landlines. They may not have a cell phone. They may not have access to Internet. So, paper resources" (patient 2). Another recognized the advantage of paper resources: "You can lose papers and documents, but sometimes, for some people, they keep onto that, so you know you have something to go back to, a hard copy" (patient 6).

Moreover, participants suggested that the shared decision-making strategy should incorporate Indigenous language translations, to be relevant specifically to older individuals who maintained their own language. Another suggestion was that text associated with shared decision-making decision aids should be simplified for better comprehension and include appropriate Indigenous symbols and images to demonstrate respect for Indigenous culture. According to a participant, "Translating it as best as we could, and utilizing, if it's a visual, utilizing Indigenous symbols so that it's meaningful to patients…so that we know you care about who we are as part of this culture, or this land area, and you're further respecting and responding to our cultural needs as well" (patient 2).

There was general interest in a shared decision-making approach that involved having a decision coach insofar as having an individual who would provide one-on-one interactions, relate with patients outside the regular clinic visits, respond to inquiries, and have knowledge of the arthritis condition and treatment options. As explained by one participant, "I would like to have somebody, talk to me and explain to me, my options about my medications. And like I said, to find out what's good and what's bad…what it does to my body and all this" (patient 7). Another participant stated, "I think that [decision coach] is a really good idea. I think someone who knows the industry, who knows the medication, who understands those who have been affected by extreme excruciating pain" (patient 2). There was resounding agreement for decision coaches to have sufficient time to discuss treatment options, reflecting the fact that time with the physician is often limited to enable extensive discussions. There were differing suggestions on who a decision coach might be. Some suggested nurses, as they work closely with physicians: "I feel like they [physicians] have a lot on their plate…I feel like nurses would have that extra time to sit, well, in between patients, right? To talk, to have the one-on-one" (patient 3).

One participant articulated the importance of the decision coach having lived experience of arthritis and welcomed having a family member as a decision coach. She trusted that a family member would have her interest at heart and provide appropriate guidance in making decisions and commented, "Even my grandmother… She said something to me 2 weeks ago the last time I'd seen her, 'Well, you have to do your exercises. There're certain exercises you have to do. I know, because I'm speaking from experience. I don't care what your doctor said. Listen to me.' And of course, I'm going to listen to an 84-year-old woman. She knows" (patient 6). Several participants noted the importance of also having a decision coach who would support holistic approaches for health.

One suggestion was to have an Elder work in collaboration with the decision coach to facilitate shared decision-making; this was expressed by one participant, "It would be new. And I would say an Elder, but an Elder may not have…specific drug, pharmacy knowledge, or how it works in your body, what the side effects could do. So it could be a combination of a doctor or, you know, someone who works closely with RA specialists, as well as an Elder who would pray and maybe translate for other members, but who would support that coach or that specialist to do it in a holistic way" (patient 2).

As a check for decision alignment, participants noted the importance of introducing the strategy before and after changing a treatment plan: "Well, I feel like you especially should start the regime…there could be that nurse that I could call and talk to, saying that, you know, 'This is where I'm at with it, and I want to know, is this where I'm supposed to be?'" (patient 3).

DISCUSSION

Our study explored whether there is a potential role for shared decision-making in arthritis care of Indigenous patients. Racial, ethnic, and cultural minority groups are more vulnerable to poor decision-making outcomes, with a majority reporting low decision satisfaction and high decision regret (6). As a framework, shared decision-making can be used to communicate with patients about health care choices and has been shown to be beneficial to patient engagement and treatment outcomes by increasing adherence to treatment plans (12). Active involvement in health care decisions promotes self-determination, especially when such decisions involve personal preferences. Urban Indigenous women with RA indicated they wanted to be involved in making decisions regarding their health and articulated the importance of health care providers working with them to ensure that they have the information they need to make complex decisions (13). This approach also supports reconciliation; the Truth and Reconciliation Commission of Canada (TRCC) provided directives to facilitate actions to be taken in Canadian society to promote healing of and reconciliation with persons directly or indirectly affected by the Indian Residential Schools system legacy (14). We propose that for a shared decision-making approach with Indigenous patients to be effective, clinicians must respond to the TRCC Calls to Action in Health (15). These include recognizing, respecting, and addressing the distinct health needs of all Indigenous peoples (Call to Action #20); ensuring that they gain training in cultural awareness and safety (Call to Action #23); understanding Indigenous health issues (Call to Action #24);

promoting and supporting Indigenous health approaches (Call to Action #22); and moving forward with reconciliatory practices in all areas and close gaps in health outcomes (Call to Action #19). Understanding Indigenous people from a historical-cultural perspective is needed, especially when striving toward health and well-being. Respecting patient perspectives and forming trustful relationships is an intentional and positive approach for shared decision-making.

Indigenous participants said that shared decision-making would be useful primarily to support decisions around treatment plans and medication changes. Patients do not always know the side effects of medications and require additional information, clear communication, and better understanding of medications. For instance, many patients suffer from medication side effects (16), rather than the actual pain and immobility of RA. Enhancing knowledge related to care may lead to medication choices better aligned with patient values and preferences in arthritis treatment (17). As patients develop a trusting relationship with their rheumatologist and educate themselves on RA and its available treatments, they gain confidence and pursue a mutual role via shared decision-making in their treatment decisions (18).

Specific content and information pertaining to Indigenous population realities were regarded as essential, particularly treatment benefits and risks unique to them, which treatment options have available medication cost coverage, and how cultural components of health could be considered in the treatment options presented. As health systems are increasingly burdened during this unprecedented time, we are reminded of the necessity of holistic health approaches, especially Elder knowledge and support, traditional herbs and medicines, and ceremonial practices and rituals that may relieve and decrease the amount of pain and discomfort felt by Indigenous patients with arthritis. Intergenerational knowledge, "blood memory" of traditional medicines and methods, is inherent in many First Nations communities, who desire to return to traditional ways. Since many Indigenous people long for traditional knowledge, particularly survivors of Indian Residential School and the Sixties Scoop, shared decision-making would be useful to support decisions around the inclusion of traditional modes of healing in care plans, especially the physiologic and spiritual aspects.

Decision support tools include decision aids that describe treatment options and their benefits and harms, and such aids may equally include a guide to decision-making (12). The tool may be web-based or printed material such as pamphlets or videos that assist patients in considering treatment options and outcomes, which proceed through the steps of deliberation and communication with the patients' health care provider. For example, in a pilot study among RA patients, patients who used the RA Choice (a print-based resource) with their doctor reported significantly improved knowledge and low decisional conflict compared to those who did not use the tool (17). The study participants reported valuing combined electronic and print shared decision-making aids that were user-friendly, attentive to health literacy challenges by applying plain language terms, and translated as needed in different Indigenous languages.

The need for conversation and accountability underlies the request for including a decision coach in the shared decision-making strategy. Decision coaching is a process that prepares patients to discuss options with their providers facilitated by a decision coach who may be a nurse, social worker, or other allied health professionals (19). Decision coaches may use decision support tools to guide the patient in deliberating about options with their health practitioner (19). In this study, the participants' suggestions for a nurse, a family member, an Elder, or a community member with a lived experience with arthritis as a decision coach points to the importance of relationship, connection with tribal members, the closeness of families, role models, and the social interaction that helps mold and guide quality of life for Indigenous people. Decision coaching may improve knowledge and increase the perceived involvement in decision-making and satisfaction with the decision-making process (20).

We are aware of prior studies that have explored shared decision-making with Indigenous populations. The Ottawa Decision Support Framework was culturally adapted to produce a tool that better met the needs of Indigenous peoples (21) and enabled shared decision-making in Western health care settings. While aiming to refine the decision tool, the study also revealed that decision coaching was required to increase engagement in the decision-making process while using the adapted framework as a talking guide (21). Also, a shared decision-making strategy called "Not Deciding Alone," developed for use by Inuit peoples in cancer care in Canada (7), included community support workers who provided peer support and facilitated the use of the tool (a booklet).

As our study was initiated in the fall of 2019, we were only able to recruit participants from the urban setting prior to the COVID-19 pandemic restrictions being enacted. Few participants were recruited, but the sample size is suitable for a phenomenologic study, which often needs 1–10 participants (9) and was sufficient to reach saturation. Participants were recruited from 3 rheumatology clinics in Calgary, and they received care from different physicians. The pandemic also limited our research population, such that individuals living in rural locations could not be included at this time. We did not include the perception of Indigenous men. Also, patients recruited to the study were actively engaged in Western health care systems and receptive to research participation; this demographic raises the possibility that the voices of Indigenous persons who have chosen to not interact with providers and researchers are not included. Future research that involves gaining insight from these groups could provide more exhaustive information on priorities and preferences for shared decision-making in RA care for Indigenous populations as we proceed to adapt and test the acceptability and effectiveness of a decision aid with this population.

Indigenous patients appreciate active engagement in decision-making for arthritis care. Our study sheds light on the importance of providing culturally safe health care practices with Indigenous patients in the health system when incorporating shared decision-making strategies. This study advances knowledge in the priority areas and specific content needed in the shared decision-making process and the preferences of shared decision-making strategies relevant and appropriate for Indigenous patients living with RA.

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. Barnabe 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. Umaefulam, Barnabe.
Acquisition of data. Umaefulam, Fox, Barnabe.
Analysis and interpretation of data. Umaefulam, Fox, Barnabe.

REFERENCES

1. Ng C, Kue Young T, Chatwood S. Arthritis in the Canadian aboriginal population: north-south differences in prevalence and correlates. Prev Chronic Dis 2011;8:A04.
2. Loyola-Sanchez A, Hazlewood G, Crowshoe L, Linkert T, Hull PM, Marshall D, et al. Qualitative study of treatment preferences for rheumatoid arthritis and pharmacotherapy acceptance: indigenous patient perspectives. Arthritis Care Res (Hoboken) 2020;72:544–52.
3. Barton JL, Trupin L, Tonner C, Imboden J, Katz P, Schillinger D, et al. English language proficiency, health literacy, and trust in physician are associated with shared decision making in rheumatoid arthritis. J Rheumatol 2014;41:1290–7.
4. Stiggelbout AM, Pieterse AH, De Haes JC. Shared decision making: concepts, evidence, and practice. Patient Educ Couns 2015;98:1172–9.
5. Li LC, Adam PM, Backman CL, Lineker S, Jones CA, Lacaille D, et al. Proof-of-concept study of a web-based methotrexate decision aid for patients with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2014;66:1472–81.
6. Hawley ST, Morris AM. Cultural challenges to engaging patients in shared decision making. Patient Educ Couns 2017;100:18–24.
7. Jull J, Hizaka A, Sheppard AJ, Kewayosh A, Doering P, Macleod L, et al. An integrated knowledge translation approach to develop a shared decision-making strategy for use by Inuit in cancer care: a qualitative study. Curr Oncol 2019;26:1–27.
8. Jull J, Giles A, Boyer Y, Lodge M, Stacey D. Shared decision making with aboriginal women facing health decisions: a qualitative study identifying needs, supports, and barriers. Altern An Int J Indig Peoples 2015;11:401–16.
9. Starks H, Trinidad SB. Choose your method: A comparison of phenomenology, discourse analysis, and grounded theory. Qual Health Res 2007;17:1372–80.
10. Stacey D, Légaré F, Boland L, Lewis KB, Loiselle MC, Hoefel L, et al. 20th anniversary Ottawa decision support framework: part 3 overview of systematic reviews and updated framework. Med Decis Mak 2020;40:379–98.
11. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16:1–13.
12. Hoffmann TC, Légaré F, Simmons MB, McNamara K, McCaffery K, Trevena LJ, et al. Shared decision making: what do clinicians need to know and why should they bother? Med J Aust 2014;201:35–9.
13. Deber RB, Kraetschmer N, Urowitz S, Sharpe N. Do people want to be autonomous patients? Preferred roles in treatment decision-making in several patient populations. Health Expect 2007;10:248–58.
14. Government of Canada. Truth and Reconciliation Commission of Canada. 2020. URL: https://www.rcaanc-cirnac.gc.ca/eng/1450124405592/1529106060525.
15. National Centre for Truth and Reconciliation. Calls to action. Winnipeg, Manitoba; 2015. URL: https://nctr.ca/records/reports/.
16. Curtis JR, Xie F, MacKey D, Gerber N, Bharat A, Beukelman T, et al. Patient's experience with subcutaneous and oral methotrexate for the treatment of rheumatoid arthritis. BMC Musculoskelet Disord 2016;17:1–11.
17. Barton JL, Trupin L, Schillinger D, Evans-Young G, Imboden J, Montori VM, et al. Use of low-literacy decision aid to enhance knowledge and reduce decisional conflict among a diverse population of adults with rheumatoid arthritis: results of a pilot study. Arthritis Care Res (Hoboken) 2016;68:889–98.
18. Mathews AL, Coleska A, Burns PB, Chung KC. Evolution of patient decision-making regarding medical treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:318–24.
19. Stacey D, Kryworuchko J, Bennett C, Murray MA, Mullan S, Légaré F. Decision coaching to prepare patients for making health decisions. Med Decis Mak 2012;32:E22–33.
20. Dawn S, Légaré F. Engaging patients using an interprofessional approach to shared decision making. Can Oncol Nurs J 2015;25:455–69.
21. Jull J, Giles A, Minwaashin Lodge, Boyer Y, Stacey D. Cultural adaptation of a shared decision making tool with aboriginal women: a qualitative study. BMC Med Inform Decis Mak 2015;15:1–13.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1332–1341
DOI 10.1002/acr.24865

Differences in the Association Between Oral Glucocorticoids and Risk of Preterm Birth by Data Source: Reconciling the Results

Kristin Palmsten,[1] Gretchen Bandoli,[2] Gabriela Vazquez-Benitez,[3] and Christina D. Chambers[2]

Objective. To investigate causes of discrepancies in the association between early pregnancy oral glucocorticoid (OGC) use and preterm birth risk among women with rheumatoid arthritis (RA) in health care utilization data from California Medicaid (Medi-Cal) and the prospective cohort MotherToBaby Pregnancy Studies.

Methods. Separately, we estimated risk ratios (RRs) between OGC exposure before gestational day 140 and preterm birth risk in data from Medi-Cal (2007–2013; n = 844) and MotherToBaby (2003–2014; n = 528). We explored differences in socioeconomic status, OGC dose distribution, exposure misclassification, and confounding by RA severity across the data sources.

Results. Preterm birth risk in women without OGC was 17.3% in Medi-Cal and was 9.7% in MotherToBaby. There was no association between OGC and preterm birth in Medi-Cal (adjusted RR 1.00 [95% confidence interval (95% CI) 0.71, 1.42]), and a 1.85-fold (95% CI 1.20, 2.84) increased preterm birth risk in MotherToBaby. When restricting each sample to women with a high-school diploma or less, preterm birth risk following no OGC exposure was 15.9% in Medi-Cal and 16.7% in MotherToBaby; adjusted RRs were 1.16 (95% CI 0.74, 1.80) in Medi-Cal and 0.81 (95% CI 0.25, 2.64) in MotherToBaby. Cumulative OGC dose was higher in MotherToBaby (median 684 mg) than in Medi-Cal (median 300 mg). An OGC dose of ≤300 mg was not associated with increased preterm birth risk. Exposure misclassification and confounding by RA severity were unlikely explanations of differences.

Conclusion. Higher baseline preterm birth risk and lower OGC dose distribution in Medi-Cal may explain the discrepancies. Studies are needed to understand the effects of autoimmune disease severity and undertreatment on preterm birth risk in low-income populations.

INTRODUCTION

Oral glucocorticoids (OGCs) may be used to manage flares/exacerbations or for chronic management of autoimmune diseases, including rheumatoid arthritis (RA) during pregnancy (1–6). Prospective cohort and health care utilization database studies have reported an increased risk of preterm birth following OGC use in women with RA and other autoimmune diseases (7–12).

The use of health care utilization databases to study medication safety during pregnancy is becoming increasingly common (13). These preexisting data sources can increase feasibility and efficiency of studying relatively rare exposure and perinatal outcomes while reducing costs compared with primary data collection (13,14). However, these data are not collected for research, and misclassification, unmeasured confounding, and restriction to patients with public or employer-based insurance are concerns (13,14). Well-designed prospective studies can collect detailed information, but participants may differ from the target population of interest. It is unclear whether these differences related to internal validity (i.e., bias) and external validity (i.e., the extent to which results from the study sample hold for the population of interest) limit comparisons of medication safety in pregnancy between the 2 types of data sources.

When studies of the same perinatal medication safety question are available from health care utilization and prospective

MotherToBaby Pregnancy Studies have been funded by research grants from AbbVie, Amgen, Apotex, Barr, Bristol-Myers Squibb, Par, Kali, Sandoz, Teva, Roche/Genentech, GlaxoSmithKline, UCB, Pfizer, Janssen, Celgene, Regeneron, Takeda, and Sanofi-Aventis/Genzyme. Dr. Palmsten's work was supported by a career development award from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH (R00-HD-082412). Dr. Bandoli's work was supported by the National Institute on Alcohol Abuse and Alcoholism, NIH (K01-AA-027811).

[1]Kristin Palmsten, ScD: HealthPartners Institute, Minneapolis, Minnesota, and University of California San Diego, La Jolla; [2]Gretchen Bandoli, PhD, Christina D. Chambers, PhD, MPH: University of California San Diego, La Jolla; [3]Gabriela Vazquez-Benitez, PhD, MSc: HealthPartners Institute, Minneapolis, Minnesota.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24865&file=acr24865-sup-0001-Disclosureform.pdf.

Address correspondence to Kristin Palmsten, ScD, HealthPartners Institute, Minneapolis, MN 55425. Email: Kristin.K.Palmsten@HealthPartners.com.

Submitted for publication December 2, 2021; accepted in revised form January 25, 2022.

SIGNIFICANCE & INNOVATIONS

- Oral glucocorticoid (OGC) exposure was not associated with preterm birth when restricting to women with lower education, but women with lower education had a high baseline risk of preterm birth of approximately 16%.
- Differing results from studies of perinatal medication safety, including studies of OGC use and preterm birth risk, may stem from differences in study populations, baseline risks, and dose distributions in addition to the typical sources of bias in observational studies (e.g., exposure misclassification, outcome misclassification, and confounding).

cohort data, results should be purposely compared, as threats to internal and external validity across data sources make clinical interpretation problematic. Therefore, we aimed to examine the same research question using both types of data, namely to what extent does OGC use early in pregnancy affect the risk of preterm birth among women with RA? We studied the association in a prospective cohort of women from the MotherToBaby Pregnancy Studies and among women enrolled in the California Medicaid Program (known as Medi-Cal). We conceptualized the target population, i.e., the population of interest, for our study question as pregnant women in the US who have RA, recognizing that our study samples were not drawn at random from this population. Instead, women enrolled in MotherToBaby primarily had higher socioeconomic status (SES), whereas women enrolled in Medi-Cal had lower SES because Medi-Cal is the joint state and US federal health insurance program for low-income individuals.

We found differing results from the 2 studies and explored potential reasons for the differences. To assess whether issues with internal validity could explain the discrepant results, we explored the potential for exposure misclassification (i.e., incorrectly classifying whether or not women used OGC) and residual confounding by RA severity within the 2 studies. We did not investigate outcome misclassification as an explanation for the observed differences because both studies used similar approaches to estimate gestational age at delivery, i.e., primarily ultrasound measurements with correction for discrepancies. Furthermore, we did not evaluate selection bias due to pregnancy loss because OGC use is not expected to increase the risk of pregnancy loss (15). To assess whether issues related to external validity could explain the discrepant results, we explored differences in SES, a potential effect modifier, and cumulative OGC dose distribution across the study samples.

PATIENTS AND METHODS

We previously conducted related studies on OGC exposure during pregnancy and the risk of preterm birth in both data sources, and the methods have been described in detail (12,16). We aimed to make the current analyses as similar as possible, given differences in data elements across the data sources. To simplify the current analyses and comparisons across studies, we focused on any OGC exposure during the first half of pregnancy. Furthermore, we limited the study populations to women with RA to reduce potential confounding by underlying disease. The Medi-Cal study was approved by the Committee for the Protection of Human Subjects, California Health and Human Services Agency, and was determined exempt by the University of California San Diego Human Research Protections Program. A data use agreement was in place with the California Department of Health Care Services. Counts of <16 were suppressed. The MotherToBaby Pregnancy Studies were approved by the University of California, San Diego Institutional Review Board and the current analysis was exempt. Informed consent was obtained in the MotherToBaby Pregnancy Studies.

We used 2007–2013 Medi-Cal enrollment and outpatient, inpatient, and pharmacy claims data linked to birth certificate and hospital discharge data for women with a live birth, continuous Medi-Cal enrollment during pregnancy, and an inpatient or outpatient International Classification of Diseases, Ninth Revision diagnosis code for RA (714.x) during pregnancy (n = 844). Women were classified as exposed to OGC if they had a pharmacy dispensing for any OGC between the last menstrual period (LMP) date and gestational day 139 (i.e., 20 gestational weeks). Gestational age at delivery was primarily determined from the birth certificate obstetric estimate (17). The LMP date was calculated from the birth certificate by subtracting the obstetric estimate of gestational age at delivery from the delivery date. Alternatively the birth certificate LMP estimate of gestational age at delivery was used when the obstetric estimate was unavailable (17), though most women (84%) meeting the inclusion criteria had the obstetric estimate available.

MotherToBaby Pregnancy Studies conduct prospective cohorts of several diseases and exposures during pregnancy, enrolling pregnant women in the US and Canada (18,19). MotherToBaby participants were self-referred, referred by health care providers, or referred by MotherToBaby, a free service of the Organization of Teratology Information Specialists providing evidence-based information on exposures in pregnancy and lactation (18,19). We included pregnant women with a live birth or stillbirth who enrolled in the MotherToBaby Autoimmune Diseases in Pregnancy Study (2003–2014) before gestational day 140 and reported having RA, and we excluded women missing information on RA-related medications, including OGC (n = 9) (12). A total of 528 women met the eligibility criteria, including 1 with stillbirth. Trained study staff conducted up to 4 semistructured telephone interviews with participants: at enrollment (before gestational week 20), at approximately 24 and 32 weeks' gestation, and after delivery. Interviewers collected data on demographics, reproductive history, prepregnancy weight and height, comorbidities, smoking, and pregnancy outcomes (20).

At study enrollment, interviewers used an open-ended prompt to obtain information on medication use such as "Have you taken

any over-the-counter medications since your last menstrual period, for example, Tylenol or Tums?" Women who reported having a specific illness or disease were asked if they took any medication for that condition, e.g., RA. For all medications reported, women were queried about dose and dates of use. During follow-up interviews, women were queried about medication use since their most recent interview and whether they were using previously reported medications (20). Interviewers administered self-assessment questionnaires to measure RA severity, including the Health Assessment Questionnaire (HAQ) disability index (a validated measure of functional status in patients with RA; range 0 = no disability, 3 = completely disabled) (21,22), pain score (pain severity rating in the past week; range 0 = no pain, 100 = severe pain), and patient's global score (overall health rating; range 0 = very well, 100 = very poor). Gestational age at delivery was estimated from the LMP date with adjustment for discrepant ultrasound measurements. Women were classified as exposed if they reported any OGC use between the LMP date and gestational day 139.

Preterm birth was classified as delivery at <37 gestational weeks, i.e., <259 days. We assessed the association between any OGC exposure and preterm birth using Poisson regression with robust variance to estimate risk ratios (RRs) and 95% confidence intervals (95% CIs) (23). We also estimated risk differences (RDs) and 95% CIs using linear regression with robust variance. We identified covariates a priori that we hypothesized to be potential confounders. We adjusted estimates for a common set of covariates available in both data sources and additional covariates unique to both data sources to further address confounding. The modeling approaches (i.e., functional form, categorization cut points) are described in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24865.

The common covariates were LMP year (<2010, ≥2010; cut point approximately halfway through the years of data for Medi-Cal), maternal age, race/ethnicity, maternal education, multiple gestation, prepregnancy body mass index, primiparity, hypertension, autoimmune comorbidities, and, as proxies of RA severity, nonsteroidal antiinflammatory drugs and disease-modifying antirheumatic drugs (DMARDs), including conventional and biologic therapies, between the LMP and day 139. In Medi-Cal, we also adjusted for being in the disability category for Medicaid eligibility, general markers of comorbidity (any hospitalizations, number of outpatient and emergency department visits) (24), and disease severity proxies between LMP and day 139 (the number of outpatient visits with RA diagnosis and inflammatory marker and rheumatoid factor laboratory results) (25). In MotherToBaby, we also adjusted for SES using Hollingshead categories of maternal and paternal education (26) to further account for confounding by SES, and we adjusted for HAQ, pain, and global scores at the time of enrollment to adjust for RA severity. In post hoc analyses, we explored the 4 factors that we expected to differ between the 2 data sources as potential explanations of discrepant results, discussed below.

SES. Pregnant women enrolled in Medi-Cal meet low-income thresholds, whereas MotherToBaby participants primarily have higher SES (12). We used education as an SES proxy because it was measured in both studies. To make the 2 populations more similar, we restricted to high-school diploma or less education in both data sources. We present characteristics for the restricted and full study populations and estimated the associations in the restricted population.

Dose. Higher OGC doses have been associated with higher preterm birth risk in both data sources (12,15). Therefore, differences in typical OGC dose across studies could contribute to differences in the association between any OGC use and preterm birth. We assessed the median total cumulative OGC dose (prednisone equivalent dose) (27) between the LMP and gestational day 139. Then we assessed the association between high and low OGC cumulative dose versus no OGC exposure between the LMP and day 139 and preterm birth, using the lower of 2 median doses as the exposure cut point.

Exposure misclassification. We anticipated greater exposure misclassification in Medi-Cal than in MotherToBaby because we could not confirm OGC use as assumed from dispensing data. Previously using MotherToBaby data, we compared prednisone use (the most common OGC during pregnancy) (28,29) in medical records versus maternal report during pregnancy in women with RA and found a sensitivity of 56% (95% CI 47, 64) and specificity of 89% (95% CI 82, 94) (20). We expect a similar degree or less exposure misclassification in claims data versus the medical records in our previous study. This is because the claims were comprised of pharmacy dispensing data, whereas the data from the medical records were from medication orders, not fills, and from active medication lists and physicians' notes, which generally required a health care visit for updates/reconciliation. We conducted a probabilistic bias analysis of exposure misclassification, simultaneously adjusting for measured covariates using the approach and macro described by Fox et al (30) to assess the degree of misclassification needed in Medi-Cal to produce the same adjusted RR in MotherToBaby (details in Supplementary Methods in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24865).

Residual confounding. Given that MotherToBaby collected validated self-reported measures of RA severity (21,22) and that we had to rely on proxies of disease severity in Medi-Cal, we were particularly concerned about confounding by RA severity in the Medi-Cal analysis. We anticipate that residual confounding by RA severity would lead to upward bias, as greater disease severity is associated with an increased risk of preterm birth (19) and greater disease severity is associated with OGC use (as observed in MotherToBaby in Table 1). We compared fully adjusted models to those without severity adjustment. Also, we conducted a bias analysis for

Table 1. Characteristics among women in the Medi-Cal study and in the MotherToBaby (MTB) study overall and by exposure status*

Characteristic	Medi-Cal			MTB		
	Overall (n = 844)	No OGC (n = 655)	OGC (n = 189)	Overall (n = 528)	No OGC (n = 269)	OGC (n = 259)
LMP year 2010 or later	556 (65.9)	430 (65.6)	126 (66.7)	219 (41.5)	124 (46.1)	95 (36.7)
Maternal age, median (IQR)	29.0 (9.0)	28 (9)	30 (9)	32 (6)	32 (6)	32 (6)
Race/ethnicity						
White (Medi-Cal), non-Hispanic White (MTB)	156 (18.5)	133 (20.3)	23 (12.2)	417 (79.0)	221 (82.2)	196 (75.7)
Black	78 (9.2)	62 (9.5)	16 (8.5)	NA	NA	NA
Hispanic	543 (64.3)	411 (62.7)	132 (69.8)	NA	NA	NA
Other or unknown	67 (7.9)	49 (7.5)	18 (9.5)	111 (21.0)	48 (17.8)	63 (24.3)
Maternal education						
High school/equivalent or less	577 (68.4)	452 (69.0)	125 (66.1)	50 (9.5)	24 (7.9)	26 (10.0)
Less than high school	290 (34.4)	242 (37.0)	48 (25.4)	NA	NA	NA
High school/equivalent	287 (34.0)	210 (32.1)	77 (40.7)	NA	NA	NA
Some college, college degree/equivalent, or higher	267 (31.6)	203 (31.0)	64 (33.9)	478 (90.5)	245 (92.1)	233 (90.0)
Some college	NA	NA	NA	116 (22.0)	55 (20.4)	61 (23.6)
College degree or higher	NA	NA	NA	362 (68.6)	190 (70.6)	172 (66.4)
Disability as source of Medi-Cal eligibility	113 (13.4)	80 (12.2)	33 (17.5)	NA	NA	NA
Socioeconomic status, median (IQR)†	NA	NA	NA	2 (1)	2 (1)	2 (1)
Gestational age at enrollment, median (IQR)	NA	NA	NA	11 (7)	11 (7)	11 (8.)
Multiple gestation	–	–	–	23 (4.4)	8 (3.0)	15 (5.8)
Prepregnancy body mass index						
Underweight to normal weight	317 (37.6)	247 (37.7)	70 (37.0)	324 (61.4)	156 (58.0)	168 (64.9)
Overweight	253 (30.0)	200 (30.5)	53 (28.0)	114 (21.6)	55 (20.4)	59 (22.8)
Obese	274 (32.5)	208 (31.8)	66 (34.9)	90 (17.0)	58 (21.6)	32 (12.4)
Primiparous	198 (23.5)	167 (25.5)	31 (16.4)	249 (47.2)	126 (46.8)	123 (47.5)
Autoimmune comorbidities‡	83 (9.8)	53 (8.1)	30 (15.9)	29 (5.5)	12 (4.5)	17 (6.6)
Hypertension (Medi-Cal)/prepregnancy hypertension (MTB)§	33. (3.9)	21 (3.2)	–	37 (7.0)	18 (6.7)	19 (7.3)
DMARD§	148 (17.5)	75 (11.5)	73 (38.6)	421 (79.7)	213 (79.2)	208 (80.3)
NSAID¶	171 (20.3)	100 (15.3)	71 (37.6)	191 (36.2)	85 (31.6)	106 (40.9)
Number of outpatient visits§						
None	292 (34.6)	227 (34.7)	65 (34.4)	NA	NA	NA
1 to 5	330 (39.1)	265 (40.5)	65 (34.4)	NA	NA	NA
≥6	222 (26.3)	163 (24.9)	59 (31.2)	NA	NA	NA
Number of emergency department visits§						
None	510 (60.4)	408 (62.3)	102 (54.0)	NA	NA	NA
1	131 (15.5)	101 (15.4)	30 (15.9)	NA	NA	NA
≥2	203 (24.1)	146 (22.3)	57 (30.2)	NA	NA	NA
Inpatient admission ≥1§	26 (3.1)	–	–	NA	NA	NA
Number of outpatient visits with RA diagnosis§						
None	592 (70.1)	475 (72.5)	117 (61.9)	NA	NA	NA
1	157 (18.6)	121 (18.5)	36 (19.1)	NA	NA	NA
≥2	95 (11.3)	59 (9.0)	36 (19.0)	NA	NA	NA
Inflammatory marker§	128 (15.1)	82 (12.5)	46 (24.3)	NA	NA	NA
Rheumatoid factor§	82 (9.7)	65 (9.9)	17 (9.0)	NA	NA	NA
HAQ at enrollment, median (IQR)#	NA	NA	NA	0.3 (0.9)	0.1 (0.6)	0.5 (1.0)
Pain score at enrollment, median (IQR)**	NA	NA	NA	20.0 (45)	10 (25)	25 (50)
Global score at enrollment, median (IQR)††	NA	NA	NA	15.0 (35)	10 (30)	25 (45)
OGC cumulative dose, median (IQR)‡‡	300 (420)	0 (0)	300 (420)	684 (645)	0	684 (645)

* Values are the number (%) unless indicated otherwise. Counts of <16 were suppressed for Medi-Cal. DMARD = disease-modifying antirheumatic drug; HAQ = Health Assessment Questionnaire; IQR = interquartile range; LMP = last menstrual period; NA = not applicable; NSAID = nonsteroidal antiinflammatory drug; OGC = oral glucocorticoid; RA = rheumatoid arthritis.
† 8 women missing socioeconomic status (Hollingshead categories, highest = 1 to lowest = 5).
‡ Inflammatory bowel disease, systemic lupus erythematosus, or ankylosing spondylitis for Medi-Cal; inflammatory bowel disease, systemic lupus erythematosus, or ankylosing spondylitis for MotherToBaby.
§ Between LMP and gestational day 139.
¶ Between LMP and gestational day 139, prescription NSAID only for Medi-Cal.
# 0 = no disability, 3 = completely disabled.
** 0 = no pain, 100 = severe pain.
†† 0 = very well, 100 = very poor.
‡‡ Between LMP and gestational day 139, among women with any OGC exposure between LMP and gestational day 139; prednisone equivalent dose.

Table 2. Oral glucocorticoid (OGC) exposure between last menstrual period (LMP) and day 139 and the risk of preterm birth among women in the Medi-Cal study*

Exposure group	No.	PTB, %	Crude RR	Adjusted RR†	Crude RD	Adjusted RD†
All Women	844	–	–	–	–	–
OGC	189	19.1	1.10 (0.78, 1.56)	1.00 (0.71, 1.42)	1.8 (–4.6, 8.2)	0.0 (–6.3, 6.4)
>300 mg cumulative dose‡	84	25.0	1.36 (1.01, 1.85)	1.23 (0.91, 1.68)	6.2 (–0.7, 13.0)	4.2 (–2.3, 10.7)
≤300 mg cumulative dose‡	105	–	–	0.81 (0.59, 1.12)	–	–3.7 (–8.9, 1.5)
No OGC	655	17.3	Ref.	Ref.	Ref.	Ref.
High school/equivalent or less	565	–	–	–	–	–
OGC	125	20.8	1.31 (0.86, 1.98)	1.16 (0.74, 1.80)	4.9 (–3.2, 13.0)	2.9 (–5.2, 11.0)
No OGC	440	15.9	Ref.	Ref.	Ref.	Ref.

* Values are the risk ratio (RR) or risk difference (RD) (95% confidence interval) unless indicated otherwise. PTB = preterm birth; Ref. = reference.
† Adjusted for LMP year (<2010, ≥2010), maternal age, race/ethnicity, maternal education, disability as source of Medi-Cal eligibility, multiple gestation, prepregnancy body mass index category, nulliparity, inflammatory bowel disease or systemic lupus erythematosus (yes/no), and hypertension, disease-modifying antirheumatic drug, prescription nonsteroidal antiinflammatory drug, any inpatient admission, any inflammatory marker laboratory results, any rheumatoid factor laboratory results, number of outpatient visits, emergency department visits, and outpatient visits with a rheumatoid arthritis diagnosis between LMP and gestational day 139.
‡ Prednisone equivalent dose.

unmeasured confounding by RA severity in Medi-Cal by adjusting the exposure-misclassification bias analysis point estimates between OGC exposure and preterm birth for an unmeasured confounder using the array approach described by Schneeweiss (31) and implemented with the episensr package in R statistical software (details in Supplementary Methods in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24865) (32). All other analyses were conducted using SAS statistical software (version 9.4).

RESULTS

Primary analysis. Before gestational day 140, 22.4% of women in Medi-Cal and 49.1% of women in MotherToBaby had OGC exposure (Table 1). DMARD use before gestational day 140, including biologic therapies, was far less common in Medi-Cal (17.5% DMARD; 6.5% biologic DMARD) than MotherToBaby (79.7% DMARD; 67.4% biologic DMARD). DMARDs were more common among OGC-exposed versus unexposed women in Medi-Cal (38.6% versus 11.5%) but not in MotherToBaby (80.3% versus 79.2%). Proxies and measures of disease severity were more common among OGC-exposed versus unexposed women in both studies. In Medi-Cal, 68.4% of women had a high-school diploma equivalent or less education, whereas only 9.5% were in this category in MotherToBaby.

Preterm birth risk in women without OGCs was 17.3% in Medi-Cal and 9.7% in MotherToBaby (Tables 2 and 3), whereas the preterm birth risk in women with OGCs was more similar across data sources (19.1% in Medi-Cal and 21.6% in MotherToBaby). Therefore, no association existed between OGC exposure and preterm birth in Medi-Cal: adjusted RR 1.00 (95% CI 0.71, 1.42), adjusted risk difference 0.0 (95% CI –6.3, 6.4), and there

Table 3. Oral glucocorticoid (OGC) exposure between last menstrual period and day 139 and the risk of preterm birth among women in the MotherToBaby study*

Exposure group	No.	PTB, %	Crude RR	Adjusted RR†	Crude RD	Adjusted RD†
All Women	528	–	–	–	–	–
OGC	259	21.6	2.24 (1.45, 3.45)	1.85 (1.20, 2.84)‡	12.0 (5.8, 18.1)	8.3 (2.6, 14.0)‡
>300 mg cumulative dose§	180	26.7	2.74 (1.77, 4.25)	2.22 (1.43, 3.45)	16.9 (9.5, 24.2)	13.0 (6.1, 19.9)
≤300 mg cumulative dose§	75	9.3	1.06 (0.50, 2.25)	1.00 (0.49, 2.06)	5.9 (–7.0, 8.2)	–1.8 (–9.2, 5.5)
No OGC	269	9.7	Ref.	Ref.	Ref.	Ref.
High school/equivalent or less	50	–	–	–	–	–
OGC	26	23.1	1.38 (0.44, 4.32)	0.81 (0.25, 2.64)¶	6.4 (–15.6, 28.4)	–9.3 (–28.4, 9.7)¶
No OGC	24	16.7	Ref.	Ref.	Ref.	Ref.

* Values are the risk ratio (RR) or risk difference (RD) (95% confidence interval) unless indicated otherwise. PTB = preterm birth; Ref. = reference.
† Adjusted for last menstrual period (LMP) year (<2010, ≥2010), maternal age, non-Hispanic White race/ethnicity, maternal education, socioeconomic status (Hollingshead categories), multiple gestation, prepregnancy body mass index category, nulliparous, prepregnancy hypertension, inflammatory bowel disease, systemic lupus erythematosus, or ankylosing spondylitis (yes/no), disease-modifying antirheumatic drug use between LMP and gestational day 139, nonsteroidal antiinflammatory drug use between LMP and gestational day 139, and health assessment questionnaire score, pain score, and global score at the time of enrollment.
‡ 8 women excluded because of missing socioeconomic status; results did not change materially when including all women without adjusting for socioeconomic status.
§ Prednisone equivalent dose.
¶ 3 women excluded because of missing socioeconomic status; results did not change materially when including all women without adjusting for socioeconomic status.

was an adjusted 1.85-fold (95% CI 1.20, 2.84) increased risk and an 8.3% (95% CI 2.6, 14.0) absolute increase in the risk for preterm birth in MotherToBaby.

Restriction. When restricting Medi-Cal to lower education, characteristics were similar to the full population with the exception of an increase in the proportion of women who were

Table 4. Characteristics among those with a high-school diploma or equivalent or less education in the Medi-Cal study and in the MotherToBaby (MTB) study*

Characteristic	Medi-Cal (n = 565)	MotherToBaby (n = 50)
LMP year 2010 or later	367 (65.0)	17 (34.0)
Maternal age, median (IQR)	28 (9)	31 (7)
Race/ethnicity		
White (Medi-Cal), non-Hispanic White (MTB)	88 (15.6)	26 (52.0)
Black	33 (5.8)	NA
Hispanic	418 (74.0)	NA
Other or unknown	26 (4.6)	24 (48.0)
Maternal education		
High school/equivalent or less	565 (100)	50 (100)
Less than high-school diploma	282 (49.9)	NA
High school/equivalent	283 (50.1)	NA
Some college, college degree, or higher	0 (0)	0 (0)
Disability as source of Medi-Cal eligibility	77 (13.6)	NA
Socioeconomic status, median (IQR)†	NA	3 (1)
Gestational age at enrollment, median (IQR)	NA	11 (8)
Multiple gestation	–	0 (0)
Prepregnancy body mass index		
Underweight to normal weight	213 (37.7)	26 (52.0)
Overweight	165 (29.2)	14 (28.0)
Obese	187 (33.1)	10 (20.0)
Primiparous	133 (23.5)	18 (36.0)
Autoimmune comorbidities‡	54 (9.6)	3 (6.0)
Hypertension (Medi-Cal)/prepregnancy hypertension (MTB)§	24 (4.2)	6 (12.0)
DMARD§	104 (18.4)	44 (88.0)
NSAID¶	118 (20.9)	17 (34.0)
Number of outpatient visits§		
None	186 (32.9)	NA
1–5	230 (40.7)	NA
≥6	149 (26.4)	NA
Emergency department visits, no.§		
None	340 (60.2)	NA
1	93 (16.5)	NA
≥2	123 (23.4)	NA
Inpatient admission ≥1§	–	NA
Outpatient visits with RA diagnosis, no.§		
None	392 (69.4)	NA
1	107 (18.9)	NA
≥2	66 (11.7)	NA
Inflammatory marker laboratory§	80 (14.2)	NA
Rheumatoid factor laboratory§	54 (9.6)	NA
HAQ at enrollment, median (IQR)#	NA	0.6 (1.1)
Pain score at enrollment, median (IQR)**	NA	28 (60)
Global score at enrollment, median (IQR)††	NA	23 (45)
OGC cumulative dose median (IQR)‡‡	300 (430)	695 (448)

* Values are the number (%) unless indicated otherwise. Counts of <16 were suppressed for Medi-Cal. DMARD = disease-modifying antirheumatic drug; HAQ = Health Assessment Questionnaire; IQR = interquartile range; LMP = last menstrual period; NA = not applicable; NSAID = nonsteroidal antiinflammatory drug; OGC = oral glucocorticoid; RA = rheumatoid arthritis.
† 8 women missing socioeconomic status (Hollingshead categories, highest = 1, lowest = 5).
‡ Inflammatory bowel disease, systemic lupus erythematosus, or ankylosing spondylitis for Medi-Cal; inflammatory bowel disease, systemic lupus erythematosus, or ankylosing spondylitis for MTB.
§ Between LMP and gestational day 139.
¶ Between LMP and gestational day 139, prescription NSAID only for Medi-Cal.
# 0 = no disability, 3 = completely disabled.
** 0 = no pain, 100 = severe pain.
†† 0 = very well, 100 = very poor.
‡‡ Between LMP and gestational day 139, among women with any OGC exposure between LMP and gestational day 139; prednisone equivalent dose.

Hispanic from 64.3% to 74.0% (Table 4). Upon restriction in MotherToBaby, the proportion of non-Hispanic White women decreased (79.0% to 52.0%), overweight/obese women increased (38.6% to 48.0%), and RA severity increased (e.g., median HAQ score increased from 0.3 to 0.6 with restriction). Upon restriction, preterm birth risk among women with no OGC exposure during the first 139 days decreased slightly to 15.9% in Medi-Cal and increased to 16.7% in MotherToBaby. The adjusted association between OGC exposure and preterm birth did not change materially in Medi-Cal (adjusted RR 1.16 [95% CI 0.74, 1.80]) and it decreased in MotherToBaby, although the estimate was imprecise (adjusted RR 0.81 [95% CI 0.25, 2.64]).

Dose and risk. Total cumulative OGC dose during the first half of pregnancy was higher in MotherToBaby (median 684 mg prednisone equivalent dose) than in Medi-Cal (median 300 mg prednisone equivalent dose). An OGC dose of ≤300 mg prednisone equivalent dose was not associated with increased preterm birth risk in either study (Tables 2 and 3). Although absolute risks for preterm birth were similarly high across both studies for OGC dose >300 mg (25.0% for Medi-Cal; 26.7% for MotherToBaby), the adjusted RRs differed across the 2 studies (Medi-Cal adjusted RR 1.23 [95% CI 0.91, 1.68]; MotherToBaby adjusted RR 2.22 [95% CI 1.43, 3.45]), as did the adjusted RDs (Medi-Cal adjusted RD –3.7% [95% CI –8.9, 1.5]; MotherToBaby adjusted RD 13.0% [95% CI 6.1, 19.9]).

Exposure misclassification adjustment. Assuming nondifferential misclassification (i.e., OGC misclassification unrelated to preterm birth status) with sensitivity = 60%, and specificity = 85%, the exposure misclassification bias analysis adjusted odds ratio (OR) was 1.40 (95% CI 0.89, 2.28) (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24865). Assuming differential misclassification (i.e., OGC misclassification differing by preterm birth status) with sensitivity = 60%, specificity = 95% for women with preterm birth, and specificity = 85% for women without preterm birth yielded a bias-adjusted OR of 3.05 (95% CI 1.71, 6.62) (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24865). With sensitivity = 60%, specificity = 85% for women with preterm birth, and specificity = 95% for women without preterm birth, the bias-adjusted OR was 0.52 (95% CI 0.26, 0.92).

Severity adjustment. Compared with adjusting for all covariates, not adjusting for disease severity increased the RR by 11% in MotherToBaby (RR 2.05 [95% CI 1.32, 3.17]). Not adjusting for severity proxies did not change the Medi-Cal results materially. The bias analysis for exposure misclassification and unmeasured confounding by RA severity in Medi-Cal indicated a reduced point estimate after adjusting for the unmeasured confounder (e.g., exposure misclassification bias analysis adjusted OR = 1.40, exposure misclassification and unmeasured confounding bias analysis RR = 1.18) (see Supplementary Tables 2 and 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24865).

DISCUSSION

We observed no association between OGC use during the first half of pregnancy and preterm birth among women with RA when using Medi-Cal data. However, a similar analytic approach with prospective cohort data yielded an 8% absolute increase in the risk for preterm birth and nearly a 2-fold increased risk for preterm birth following OGC exposure during the first half of pregnancy. Based on post hoc bias analyses, difference in the results across studies seemed unlikely to be related to threats to internal validity. Instead, differences in the study samples related to SES and OGC dose distribution may have contributed to the discrepancy in the associations across studies.

Preterm birth risk following OGC exposure was similar in both studies; the disparity in results from the full populations originates in the reference groups. Women in Medi-Cal were low income and primarily Hispanic. Furthermore, most women were overweight/obese, and few used DMARDs, several of which are recommended to control disease activity and reduce the risks of flares during pregnancy (e.g., hydroxychloroquine) (33,34), likely resulting in increased disease activity. These factors may have contributed to the high observed baseline risk of preterm birth. Women in MotherToBaby self-selected into the study, which may be a proxy for health-seeking behaviors protective for preterm birth, had high SES, and were primarily non-Hispanic White. Most women were normal-weight/underweight, had relatively low disease severity, and used DMARDs including biologic therapies (which were not associated with an increased risk of preterm birth in these studies [12,15]), resulting in lower baseline preterm birth risk compared with Medi-Cal. Among RA patients in the general population, major disparities in access to DMARDs related to race and SES status have been described, with Medicaid patients being far less likely to receive DMARDs than patients with private insurance (35,36). Therefore, SES may influence the baseline risk for preterm birth among pregnant women with RA through a variety of pathways, e.g., decreased access to DMARDs resulting in increased RA severity.

OGC exposure was not associated with preterm birth when restricting to women with lower education in either study, although the point estimate was imprecise for MotherToBaby, as <10% had lower education. After restriction to women with lower education, the impact of OGC exposure on the development of preterm birth may have been negligible, given high baseline risk for preterm birth (approximately 16% in both studies). Increased RA severity due to less DMARD use and other factors, e.g., inadequate prenatal care, environmental pollution, or experiences of racism (37–40), may have been more impactful contributors to preterm birth than OGC

in a population with lower SES. Therefore, SES appears to be an important effect modifier of the association between OGC use during pregnancy and preterm birth and should be considered when generalizing estimates to the target population or transporting estimates to other populations (41).

The median total cumulative dose of OGC during the first half of pregnancy in Medi-Cal was less than half that observed in MotherToBaby. Furthermore, lower OGC doses were not associated with an increased preterm birth risk in either population. Therefore, the lower distribution of cumulative OGC dose in Medi-Cal may have contributed to the null association observed between OGC and preterm birth in Medi-Cal. Differences in typical OGC regimes with respect to dose across the study samples may have contributed to the discrepancy across the studies.

We assumed that OGC exposure was captured with higher accuracy in MotherToBaby than in Medi-Cal, given the careful collection of medication use information via semistructured interviews at multiple time points during pregnancy in MotherToBaby versus the reliance on pharmacy dispensing data in Medi-Cal. Using estimates from our validation study of OGC exposure during pregnancy (20), correcting for OGC misclassification unrelated to preterm birth status in Medi-Cal led to point estimates that were weaker than those observed in MotherToBaby. Correcting for OGC misclassification related to preterm birth status in Medi-Cal led to point estimates that were stronger than in MotherToBaby when accuracy was higher in women with preterm birth than without preterm birth. However, plausibility is low given that exposure was classified during the first half of pregnancy, well before preterm birth occurred. Therefore, misclassified OGC status in Medi-Cal seems unlikely to have explained the observed discrepancy across the studies.

We assumed that we were able to more fully adjust for RA severity in MotherToBaby, given validated measures of RA severity (21,22) versus proxies of severity in Medi-Cal. Quantitative bias analysis for residual confounding by disease severity following correction for exposure misclassification in Medi-Cal resulted in a weakened association between OGC and preterm birth that was lower than the MotherToBaby point estimate. Therefore, residual confounding by RA severity in Medi-Cal was not a likely explanation for the observed differences across the studies.

A limitation of our study is the small number of women with lower education in MotherToBaby, resulting in imprecise estimates when exploring the impact of making the study populations more similar with respect to SES. Furthermore, although we aimed to create more comparable study populations by restricting to women with the same education level, we acknowledge that education level is a proxy of SES, and there are socioeconomic and other sources of variability across the 2 restricted populations (e.g., health behaviors). All of the women in the Medi-Cal study met low-income eligibility criteria, whereas some women with a high school education or less in MotherToBaby were still classified as having higher SES according to Hollingshead's categories. Medicaid status was not available in MotherToBaby. Furthermore, we had to rely on diagnostic codes to classify RA in Medi-Cal, which may have resulted in the inclusion of some women without RA.

Our study intentionally analyzed the same medication safety question in prospective cohort data and health care utilization data and investigated potential reasons for discrepant answers. We used a similar analytic approach across the data sources, and the data sources had complementary strengths that allowed us to investigate several reasons for differing results. Medi-Cal allowed us to investigate the association of interest in a low-income population without selection of volunteers. MotherToBaby had careful ascertainment of OGC use with maternal report and medical record confirmation and self-reported validated measures of disease severity.

Results of individual studies of medication safety during pregnancy are often at odds with each other. Discrepant results can make counseling on medication safety complex for providers and decision-making fraught for patients. When discrepant results arise, ideally investigators could quantitatively explore threats to internal validity, including exposure misclassification, outcome misclassification, confounding, and selection bias, as well as external validity issues, including differences in study populations related to baseline risks, the distribution of effect modifiers, and treatment regimens, including daily and/or cumulative dose, as possible explanations. In our comparison, differences in results across the data sources may be due to the underlying risk of the outcome in the referent groups of each study. We also observed differences in the distribution of OGC dose, which may have contributed to differences in the association between OGC modeled as a binary yes/no variable and preterm birth (42). However, we could not attribute the differences in results to expected biases in Medi-Cal data (i.e., exposure misclassification, confounding). Our findings underscore the need 1) for authors to describe and contextualize study samples, assess medication dose, and present stratum-specific results for potential effect modifiers, and 2) for readers to consider characteristics of the study sample, baseline risks, and medication dose distribution when comparing discrepant answers to the same perinatal medication safety question.

Given the high baseline risk for preterm birth among Medi-Cal enrollees with RA and replicated among women with a high school education or less in MotherToBaby, and also in enrollees with asthma, systemic lupus erythematosus, and inflammatory bowel disease as described previously (16), studies are needed to understand the effects of autoimmune disease severity and undertreatment of autoimmune diseases on preterm birth risk before and during pregnancy in low-income populations.

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. Palmsten 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. Palmsten, Bandoli, Vazquez-Benitez, Chambers.

Acquisition of data. Chambers.

Analysis and interpretation of data. Palmsten, Bandoli, Vazquez-Benitez, Chambers.

REFERENCES

1. Ferguson CB, Mahsud-Dornan S, Patterson RN. Inflammatory bowel disease in pregnancy. BMJ 2008;337:a427.
2. Ostensen M, Forger F. Management of RA medications in pregnant patients. Nat Rev Rheumatol 2009;5:382–90.
3. Mitchell K, Kaul M, Clowse ME. The management of rheumatic diseases in pregnancy. Scand J Rheumatol 2010;39:99–108.
4. Ateka-Barrutia O, Khamashta MA. The challenge of pregnancy for patients with SLE. Lupus 2013;22:1295–308.
5. Gregersen TL, Ulrik CS. Safety of bronchodilators and corticosteroids for asthma during pregnancy: what we know and what we need to do better. J Asthma Allergy 2013;6:117–25.
6. Namazy JA, Schatz M. The safety of asthma medications during pregnancy: an update for clinicians. Ther Adv Respir Dis 2014;8:103–10.
7. Schatz M, Dombrowski MP, Wise R, Momirova V, Landon M, Mabie W, et al. The relationship of asthma medication use to perinatal outcomes. J Allergy Clin Immunol 2004;113:1040–5.
8. Boyd HA, Basit S, Harpsoe MC, Wohlfahrt J, Jess T. Inflammatory bowel disease and risk of adverse pregnancy outcomes. PLoS One 2015;10:e0129567.
9. Broms G, Granath F, Stephansson O, Kieler H. Preterm birth in women with inflammatory bowel disease: the association with disease activity and drug treatment. Scand J Gastroenterol 2016;51:1462–9.
10. De Man YA, Hazes JM, van der Heide H, Willemsen SP, de Groot CJ, Steegers EA, et al. Association of higher rheumatoid arthritis disease activity during pregnancy with lower birth weight: results of a national prospective study. Arthritis Rheum 2009;60:3196–206.
11. Bandoli G, Palmsten K, Forbess Smith CJ, Chambers CD. A review of systemic corticosteroid use in pregnancy and the risk of select pregnancy and birth outcomes. Rheum Dis Clin North Am 2017;43:489–502.
12. Palmsten K, Bandoli G, Vazquez-Benitez G, Xi M, Johnson DL, Xu R, et al. Oral corticosteroid use during pregnancy and risk of preterm birth. Rheumatology (Oxford) 2020;59:1262–71.
13. Huybrechts KF, Bateman BT, Hernandez-Diaz S. Use of real-world evidence from healthcare utilization data to evaluate drug safety during pregnancy. Pharmacoepidemiol Drug Saf 2019;28:906–22.
14. Palmsten K, Chambers CD. Making the best use of data not created for research. Paediatr Perinat Epidemiol 2018;32:287–9.
15. Wang NF, Kolte AM, Larsen EC, Nielsen HS, Christiansen OB. Immunologic abnormalities, treatments, and recurrent pregnancy loss: what is real and what is not? Clin Obst Gynecol 2016;59:509–23.
16. Palmsten K, Bandoli G, Watkins J, Vazquez-Benitez G, Gilmer TP, Chambers CD. Oral corticosteroids and risk of preterm birth in the California Medicaid program. J Allergy Clin Immunol Pract 2021;9:375–84.
17. Barradas DT, Dietz PM, Pearl M, England LJ, Callaghan WM, Kharrazi M. Validation of obstetric estimate using early ultrasound: 2007 California birth certificates. Paediatr Perinat Epidemiol 2014;28:3–10.
18. Chambers C, Johnson DL, Kiernan E. Approach to evaluating pregnancy safety of anti-rheumatic medications in the OTIS MotherToBaby pregnancy studies: what have we learned? Rheumatology (Oxford) 2018;57:v34–9.
19. Bharti B, Lee SJ, Lindsay SP, Wingard DL, Jones KL, Lemus H, et al. Disease severity and pregnancy outcomes in women with rheumatoid arthritis: results from the Organization of Teratology Information Specialists Autoimmune Diseases in Pregnancy Project. J Rheumatol 2015;42:1376–82.
20. Palmsten K, Hulugalle A, Bandoli G, Kuo GM, Ansari S, Xu R, et al. Agreement between maternal report and medical records during pregnancy: medications for rheumatoid arthritis and asthma. Paediatr Perinat Epidemiol 2018;32:68–77.
21. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137–45.
22. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: dimensions and practical applications. Health Qual Life Outcomes 2003;1:20.
23. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159:702–6.
24. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. Am J Epidemiol 2001;154:854–64.
25. Ting G, Schneeweiss S, Scranton R, Katz JN, Weinblatt ME, Young M, et al. Development of a health care utilisation data-based index for rheumatoid arthritis severity: a preliminary study. Arthritis Res Ther 2008;10:R95.
26. Hollingshead AB. Four factor index of social status, unpublished working paper. New Haven (CT): Yale University; 1975.
27. Schimmer BP, Funder JW. ACTH, adrenal steroids, and pharmacology of the adrenal cortex. In: Brunton LL, Chabner BA, Knollmann BC, eds. Goodman & Gilman's the pharmacological basis of therapeutics. 12th ed. New York: McGraw-Hill Medical; 2011. p. 1209–36.
28. Andrade SE, Gurwitz JH, Davis RL, Chan KA, Finkelstein JA, Fortman K, et al. Prescription drug use in pregnancy. Am J Obstet Gynecol 2004;191:398–407.
29. Palmsten K, Hernandez-Diaz S, Chambers CD, Mogun H, Lai S, Gilmer TP, et al. The most commonly dispensed prescription medications among pregnant women enrolled in the U.S. Medicaid program. Obstet Gynecol 2015;126:465–73.
30. Fox MP, Lash TL, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. Int J Epidemiol 2005;34:1370–6.
31. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf 2006;15:291–303.
32. Haine D. Package 'episensr.' 2019. URL: https://cran.r-project.org/web/packages/episensr/episensr.pdf.
33. Sammaritano LR, Bermas BL, Chakravarty EE, Chambers C, Clowse ME, Lockshin MD, et al. 2020 American College of Rheumatology guideline for the management of reproductive health in rheumatic and musculoskeletal diseases. Arthritis Rheumatol 2020;72:529–56.
34. Balevic SJ, Cohen-Wolkowiez M, Eudy AM, Green TP, Schanberg LE, Clowse ME. Hydroxychloroquine levels throughout pregnancies complicated by rheumatic disease: implications for maternal and neonatal outcomes. J Rheumatol 2019;46:57–63.
35. Barnabe C. Disparities in rheumatoid arthritis care and health service solutions to equity. Rheum Dis Clin North Am 2020;46:685–92.
36. Cifaldi M, Renaud J, Ganguli A, Halpern MT. Disparities in care by insurance status for individuals with rheumatoid arthritis: analysis of the medical expenditure panel survey, 2006–2009. Curr Med Res Opin 2016;32:2029–37.
37. Vintzileos AM, Ananth CV, Smulian JC, Scorza WE, Knuppel RA. The impact of prenatal care in the United States on preterm births in the presence and absence of antenatal high-risk conditions. Am J Obstet Gynecol 2002;187:1254–7.

38. Padula AM, Huang H, Baer RJ, August LM, Jankowska MM, Jellife-Pawlowski LL, et al. Environmental pollution and social factors as contributors to preterm birth in Fresno County. Environ Health 2018; 17:70.
39. Bower KM, Geller RJ, Perrin NA, Alhusen J. Experiences of racism and preterm birth: findings from a pregnancy risk assessment monitoring system, 2004 through 2012. Womens Health Issues 2018;28: 495–501.
40. Chambers BD, Baer RJ, McLemore MR, Jelliffe-Pawlowski LL. Using index of concentration at the extremes as indicators of structural racism to evaluate the association with preterm birth and infant mortality: California, 2011-2012. J Urban Health 2019; 96:159–70.
41. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR. Generalizing study results: a potential outcomes perspective. Epidemiology 2017;28:553–61.
42. Wood ME, Lupattelli A, Palmsten K, Bandoli G, Hurault-Delarue C, Damase-Michel C, et al. Longitudinal methods for modeling exposures in pharmacoepidemiologic studies in pregnancy. Epidemiol Rev 2022;43:130–46.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1342–1348
DOI 10.1002/acr.24559

Development of a Medicare Claims–Based Model to Predict Persistent High-Dose Opioid Use After Total Knee Replacement

Chandrasekar Gopalakrishnan,[1] Rishi J. Desai,[1] Jessica M. Franklin,[1] Yinzhu Jin,[1] Joyce Lii,[1] Daniel H. Solomon,[1] Jeffrey N. Katz,[1] Yvonne C. Lee,[2] Patricia D. Franklin,[2] and Seoyoung C. Kim[1]

Objective. To develop a claims-based model to predict persistent high-dose opioid use among patients undergoing total knee replacement (TKR).

Methods. Using Medicare claims (2010–2014), we identified patients ages ≥65 years who underwent TKR with no history of high-dose opioid use (mean >25 morphine milligram equivalents [MMEs]/day) in the year prior to TKR. We used group-based trajectory modeling to identify distinct opioid use patterns. The primary outcome was persistent high-dose opioid use in the year after TKR. We split the data into training (2010–2013) and test (2014) sets and used logistic regression with least absolute shrinkage and selection operator regularization, utilizing a total of 83 preoperative patient characteristics as candidate predictors. A reduced model with 10 prespecified variables, which included demographic characteristics, opioid use, and medication history was also considered.

Results. The final study cohort included 142,089 patients who underwent TKR. The group-based trajectory model identified 4 distinct trajectories of opioid use (group 1: short-term, low-dose; group 2: moderate-duration, low-dose; group 3: moderate-duration, high-dose; and group 4: persistent high-dose). The model predicting persistent high-dose opioid use achieved high discrimination (receiver operating characteristic area under the curve [AUC] 0.85 [95% confidence interval (95% CI) 0.84–0.86]) in the test set. The reduced model with 10 predictors performed equally well (AUC 0.84 [95% CI 0.84–0.85]).

Conclusion. In this cohort of older patients, 10.6% became persistent high-dose (mean 22.4 MME/day) opioid users after TKR. Our model with 10 readily available clinical factors may help identify patients at high risk of future adverse outcomes from persistent opioid use after TKR.

INTRODUCTION

The economic burden of prescription opioid overdose, abuse, and dependence is estimated to be $78.5 billion each year in the US (1). An estimated 2.0 million people in the US had an opioid-use disorder (defined in the Diagnostic and Statistical Manual of Mental Disorders [2] as a problematic pattern of opioid use leading to clinically significant impairment or distress) associated with prescription opioids in 2015, and nearly half of all opioid-related deaths involved a prescription opioid (3,4).

Patients undergoing major surgical procedures, such as total knee replacements (TKRs), are often prescribed opioids before and/or after surgery for pain relief. A recent study in a commercially insured population revealed that 87.1% of patients filled at least 1 prescription for opioids in the year prior to hip or knee arthroplasty (5,6). In a large cohort of older Medicare enrollees with osteoarthritis, 58.3% had used opioids at least once in the year prior to TKR, and 7.2% had continuous opioid use, defined by a dispensing for opioids at least once every month for 12 months before the surgical procedure (7). Some studies

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant R01AR069557). Dr. Solomon's work was supported by AbbVie, Amgen, Corrona, Genentech, Janssen, and Pfizer (grants to Brigham and Women's Hospital).

[1]Chandrasekar Gopalakrishnan, MD, MPH, Rishi J. Desai, PhD, Jessica M. Franklin, PhD, Yinzhu Jin, MS, MPH, Joyce Lii, MS, MA, Daniel H. Solomon, MD, MPH, Jeffrey N. Katz, MD, MSc, Seoyoung C. Kim, MD, ScD, MSCE: Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; [2]Yvonne C. Lee, MD, MMSc, Patricia D. Franklin, MD, MBA, MPH: Northwestern University, Chicago, Illinois.

Dr. Lee owns stock in Cigna-Express Scripts and serves as an advisory board member for Eli Lilly. No other disclosures relevant to this article were reported.

Address correspondence to Seoyoung C. Kim, MD, ScD, MSCE, Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 1620 Tremont Street, Suite 3030, Boston, MA 02120. Email: sykim@bwh.harvard.edu.

Submitted for publication July 29, 2020; accepted in revised form January 7, 2021.

SIGNIFICANCE & INNOVATIONS

- Using group-based trajectory modeling, our study identified 4 distinct patterns of opioid use among Medicare patients who underwent total knee replacement (TKR) in the year after their surgery.
- A total of 10.6% of these patients became persistent high-dose (mean 22.4 morphine milligram equivalents/day) opioid users in the year after their TKR.
- Our model with 10 readily available preoperative clinical factors achieved excellent predictive performance (receiver operating characteristic area under the curve 0.84) and may help identify patients at high risk of future adverse outcomes from persistent opioid use after TKR.

suggest that use of preoperative or perioperative opioids increases risk for persistent opioid use and opioid dependence following the surgery (8,9). Patients who were continuous users of opioids prior to surgery were found to have poorer surgical outcomes after surgery and nearly a 5-fold increased risk for opioid overdose compared to those who did not use any opioids prior to surgery (7).

Few studies have characterized the longitudinal patterns of opioid use after TKR (5), and those that do evaluate such patterns have not studied in detail the dynamic patterns of use over time. Group-based trajectories have been used to model complex longitudinal outcomes or behaviors such as health care spending (10), postoperative pain (11), and medication adherence to chronic medications (12,13) and may aid in characterizing longitudinal patterns of opioid use over time (14).

The objectives of our study were to: 1) characterize the dynamic patterns of opioid use in the year following TKR of patients who underwent the surgery using group-based trajectory modeling to classify patients with persistent high-dose opioid use; and 2) develop a prognostic clinical prediction model to identify persistent high-dose opioid users after TKR using preoperative patient characteristics.

PATIENTS AND METHODS

Data source and study population. Using Medicare Parts A (inpatient), B (outpatient), and D (prescription) claims (2010–2014) and a cohort design, we defined cohort eligibility as patients ages ≥65 years who underwent TKR (see Supplementary Table 1,* available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24599) and who were continuously enrolled in Medicare for ≥360 days prior to TKR and for a minimum of 30 days after TKR. To prevent inclusion of patients with bilateral TKRs, patients with 2 codes for TKR on the same day or with a history of TKR in the 360 days prior were not included. We defined 2 exclusion criteria: 1) patients with any prior history of cancer, as these patients may be more likely to have an indication for persistent opioid use; and 2) patients with high-dose opioid use at baseline (mean >25 morphine milligram equivalents [MMEs]/day), as previous use of high-dose opioids itself is a strong predictor of future high-dose opioids (15,16). Opioid use was assessed in the 360 days after TKR in 30-day intervals. To calculate MMEs/day, we utilized the "Opioid NDC and Oral MME Conversion File" provided by the Centers for Disease Control and Prevention National Center for Injury Prevention and Control (17) that contains opioid national drug codes with their linked oral MME conversion factors (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24599). If patients filled multiple prescriptions of the same drug on the same day, we only considered the fill with a higher days' supply; however, prescriptions filled for different opioid medications on the same day or in combination with other medications (such as nonsteroid antiinflammatory drugs [NSAIDs]) contributed to the MME calculation.

A signed data use agreement with the Center for Medicare and Medicaid services was available, and the Brigham and Women's Hospital Institutional Review Board approved the protocol for this study. To protect patient privacy, the individual patient-level data will not be made available to other researchers for purposes of reproducing the results or replicating the procedure. Programming codes will be made available upon request on publication of this study to enable other researchers to implement the model proposed in our study.

Group-based trajectory modeling. A trajectory model estimated several regression models simultaneously through maximization of a likelihood function that combined the information from all models (14). Within each opioid use group, usage patterns were modeled as a smooth function of time. The output of a group-based trajectory included estimated probabilities of group membership for each individual and an estimated trajectory curve over time for each group (14). Group-based trajectory models were built using opioid filling patterns in the 360 days after TKR. We modeled opioid use as a continuous variable (MME) in every 30-day interval using Proc Traj in SAS, version 9.4, and specifying a censored normal model. Multiple models were developed varying the number of groups from 2 to 6; the 4-group model was selected as the final model based on assessment of model fit. This included a combination of factors including the Bayesian information criterion, the predicted probability of group membership (typically >0.9) (14), and having an adequate number of patients in each trajectory group.

Outcomes: defining persistent high-dose opioid users. The group-based trajectory model classified patients into 4 distinct trajectories (group 1: short-term, low-dose; group 2: moderate-duration, low-dose; group 3: moderate-duration,

*[Correction added on 30 May 2022, after first online publication: The order of the supplementary material was corrected in this version.]

high-dose; and group 4: persistent, high-dose) of opioid use in the year after TKR. The primary outcome was persistent high-dose opioid use defined as patients in trajectory group 4. Patients in group 4 had the longest average duration (164.4 days) of opioid use and high average dose (22.4 MMEs/day) compared to patients in trajectory group 1 (14.3 days and 20.9 MMEs/day), group 2 (44.9 days and 14.4 MMEs/day), and group 3 (51.1 days and 38.8 MMEs/day) (Figure 1).

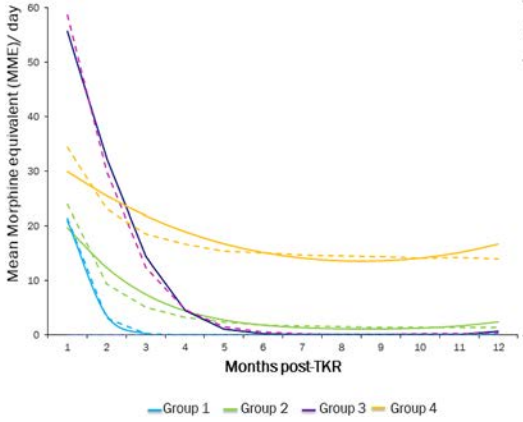
Development of prediction model. We used logistic regression to predict membership in trajectories with persistent high-dose opioid use (group 4) versus other groups (groups 1, 2, and 3) as a binary outcome utilizing all candidate predictors and implemented the least absolute shrinkage and selection operator (LASSO) regression for variable selection using the Glmnet package in R (18). We performed 10-fold cross-validation to select the value of the penalty parameter in a way that minimized the mean cross-validated error. The study cohort was split into a training set (2010–2013 [71.7% of the sample]) and a test set (2014 [28.3% of the sample]) based on the year of cohort entry. We used the regression coefficients for each of the selected predictor variables estimated in the training set to predict persistent high dose versus other groups in the test set.

Predictors. Based on Medicare medical, procedure, or pharmacy claims, we defined a total of 83 investigator-specified candidate predictors in the 360 days prior to the date of TKR (the index date). These predictors included demographic characteristics (such as age, sex, and race), comorbidities (such as substance use disorders, depression, arthritis), co-medications (such as baseline opioid use, benzodiazepines, and antidepressants), health care utilization variables (such as emergency room visits, hospitalizations, and physician office visits), comorbidity index (19), and markers of frailty captured using a validated claims-based frailty index (20). We also considered a reduced model with only 10 prespecified variables, which included demographic characteristics, history of opioid use and other substance use, and prior medication history. The rationale for these 10 variables was to include variables that are readily available for clinical use but also reliably captured in claims data (such as medication use) that may enhance the external validity of the model in other settings.

Performance metrics and diagnostics. Predictive performance of the model was assessed by calculating the receiver operating characteristic (ROC) area under the curve (AUC), with values close to 0.5 indicating an uninformative model and values close to 1 indicating perfect prediction. Since the ROC AUC can be too optimistic, we also calculated the area under the precision-recall curve (AUPRC), which factors in the outcome prevalence and visually plots the tradeoff between positive predictive value (PPV) and sensitivity (21). In contrast to the ROC AUC, the lower bound of the AUPRC is determined by the outcome prevalence (~0.10 in our study given 10% outcome prevalence), which indicates an uninformative model with values close to 1 indicating perfect prediction. Plots of ROC and the precision recall curve were generated in both the training and test data. Finally, calibration plots were produced to assess if the predicted probabilities match the actual probabilities by visual inspection and report of the intercept and slope. Intercept values close to 0 and slope values close to 1 indicate perfect prediction (22–24).

RESULTS

We identified 142,089 patients ages ≥65 years who underwent TKR and met all eligibility criteria (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24599). The 4-group trajectory model identified clusters of patients (Figure 1) with clinically meaningful discrimination between patients' opioid use. Patient characteristics between patients in group 4 (long-



Group	N (%)	Average MME/day in 360-day follow-up (Mean, SD)	Total opioid days covered in 360-day follow-up (Mean, SD)
4	15,044 (10.6)	22.4 (12.7)	164.4 (86.9)
3	8,020 (5.6)	38.8 (16.6)	51.1 (21.3)
2	38,224 (26.9)	14.4 (6.4)	44.9 (29.5)
1	80,801 (56.9)	20.9 (11.9)	14.3 (10.4)

Figure 1. Trajectories of opioid-filling patterns after total knee replacement (TKR). Graph shows the observed mean morphine milligram equivalents (MMEs) per day (dotted lines) and the predicted values of MMEs from the model (solid lines).

Table 1. Selected baseline patient characteristics by opioid use trajectory groups*

	Group 1 (short-term, low-dose)	Group 2 (moderate duration, low-dose)	Group 3 (moderate duration, high-dose)	Group 4 (persistent, high-dose)
Total no. of patients	80,801	38,224	8,020	15,044
Age, mean ± SD years	73.09 ± 5.35	72.68 ± 5.23	70.64 ± 4.09	72.03 ± 5.09
Male sex	30,025 (37.2)	11,887 (31.1)	3,018 (37.6)	4,089 (27.2)
White race	74,686 (92.4)	33,757 (88.3)	7,445 (92.8)	13,228 (87.9)
Combined Comorbidity Index (ref. 19), mean ± SD	0.50 ± 1.64	0.71 ± 1.78	0.45 ± 1.51	0.95 ± 1.94
Frailty categories (ref. 20)				
Mild (<0.13)	21,110 (26.1)	7,000 (18.3)	2,136 (26.6)	1,756 (11.7)
Moderate (0.13 to <0.16)	32,015 (39.6)	14,032 (36.7)	3,027 (37.7)	4,725 (31.4)
Severe (≥0.16)	27,676 (34.3)	17,192 (45.0)	2,857 (35.6)	8,563 (56.9)
Baseline MMEs/day, mean ± SD	0.77 ± 1.95	2.18 ± 3.54	2.27 ± 4.23	9.02 ± 7.00
Tobacco use	8,230 (10.2)	4,401 (11.5)	1,082 (13.5)	2,279 (15.1)
Alcohol abuse	482 (0.6)	277 (0.7)	71 (0.9)	149 (1.0)
Anxiety	6,112 (7.6)	3,927 (10.3)	837 (10.4)	2,351 (15.6)
Falls	2,569 (3.2)	1,600 (4.2)	243 (3.0)	904 (6.0)
Back pain	30,837 (38.2)	18,019 (47.1)	3,430 (42.8)	8,799 (58.5)
Depression	8,051 (10.0)	5,164 (13.5)	1,083 (13.5)	2,933 (19.5)
Diabetes	21,488 (26.6)	12,173 (31.8)	2,158 (26.9)	5,131 (34.1)
Drug abuse	39 (0.0)	37 (0.1)	11 (0.1)	54 (0.4)
No. of ER visits, mean ± SD	0.25 ± 0.63	0.35 ± 0.80	0.29 ± 0.78	0.47 ± 1.01
No. of drugs, mean ± SD	8.46 ± 4.86	10.60 ± 5.39	9.66 ± 5.20	12.77 ± 5.83
No. of office visits, mean ± SD	10.39 ± 6.25	11.94 ± 7.03	11.09 ± 6.51	13.18 ± 7.75

* Values are the number (%) unless indicated otherwise. ER = emergency room; MME = mean morphine milligram equivalents.

term, high-dose opioid use) and group 1 (short-term, low-dose opioid use) differed greatly (Table 1). Patients in group 4 versus those in group 1 were more likely to be younger and female, have higher baseline opioid use (9.0 MMEs/day versus 0.8 MMEs/day, respectively) in the year prior to TKR, use tobacco/alcohol, have a greater burden of comorbidities (such as depression, anxiety disorders, and back pain), and use multiple medications. Patients in group 4 were also more likely to be frail, as assessed by a claims-based frailty index (20), and were more likely to use health care services in general (such as number of emergency room and physician office visits).

Using logistic regression and LASSO, we predicted the probability of persistent high-dose opioid use (n = 11,607) in the training data (n = 101,810) for an AUC of 0.88 (95% confidence

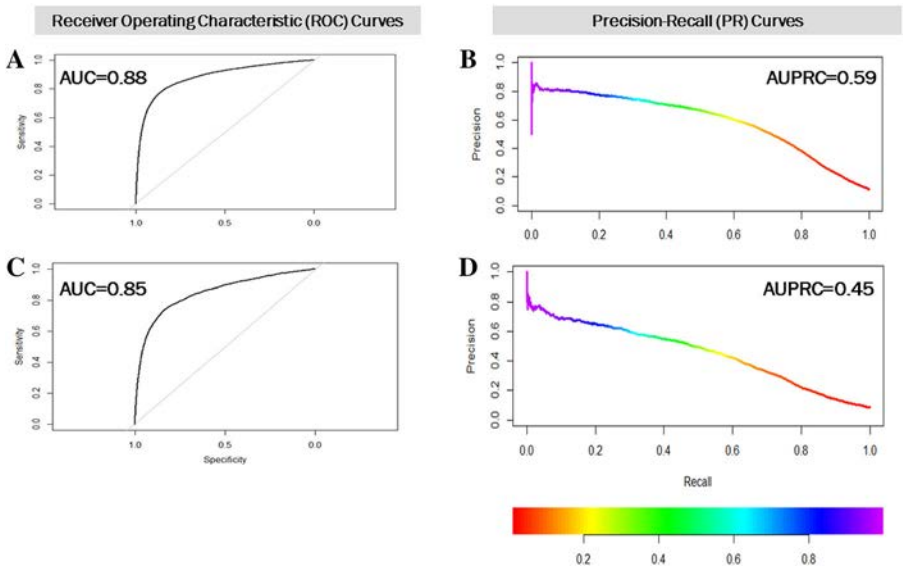


Figure 2. Comparison of the receiver operating characteristic (ROC) and precision-recall (PR) curves in the training (**A** and **B**) and test (**C** and **D**) data for the full model (group 4 versus groups 1, 2, and 3). AUC = area under the curve; AUPRC = area under the PR curve. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24559/abstract.

interval [95% CI] 0.88–0.89). The AUC in the test data (n = 40,279) predicting high-dose opioid use (n = 3,437) was 0.85 (95% CI 0.84–0.86) (Figure 2; see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24599). The AUPRC was 0.59 and 0.45, respectively, in the training and test data (Figure 2). Visual inspection of the calibration plot suggested that the model tended to overestimate the predicted probability of persistent high-dose opioid use when the actual probabilities were >40% (calibration slope = 0.82) (see Supplementary Figure 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24599). The final model with the lowest mean cross-validated error selected 22 predictors (see Supplementary Figure 3). Figure 3 shows the variables selected in the full model and their coefficients predicting persistent high-dose opioid use. The strongest positive predictor of persistent high-dose opioid use was baseline opioid use. The other variables selected in the final model included demographic characteristics (such as age, region, and race), medication use (such as use of benzodiazepines, anxiolytic medications, NSAIDs, and antidepressants), substance use, and comorbidities (such as migraine, anxiety, and back pain).

A reduced model with only 10 investigator-specified predictors, which included demographic characteristics (age, sex, and race), history of substance abuse (opioids, alcohol, and tobacco) and medication use (benzodiazepines, anxiolytics, antidepressants, anticonvulsants and NSAIDs), also showed comparable predictive ability in terms of discrimination (AUC 0.84 [95% CI 0.84–0.85]), precision (AUPRC 0.45), and calibration (calibration slope = 0.79) (see Supplementary Tables 1 and 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24599). The strongest predictors of persistent high-dose opioid use were baseline opioid use (modeled as a continuous variable) (odds ratio [OR] 1.33 [95% CI 1.33–1.34]), Black race (OR 1.50 [95% CI 1.36–1.66]), and history of benzodiazepine use (OR 1.44 [95% CI 1.32–1.57]) (Table 2). The coefficients for the final predictive model are reported (see Supplementary Table 5, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24599), and an online risk calculator that will be publicly available on our website (https://www.bwhprime.org/).

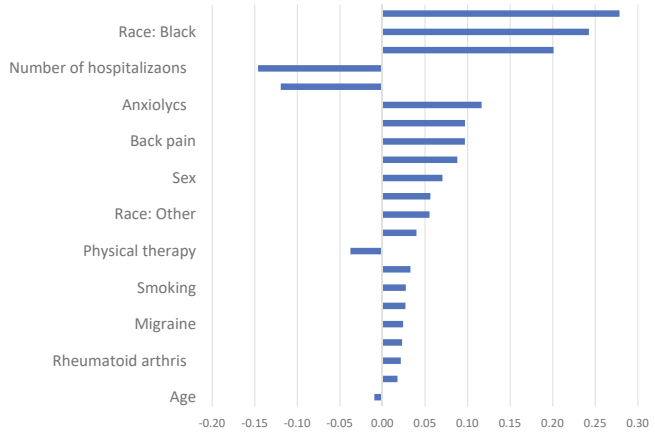


Figure 3. Bar graph of the variables selected in the full model and the range of the coefficients from the least absolute shrinkage and selection operator model predicting persistent high-dose opioid use. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24559/abstract.

Table 2. Predictors of trajectories of persistent high-dose opioid use in the reduced model in the training data*

Variable	Multivariable OR (95% CI) persistent high dose (group 4) vs. groups 1, 2, and 3	*P*†
Age, years	0.98 (0.98–0.99)	<0.001
Female sex (ref. = male sex)	0.83 (0.79–0.88)	<0.001
Race		
Other	1.15 (1.04–1.28)	0.008
Black	1.50 (1.36–1.66)	<0.001
White (ref.)	1.00	
Baseline opioid use (MMEs/day)	1.33 (1.33–1.34)	<0.001
Substance use (yes/no)	1.14 (1.06–1.22)	<0.001
Benzodiazepine use (yes/no)	1.44 (1.32–1.57)	<0.001
Anxiolytic use (yes/no)	1.27 (1.19–1.36)	<0.001
Anticonvulsant use (yes/no)	1.25 (1.17–1.33)	<0.001
Antidepressant use (yes/no)	1.16 (1.10–1.23)	<0.001
NSAID use (yes/no)	1.21 (1.15–1.27)	<0.001

* 95% CI = 95% confidence interval; MMEs = morphine milligram equivalents; NSAID = nonsteroidal antiinflammatory drug; OR = odds ratio; ref. = reference.
† Values are significant.

DISCUSSION

In this large cohort of older Medicare enrollees who underwent TKR (mean age 72.7 years), group-based trajectory modeling identified groups of patients with 4 distinct patterns of opioid use in the year following the surgery. We classified 1 of these groups, comprising 10.6% of the population, as persistent high-dose users (average dose of 22.4 MMEs/day for an average duration of 164.4 days) and developed a prediction model using only preoperative patient characteristics to predict membership in the empirically identified trajectory group, which showed excellent predictive performance. Using a reduced model with only 10 investigator-specified predictors that would be more readily available in clinical settings also resulted in comparable predictive performance.

The key predictors of persistent high-dose opioid use were baseline use of opioids, benzodiazepines, anxiolytics, antidepressants and diagnosis or treatments for chronic painful conditions. Numerous prior studies (5,7,15,25) have identified

baseline opioid use as an extremely strong predictor of future opioid use, and this finding was also observed in our study, even after excluding patients with high-dose opioid use at baseline. Although the final model selected 22 predictors, only minimal gains were observed beyond adding baseline opioid use, with the most regularized model only selecting 13 predictors. Black race (compared to White race) was the second strongest predictor identified and is consistent with recent trends that suggest large increases in rates of opioid overdose among older Black men and women (26,27).

Other influential predictors were use of benzodiazepines, which have been associated with higher rates of opioid dependence and overdose in prior studies (5,28), and the presence of chronic painful conditions (the majority of patients in the cohort [99.5%] had a diagnosis of osteoarthritis or rheumatoid arthritis at baseline) that require non-opioid pain medications (such as NSAIDs or gabapentin), which have also been identified in other studies as a predictor of future persistent opioid use (5,25).

The predictors in the reduced model focused mainly on variables that are well captured in claims, such as demographic characteristics and prior medication use (including opioid use), and thus may be more generalizable when used in other settings. Using a simple algorithm to preoperatively identify patients at high risk of persistent opioid use may help providers exercise appropriate caution before prescribing opioids after surgery. The predictive performance of our model was comparable to other studies that have aimed to predict prolonged opioid use after orthopedic surgeries (15,25). The tradeoff for gains in positive predictive value to loss of sensitivity may be a worthwhile compromise, given the scale of the opioid epidemic in the US. A greater emphasis on perioperative interventions to improve early mobilization (such as using an adductor canal block compared to femoral nerve block) (29,30) and manage postoperative pain using non-opioid pain management strategies after TKR (such as use of NSAIDs, gabapentin [31], local infiltrating analgesia [32], or knee braces [32]) may be warranted for patients at high risk of future opioid dependence.

This study includes older patients enrolled in Medicare and thus, our findings may not be generalizable to other populations. Patients who did not have a minimum of 30 days of follow-up after TKR were excluded. Most of these patients (90.2%) were admitted to a nursing home where it was no longer possible to track their prescription claims; however, the demographic characteristics of patients excluded were largely similar to those in the final cohort, thus minimizing the potential for bias. While the trajectory model empirically identified a group of patients with persistent high-dose opioid use, it remains possible that such use is specific to our study population and thus our prediction model may be prone to overfitting, which highlights the need for future work to externally validate our model using alternate data sources before it can be translated into practice. Another limitation is that we were not aware of the specific indication for opioid use, although we tried to minimize this issue by excluding patients with cancer and high-dose opioid use prior to surgery. Finally, opioid use is captured in claims data based on filled prescriptions rather than actual consumption, which may be susceptible to some misclassification; however, this method of assessing exposure is generally superior to physician prescribing records and patient self-reporting (33,34).

In conclusion, in this cohort of 142,089 older patients with no history of cancer or high-dose opioid use at baseline, 10.6% became persistent high-dose (mean 22.4 MMEs/day) opioid users during the year after TKR. Our prediction model developed using Medicare claims with 10 readily available clinical factors may help identify patients at high risk of future adverse outcomes from persistent opioid use and dependence after TKR.

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. Kim 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. Gopalakrishnan, Desai, Solomon, Katz, Lee, P.D. Franklin, Kim.

Acquisition of data. Kim.

Analysis and interpretation of data. Gopalakrishnan, Desai, J.M. Franklin, Jin, Lii, Kim.

REFERENCES

1. Florence CS, Zhou C, Luo F, Xu L. The economic burden of prescription opioid overdose, abuse, and dependence in the United States, 2013. Med Care 2016;54:901–6.
2. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Arlington: American Psychiatric Association; 2013.
3. Huang B, Dawson DA, Stinson FS, Hasin DS, Ruan WJ, Saha TD, Smith SM, Goldstein RB, Grant BF. Prevalence, correlates, and comorbidity of nonmedical prescription drug use and drug use disorders in the United States: results of the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry 2006;67:1062–73.
4. Guy GP. Vital signs: changes in opioid prescribing in the United States, 2006–2015. MMWR Morb Mortal Wkly Rep 2017;66.
5. Kim SC, Choudhry N, Franklin JM, Bykov K, Eikermann M, Lii J, et al. Patterns and predictors of persistent opioid use following hip or knee arthroplasty. Osteoarthritis Cartilage 2017;25:1399–406.
6. Wright EA, Katz JN, Abrams S, Solomon DH, Losina E. Trends in prescription of opioids from 2003-2009 in persons with knee osteoarthritis. Arthritis Care Res (Hoboken) 2014;66:1489–95.
7. Kim SC, Jin Y, Lee YC, Lii J, Franklin PD, Solomon DH, et al. Association of preoperative opioid use with mortality and short-term safety outcomes after total knee replacement. JAMA Netw Open 2019;2:e198061.
8. Chou R, Fanciullo GJ, Fine PG, Adler JA, Ballantyne JC, Davies P, American Pain Society-American Academy of Pain Medicine Opioids Guidelines Panel, et al. Clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. J Pain Off J Am Pain Soc 2009;10:113–30.

9. Goesling J, Moser SE, Zaidi B, Hassett AL, Hilliard P, Hallstrom B, et al. Trends and predictors of opioid use after total knee and total hip arthroplasty. Pain 2016;157:1259–65.
10. Lauffenburger JC, Franklin JM, Krumme AA, Shrank WH, Brennan TA, Matlin OS, et al. Longitudinal patterns of spending enhance the ability to predict costly patients: a novel approach to identify patients for cost containment. Med Care 2017;55:64–73.
11. Singh JA, Lemay CA, Nobel L, Yang W, Weissman N, Saag KG, et al. Association of early postoperative pain trajectories with longer-term pain outcome after primary total knee arthroplasty. JAMA Netw Open 2019;2:e1915105.
12. Franklin JM, Gopalakrishnan C, Krumme AA, Singh K, Rogers JR, Kimura J, et al. The relative benefits of claims and electronic health record data for predicting medication adherence trajectory. Am Heart J 2018;197:153–62.
13. Franklin JM, Krumme AA, Tong AY, Shrank WH, Matlin OS, Brennan TA, et al. Association between trajectories of statin adherence and subsequent cardiovascular events. Pharmacoepidemiol Drug Saf 2015;24:1105–13.
14. Franklin JM, Shrank WH, Pakes J, Sanfélix-Gimeno G, Matlin OS, Brennan TA, et al. Group-based trajectory models: a new approach to classifying and predicting long-term medication adherence. Med Care 2013;51:789–96.
15. Karhade AV, Schwab JH, Bedair HS. Development of machine learning algorithms for prediction of sustained postoperative opioid prescriptions after total hip arthroplasty. J Arthroplasty 2019;34:2272–7.e1.
16. Politzer CS, Kildow BJ, Goltz DE, Green CL, Bolognesi MP, Seyler TM. Trends in opioid utilization before and after total knee arthroplasty. J Arthroplasty 2018;33:S147–153.e1.
17. Centers for Disease Control and Prevention. Opioid prescription data resources: analyzing opioid prescription data and oral morphine milligram equivalents (MME). 2020. URL: https://www.cdc.gov/drugoverdose/resources/data.html.
18. Friedman J, Hastie T, Tibshirani R. Glmnet: LASSO and elastic-net regularized generalized linear models. R package version. 2009;1.
19. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol 2011;64:749–59.
20. Kim DH, Glynn RJ, Avorn J, Lipsitz LA, Rockwood K, Pawar A, Schneeweiss S. Validation of a claims-based frailty index against physical performance and adverse health outcomes in the health and retirement study. J Gerontol Ser A. 2019;74:1271–6.
21. Davis J, Goadrich M. The relationship between precision-recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning. ICML 2006. Association for Computing Machinery; 2006:233–40.
22. Steyerberg EW, Harrell FE Jr. Regression modeling strategies: with applications, to linear models, logistic and ordinal regression, and survival analysis, 2nd ed. New York: Springer; 2016.
23. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiol Camb Mass 2010;21:128–38.
24. Desai RJ, Wang SV, Vaduganathan M, Evers T, Schneeweiss S. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. JAMA Netw Open 2020;3:e1918962.
25. Karhade AV, Cha TD, Fogel HA, Hershman SH, Tobert DG, Schoenfeld AJ, et al. Predicting prolonged opioid prescriptions in opioid-naïve lumbar spine surgery patients. Spine J 2020;20:888–95.
26. Lippold KM, Jones CM, Olsen EO, Giroir BP. Racial/ethnic and age group differences in opioid and synthetic opioid-involved overdose deaths among adults aged ≥18 years in metropolitan areas: United States, 2015-2017. MMWR Morb Mortal Wkly Rep 2019;68:967–73.
27. Shiels MS, Freedman ND, Thomas D, Berrington de Gonzalez A. Trends in U.S. drug overdose deaths in non-hispanic black, hispanic, and non-hispanic white persons, 2000-2015. Ann Intern Med 2018;168:453–5.
28. Burke LG, Zhou X, Boyle KL, Orav EJ, Bernson D, Hood ME, et al. Trends in opioid use disorder and overdose among opioid-naive individuals receiving an opioid prescription in Massachusetts from 2011 to 2014. Addiction 2020;115:493–504.
29. Kim DH, Lin Y, Goytizolo EA, Kahn RL, Maalouf DB, Manohar A, et al. Adductor canal block versus femoral nerve block for total knee arthroplasty: a prospective, randomized, controlled trial. Anesthesiology 2014;120:540–50.
30. Krishnan SH, Gilbert LA, Ghoddoussi F, Applefield DJ, Kassab SS, Ellis TA. Addition of buprenorphine to local anesthetic in adductor canal blocks after total knee arthroplasty improves postoperative pain relief: a randomized controlled trial. J Clin Anesth 2016;33:432–7.
31. Hah J, Mackey SC, Schmidt P, McCue R, Humphreys K, Trafton J, et al. Effect of perioperative gabapentin on postoperative pain resolution and opioid cessation in a mixed surgical cohort: a randomized clinical trial. JAMA Surg 2018;153:303–11.
32. Derogatis MJ, Sodhi N, Anis HK, Ehiorobo JO, Bhave A, Mont MA. Pain management strategies to reduce opioid use following total knee arthroplasty. Surg Technol Int 2019;35:301–10.
33. West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A. Recall accuracy for prescription medications: self-report compared with database information. Am J Epidemiol 1995;142:1103–12.
34. West SL, Strom BL, Freundlich B, Normand E, Koch G, Savitz DA. Completeness of prescription recording in outpatient medical records from a health maintenance organization. J Clin Epidemiol 1994;47:165–71.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1349–1358
DOI 10.1002/acr.24581

American College of Rheumatology
Empowering Rheumatology Professionals

Societal Cost of Opioid Use in Symptomatic Knee Osteoarthritis Patients in the United States

Jamie L. Huizinga,[1] Elizabeth E. Stanley,[1] James K. Sullivan,[1] Shuang Song,[1] David J. Hunter,[2] A. David Paltiel,[3] Tuhina Neogi,[4] Robert R. Edwards,[5] Jeffrey N. Katz,[6] and Elena Losina[7]

Objective. Symptomatic knee osteoarthritis (SKOA) is a chronic, disabling condition, requiring long-term pain management; over 800,000 SKOA patients in the US use opioids on a prolonged basis. We aimed to characterize the societal economic burden of opioid use in this population.

Methods. We used the Osteoarthritis Policy Model, a validated computer simulation of SKOA, to estimate the opioid-related lifetime and annual cost generated by the US SKOA population. We included direct medical, lost productivity, criminal justice, and diversion costs. We modeled the SKOA cohort with a mean ± SD age of 54 ± 14 years and Western Ontario and McMaster Universities Osteoarthritis Index pain score of 29 ± 17 (0–100, 100 = worst). We estimated annual costs of strong ($1,381) and weak ($671) opioid regimens using Medicare fee schedules, Red Book, the Federal Supply Schedule, and published literature. The annual lost productivity and criminal justice costs of opioid use disorder (OUD), obtained from published literature, were $11,387 and $4,264, per-person, respectively. The 2015–2016 Medicare Current Beneficiary Survey provided OUD prevalence. We conducted sensitivity analyses to examine the robustness of our estimates to uncertainty in input parameters.

Results. Assuming 5.1% prevalence of prolonged strong opioid use, the total lifetime opioid-related cost generated by the US SKOA population was estimated at $14.0 billion, of which only $7.45 billion (53%) were direct medical costs.

Conclusion. Lost productivity, diversion, and criminal justice costs comprise approximately half of opioid-related costs generated by the US SKOA population. Reducing prolonged opioid use may lead to a meaningful reduction in societal costs that can be used for other public health causes.

INTRODUCTION

Opioid prescription practices in the US contributed to the current opioid epidemic (1,2). In 2018, an estimated 9.9 million persons misused prescription opioids, and 2 million met the Diagnostic and Statistical Manual of Mental Disorders criteria for opioid use disorder (OUD), a problematic pattern of opioid use leading to significant impairment (3). Drug overdose is the leading cause of accidental death in the US, killing over 70,000 individuals in 2017 alone (4,5). Furthermore, the opioid epidemic is costly; the annual cost of opioid misuse in the US is $61.1 billion (6). Today, opioid prescription practices contribute to opioid misuse through increased risk of OUD among patients prescribed opioids (7) and through diversion, the illicit sharing of opioids by prescription recipients with others who misuse the drugs (8).

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grants R01-AR-074290, K24-AR-057827, and P30-AR-072577) and by a Burden of Disease grant from Pfizer.

[1]Jamie L. Huizinga, BS, Elizabeth E. Stanley, BSE, James K. Sullivan, BA, Shuang Song, DVM, MPH: Brigham and Women's Hospital, Boston, Massachusetts; [2]David J. Hunter, MBBS, PhD, MSc: University of Sydney and Royal North Shore Hospital, Sydney, New South Wales, Australia; [3]A. David Paltiel, PhD, MBA: Yale School of Public Health, New Haven, Connecticut; [4]Tuhina Neogi, MD, PhD: Boston University School of Medicine, Boston, Massachusetts; [5]Robert R. Edwards, PhD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [6]Jeffrey N. Katz, MD, MSc: Brigham and Women's Hospital, Harvard Medical School, and Harvard T. H. Chan School of Public Health, Boston, Massachusetts; [7]Elena Losina, PhD: Brigham and Women's Hospital, Harvard Medical School, and Boston University School of Public Health, Boston, Massachusetts.

Ms. Huizinga and Ms. Stanley contributed equally to this work.

Dr. Hunter has received consulting fees, speaking fees, and/or honoraria from Merck Serono, TLCBio, Pfizer, and Eli Lilly (less than $10,000 each). Dr. Neogi has received consulting fees from Pfizer/Lilly, EMD Serono, and Novartis (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Elena Losina, PhD, Policy and Innovation Evaluation in Orthopaedic Treatments Center, Orthopaedic and Arthritis Center for Outcomes Research, Department of Orthopaedic Surgery, Brigham and Women's Hospital, 75 Francis Street, BTM 5016, Boston, MA 02115. Email: elosina@bwh.harvard.edu.

Submitted for publication July 13, 2020; accepted in revised form February 12, 2021.

SIGNIFICANCE & INNOVATIONS

- Although opioids are not recommended for knee osteoarthritis (OA) pain management, recent data suggest that 858,000 knee OA patients in the US use opioids regularly.
- We used the Osteoarthritis Policy Model to estimate the opioid-related and overall costs of knee OA and the distribution of those costs between direct medical outlays, lost time and productivity, diversion, and criminal justice.
- We estimate that almost half of the $14 billion total societal cost of opioid use among persons with symptomatic knee OA is being used to pay for lost work productivity, criminal justice, and diversion activities that are not directly connected to pain management and clinical care.
- Using a novel computer simulation model and data from national sources, this study offers new evidence of the magnitude of the societal burden generated by opioid use and misuse.

Opioid use for long-term pain management has driven increases in opioid prescriptions for musculoskeletal conditions (9), which are estimated to represent 53% of opioid prescriptions in the US (10). Symptomatic knee osteoarthritis (SKOA) is a leading cause of chronic musculoskeletal pain for which opioids are frequently prescribed in the US (11,12); as many as 858,000 patients with knee OA may use opioids regularly for pain management (11,13,14). Current treatment guidelines do not recommend opioid use for knee OA (15,16) based on a growing body of evidence questioning the utility of opioids for long-term pain management (17–21). Additionally, opioid use is not cost-effective for management of knee OA pain (21). Lastly, Zhao et al demonstrated that, compared to other OA patients in the US, those using opioids missed significantly more workdays, resulting in annual incremental wage loss of $1,395 (22).

A recent analysis of data from the Medicare Current Beneficiary Survey (MCBS) found that prolonged opioid use among knee OA patients decreased from 12% in 2013 to 4% in 2016 (23). Despite these encouraging changes in opioid prescription practices, even a small proportion of SKOA patients in the US using opioids on a prolonged basis likely results in a large societal cost. We aimed to estimate the annual and lifetime contribution of opioids to knee OA–related costs. We considered the direct medical cost of treatment and the cost of lost work productivity, as well as the cost of interactions with the criminal justice system among opioid users experiencing OUD. We also estimated the cost of prescriptions diverted from knee OA patients to others with OUD. Further, we aimed to illustrate how changes in opioid prescription practices influence the overall lifetime cost of knee OA.

MATERIALS AND METHODS

Analytic overview. We used the Osteoarthritis Policy (OAPol) model (21,24,25), a validated microsimulation of the natural history and treatment of SKOA, to estimate average lifetime and annual costs following initial presentation to care with SKOA. We considered the direct medical costs of SKOA treatment and treatment-related adverse events, costs of lost work productivity related to SKOA, criminal justice costs resulting from OUD among SKOA patients, and cost of opioids diverted from those with SKOA for illicit use by others. We assessed opioid-specific costs, as well as the overall costs of SKOA. Outcomes included the lifetime and annual average cost per-person with knee OA and national cost in the US. Annual costs were estimated as the lifetime cost divided by the disease duration. Costs were discounted 3% annually and are reported in 2018 US$.

OAPol model. The OAPol model is a validated, state-transition microsimulation (21,25). At the start of analysis, it generates a cohort of hypothetical subjects whose characteristics, including age, sex, race, body mass index (BMI), and OA severity, are determined based on user-specified distributions. We define subjects' radiographic OA severity using the Kellgren/Lawrence grading scale (26) and pain severity using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale (0 = no pain, 100 = worst pain) (27). The OAPol model uses Monte Carlo simulation to transition persons among health states defined by knee OA severity and comorbidities and tracks the associated annual costs throughout the model analysis.

Subjects with SKOA experience a progression of OA treatment, including pharmacologic pain management, glucocorticoid injections, and total knee arthroplasty (TKA) (Figure 1 and Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581, with Supplementary Figure 1 and Supplementary Tables 1–12). These treatments reduce subjects' pain and are associated with costs and adverse events. Each year, the model determines the cost each subject has incurred from SKOA pain and treatment. Model outcomes include the average lifetime opioid-attributable and knee OA–attributable medical and nonmedical costs in persons with SKOA.

Cohort characteristics. In this analysis, we simulated subjects from initial presentation to care for SKOA until death. The demographic characteristics were derived from National Health Interview Survey respondents with knee OA (see Supplementary Table 13, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581) (25). Average pain and pain levels were derived from the Osteoarthritis Initiative after adjusting for age, sex, race, and BMI. At simulation

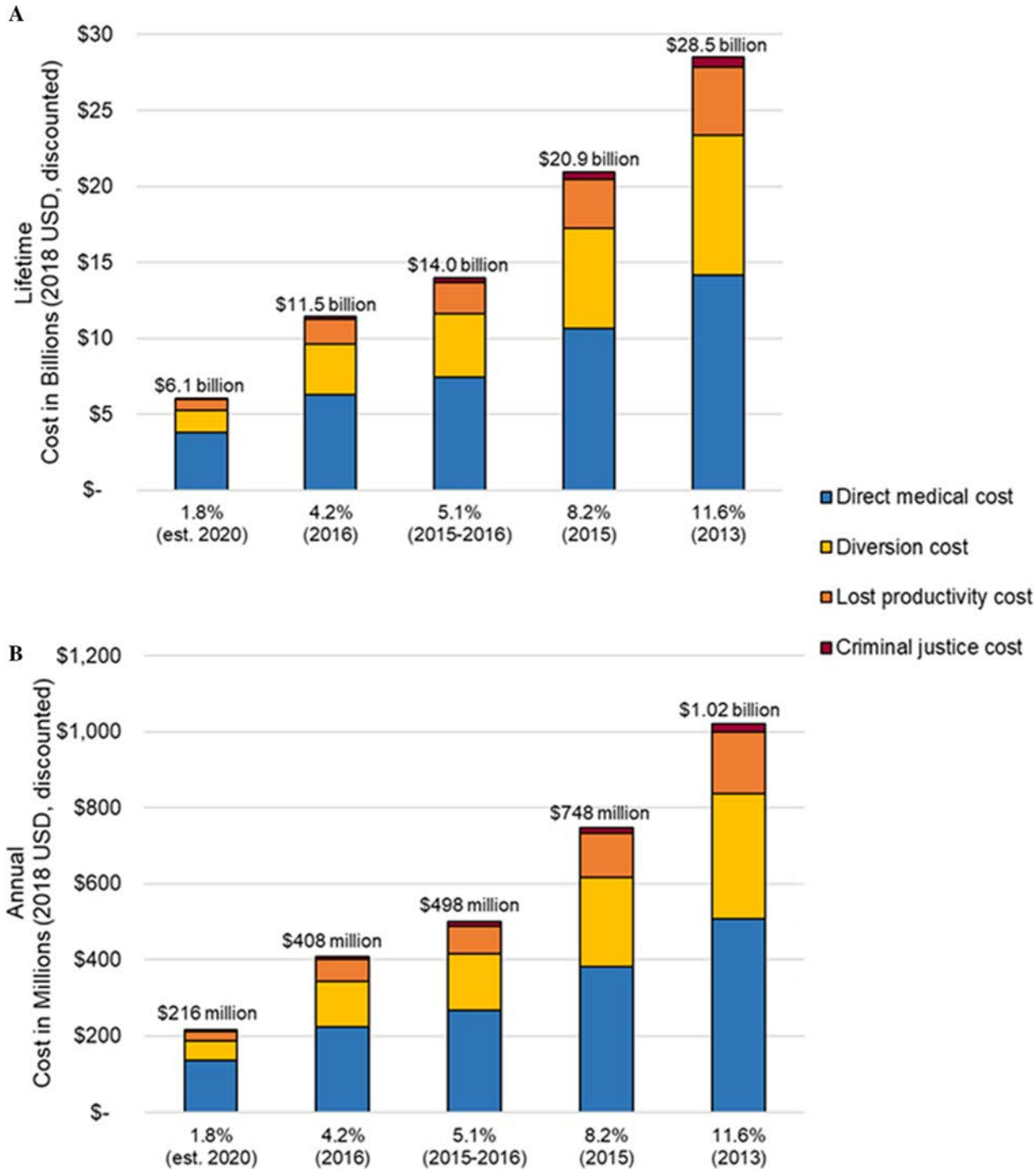


Figure 1. Percent of knee osteoarthritis (OA) population using strong opioids. **A**, Total lifetime opioid-related cost. **B**, Total annual opioid-related cost. The percent of the knee OA population using strong opioids varies at 1.8%, 4.2%, 5.1% (base case), 8.2%, and 11.6%, with values based on data from the years 2020 (estimate), 2016, 2015–2016, 2015, and 2013, respectively, of the Medicare Current Beneficiary Survey. The total lifetime (**A**) and annual (**B**) opioid-related cost for the knee OA population increases as strong opioid utilization increases, as does each component of the total cost: direct medical, diversion, lost productivity, and criminal justice cost. The total opioid-related cost is labeled above each bar.

start, the cohort had mean ± SD age of 54 ± 14 years, mean ± SD BMI of 30.5 ± 6.8 kg/m^2, and mean ± SD WOMAC pain score of 29 ± 17 (see Supplementary Table 13) (25).

OA progression. We estimated the annual probability of progressing to a higher Kellgren/Lawrence grade, stratified by sex and obesity status, using published data from the Johnston

County Osteoarthritis Project calibrated to published literature (28). We modeled pain trajectories derived from Osteoarthritis Initiative data (see section 1 of Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581).

Knee OA treatment characteristics. Details of the derivation of input parameters for the oral nonsteroidal antiinflammatory drugs (NSAIDs)/physical therapy/braces, glucocorticoid injection, and TKA regimens have been published (25). We modeled intermittent analgesic use (ibuprofen, acetaminophen, opioids) to control pain flares (annual cost $65) when patients did not respond to 1 intervention but had not yet started the next (29–32). Below, we highlight data pertinent to opioid use. Section 3 of Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581, describes relevant derivations.

Opioids. In this analysis, we modeled a scenario in which opioid-using patients start with weak opioids (e.g., tramadol) and those whose pain is not controlled proceed to strong opioids (e.g., oxycodone). Only subjects with a WOMAC pain score of >40 were eligible for opioid interventions. Using data from the MCBS, we estimated that 5.13% of knee OA patients use strong opioids on a prolonged basis (>5 prescriptions annually of fentanyl, hydrocodone/acetaminophen, hydromorphone, morphine, oxycodone, or oxycodone/acetaminophen) (14).

The efficacy, adverse events, and costs of opioids (weak and strong) are presented in Supplementary Table 13, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581. We estimated the decrement in pain during opioid use from a published meta-analysis, modeling weak opioids after tramadol and strong opioids after oxycodone and hydromorphone (17). We assumed that the annual likelihood that weak and strong opioids would stop providing pain relief was 12% and 24%, respectively (33). Based on expert clinical opinion, we assumed that subjects would visit their health care provider 2 and 6 times annually while using weak and strong opioids, respectively (21). We estimated the annual pharmaceutical cost of weak and strong opioids using data for tramadol and oxycodone, respectively. We modeled the annual cost of provider visits and pharmaceuticals to be $671 and $1,381 while using weak and strong opioids, respectively (29–31,34).

Weak opioid use carried a risk of fracture, excess all-cause mortality, somnolence, nausea, and constipation (21,35). Strong opioid use carried these same toxicities, as well as cardiovascular events (myocardial infarction, stroke, congestive heart failure, and cardiac death), OUD, and opioid overdose (18,21,35). The probabilities and costs of these toxicities are reported in Supplementary Table 13, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581. Toxicity probabilities were estimated using published data and data from the MCBS, as outlined in section 3 of Supplementary Appendix A, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581. Based on a published study of opioids, we modeled that 22% of subjects would discontinue opioid use due to toxicity in the first year (36); we assumed that these subjects would cease to incur opioid-related costs and would not subsequently be prescribed another opioid.

TKA following prolonged opioid use. The details of the TKA efficacy, adverse events, and costs used in OAPol have been published (25). In this analysis, based on published data, subjects who used opioids on a prolonged basis prior to TKA experienced an 8.9% lower decrement in pain following TKA (19). The rate of revision in the year following surgery was 58% higher among those who used opioids on a prolonged basis prior to TKA compared to those who did not receive opioid therapy (20). Based on published data, we modeled a 0.15% and 0.02% probability of opioid overdose in the year of TKA for those who used opioids on a prolonged basis prior to TKA and those who did not, respectively (20).

Nonmedical costs. *Costs of lost work productivity.* In this analysis, we considered the cost from lost wages due to knee OA pain, opioid use, OUD, and TKA. Based on published literature, we estimated that workers with a WOMAC pain score of >40 incurred an annual cost of $1,785 as a result of knee pain–related absenteeism and that workers with pain between 15 and 40 incurred 50% of this cost annually (see section 4.1 of Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581). We assumed a $0 cost for knee OA patients not participating in the labor market. In the simulation, subjects were assigned the average annual cost (across workers and nonworkers) for their age and pain group (see section 4.1 of Supplementary Appendix A and Supplementary Table 13, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581). Using the same methodology, we estimated the average cost in the year of primary and revision TKA to be $3,631 and $3,947, respectively (see section 4.2 of Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581), and the annual average cost during strong opioid use to be $687 (see section 4.3 of Supplementary Appendix A) (22). We modeled that those who experience OUD would incur the average 1-year cost associated with OUD in the US, which was $11,387 (6).

Criminal justice. OUD has been shown to lead to costs via property damage, police protection, legal proceedings, and imprisonment resulting from increased rates of drug law violations, including unlawful possession or distribution of narcotics, and other opioid-attributable crimes, such as burglary (6). We modeled criminal justice costs among SKOA patients with OUD as the average annual criminal justice cost resulting from OUD in the US. We estimated this value to be $4,264 (the reported national OUD-related criminal justice cost [$8.25 billion] divided by the number of persons experiencing OUD [1.935 million]) (6).

Diversion. We considered the cost of opioid misuse by persons who illicitly obtained opioids prescribed to SKOA patients. The 2013 National Survey on Drug Use and Health reported that 67.6% of persons misusing prescription pain relievers obtained their last pill from a friend or family member (37), which we assumed to be the percentage of misused opioids obtained through diversion. We calculated the cost of diversion as the national cost of OUD ($61.1 billion) (6) multiplied by the percentage of misused opioids obtained through diversion (67.6%) (37). Using these data and the estimated number of annual opioid prescriptions in the US (207 million) (1), we estimated the national annual cost of diversion ($41.3 billion) (6,37) and the cost of diversion per opioid prescription ($199). We estimated the average annual cost of diversion attributable to SKOA patients by multiplying the per-prescription diversion cost with the average number of annual strong opioid prescriptions (11.23) among prolonged opioid users with SKOA (14). Of the $2,241 total annual cost of opioid diversion from SKOA patients, $1,130, $808, and $303 were attributed to direct medical, lost work productivity, and criminal justice costs, respectively (1,6,14,37).

Population-level estimates. We estimated the total knee OA–attributable and opioid-attributable costs generated by SKOA patients in the US by multiplying the average cost per-person (predicted by the OAPol model) by the number of persons with SKOA in the US, estimated at 16.8 million using published methods (11) and data from the US Census Bureau (13) and National Health and Nutrition Examination Survey (38). We estimated annual costs as the lifetime cost divided by the disease duration predicted by the OAPol model.

Sensitivity analyses. To examine the robustness of our findings in the face of data uncertainty, we varied key model input parameters individually in sensitivity analyses (see section 5 of Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581). We varied the percentage of SKOA patients using opioids, considering opioid utilization rates from 2013 (11.6%), 2015 (8.2%), and 2016 (4.2%) (14), as well as further decline to 1.8%, the predicted utilization in 2020 (see section 5.1 of Supplementary Appendix A). We ran simulations wherein we used alternative methods to derive the cost of opioids, using data from the National Average Drug Acquisition Cost (NADAC) (39) and from the Centers for Medicare and Medicaid Services (CMS) prescription drug formulary (40). We estimated costs in the scenario wherein weak opioid use was associated with the same annual lost work productivity cost as strong opioid use. As an analysis reported that undiagnosed OUD cases may outnumber diagnosed cases 2 to 1 (41), we conducted analyses wherein subjects experienced undiagnosed OUD at twice the rate of diagnosed OUD. We assumed that undiagnosed OUD would result in the same cost as diagnosed OUD but would not cause subjects to cease receiving opioid prescriptions for knee OA. As the average criminal justice cost among persons with OUD was derived from a population younger than the SKOA population (6,28,42), we conducted a sensitivity analysis with reduced criminal justice cost of OUD and overdose to half the base-case value. We varied the number of opioid prescriptions diverted from strong opioid users annually within our base-case estimates' 95% confidence interval (8.8 to 13.7 prescriptions) (14). Additionally, we created the best, or least costly, scenario (NADAC-derived opioid cost, no undiagnosed OUD, lower estimate of diverted opioid prescriptions, and no annual lost productivity cost of weak opioids) and the worst, or costliest, scenario (CMS-derived opioid cost, undiagnosed OUD, annual lost productivity cost of weak opioids, and highest estimate of diverted opioid prescriptions). As the prevalence of SKOA may be higher (43), we estimated national costs using a 50% higher prevalence of SKOA.

RESULTS

Lifetime opioid-related cost per SKOA opioid user. We estimated a total lifetime opioid-related cost per opioid user in the knee OA population of $13,770 (Table 1). Direct medical

Table 1. Lifetime opioid-related and knee OA overall cost*

		For knee OA population		Per knee OA patient, regardless of opioid use		
Costs	Per knee OA opioid user opioid-related, $†	Opioid-related, $B‡	OA overall, $B‡	Opioid-related, $	OA overall, $	Attribution of opioid-related to OA overall cost, %§
Direct medical	6,390 (46)	7.45 (53)	188 (57)	450	11,250	4
Lost productivity	2,300 (17)	2.04 (15)	137 (42)	120	8,210	1
Criminal justice	320 (2)	0.288 (2)	0.288 (0.1)	20	20	100
Diversion	4,760 (35)	4.19 (30)	4.19 (1)	250	250	100
Total, $	13,770	14.0 $B	330 $B	840	19,730	4

* Values are the cost (%) unless indicted otherwise. OA = osteoarthritis.
† Percent of the total cost of the column.
‡ Percent of the total cost of the column. Costs are in billions of US$ (2018 $).
§ Percent opioid-related cost of the OA overall cost for each knee OA subject.

Table 2. Annual opioid-related and OA overall cost*

Costs	Per knee OA patient, regardless of opioid use		For knee OA population	
	Opioid-related, $	OA overall, $	Opioid-related, $B†	OA overall, $B†
Direct medical	16	400	0.266	6.71
Lost productivity	4	290	0.073	4.90
Criminal justice	1	1	0.010	0.010
Diversion	9	9	0.149	0.149
Total	30	700	0.498	11.8

* OA = osteoarthritis.
† Values are in billions of US$ (2018 $).

costs comprised 46% ($6,390) of opioid-related cost, costs due to lost work productivity comprised 17% ($2,300), diversion costs comprised 35% ($4,760), and criminal justice costs comprised 2% ($320).

Lifetime opioid-related and SKOA-related cost per SKOA patient. We estimated the average total lifetime opioid-related cost per SKOA patient (including both those who did and did not use opioids) at $840, which comprised 4% of the overall lifetime cost of SKOA per SKOA patient ($19,730) (Table 1). We estimated the direct medical cost of opioids at $450 per SKOA patient, contributing 4% of the direct medical cost of SKOA overall ($11,250). At $120 per SKOA patient, the opioid-attributable cost of lost work productivity comprised 1% of the cost of lost productivity of SKOA overall ($8,210). We estimated the diversion cost at $250 per SKOA patient and the criminal justice cost at $20.

Population-based lifetime opioid-related and SKOA-related cost. Assuming 5.1% of the SKOA population used strong opioids, the total lifetime opioid-related cost in the US SKOA population was estimated at $14.0 billion (Table 1). Of the opioid-related cost, $7.45 billion (53%) were due to direct medical costs, $2.04 billion (15%) to lost productivity costs, $4.19 billion (30%) to diversion costs, and $288 million (2%) to criminal justice costs. The national lifetime cost of SKOA was estimated at $330 billion: $188 billion of direct medical costs (including the cost of opioids and other SKOA treatments such as TKA), $137 billion of lost productivity costs, $4.19 billion of diverted opioid prescriptions, and $288 million of criminal justice costs (Table 1).

Annual opioid-related and SKOA-related cost for SKOA population and per SKOA patient. The total annual opioid-related cost for the knee OA population was estimated at $498 million per year (Table 2); the annual cost of SKOA overall was $11.8 billion. A total of $266 million of the annual opioid-related cost was due to the direct medical costs ($6.71 billion for SKOA overall), $73 million to the lost productivity costs ($4.90 billion SKOA overall), $149 million to the diversion costs (for opioid-related and SKOA overall), and $10 million to the criminal justice costs (for opioid-related and SKOA overall). The average annual opioid-related cost per SKOA patient, regardless of opioid use, was $30 and the average annual SKOA-related cost per SKOA was $700 (Table 2).

Sensitivity analyses. *Varying the prevalence of SKOA in the US population.* When we increased the prevalence of SKOA in the US by 50%, the total lifetime opioid-related cost for the knee OA population increased from $14.0 billion to $21.0 billion: $11.2 billion of direct medical cost, $3.06 billion of lost productivity, $431 million of criminal justice, and $6.29 billion of diversion. Additionally, the average lifetime SKOA-related cost increased from $330 billion to $495 billion.

Varying strong opioid utilization by SKOA patients. As we varied the percent of subjects using strong opioids from 1.8% (2020) to 11.6% (2013), the total lifetime opioid-related cost for

Table 3. Lifetime population cost of knee osteoarthritis overall by varying percent of population using strong opioids*

Costs	1.8% (est. 2020)†	4.2% (2016)	Base case 5.1% (2015–2016)	8.2% (2015)	11.6% (2013)
Direct medical	187	188	188	189	190
Lost productivity	137	137	137	138	138
Criminal justice	0.102	0.228	0.288	0.451	0.628
Diversion	1.42	3.30	4.19	6.59	9.22
Total	326	329	330	334	338

* Values are in billions of US$ (2018 $).
† Year of Medicare Current Beneficiary Survey for opioid utilization data.

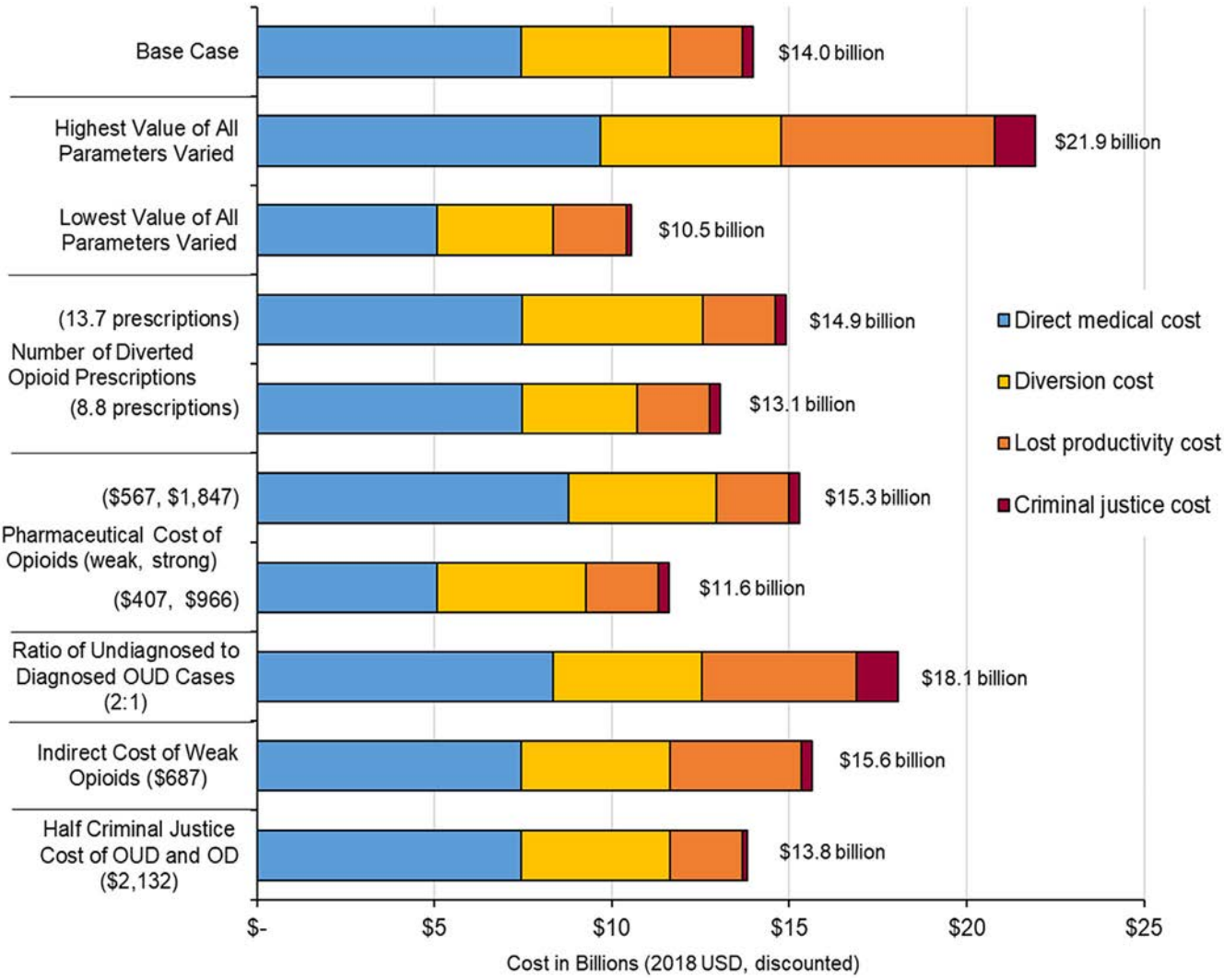


Figure 2. Sensitivity analysis of lifetime opioid-related cost in the knee osteoarthritis (OA) population. In all scenarios, 5.1% of the knee OA population used strong opioids. The base-case scenario included a 0:1 ratio of undiagnosed to diagnosed cases of opioid use disorder (OUD), 11.2 diverted strong opioid prescriptions per year, annual criminal justice of OUD and overdose of $4,264, annual pharmaceutical costs of weak opioids of $671 and strong opioids of $1,381, and annual lost productivity cost of weak opioids of $0. These parameters were varied individually. The cost was also calculated when the highest value of all parameters was used and the lowest value of all parameters was used. The total opioid-related cost is labeled next to each bar. OD = overdose.

the SKOA population increased from $6.1 billion to $28.5 billion (Figure 1A). The overall lifetime SKOA-attributable cost for the SKOA population increased from $326 billion to $338 billion as opioid utilization increased from 1.8% to 11.6%; the direct medical, diversion, lost productivity, and criminal justice costs each increased to varying degrees with the increase in opioid utilization (Table 3). The non-opioid related estimated lifetime cost of SKOA decreased from $320 billion to $309 billion as opioid use increased. The total annual opioid-related cost for the SKOA population increased from $216 million to $1.02 billion as opioid utilization increased from 1.8% to 11.6% (Figure 1B).

Varying prevalence of undiagnosed OUD, pharmaceutical cost of opioids, lost productivity cost of weak opioids, number of diverted opioid prescriptions, and criminal justice cost of OUD and overdose. Under the best-case scenario with the lowest value of all parameters varied, the total lifetime opioid-related cost generated by the knee OA population was estimated at $10.5 billion and the worst-case estimate with the highest value of all parameters varied at $21.9 billion (Figure 2). As we increased the ratio of undiagnosed to diagnosed cases of OUD from 0:1 to 2:1, the direct medical costs increased by 12%, the lost productivity costs doubled, and the criminal justice opioid-related costs quadrupled. The lost productivity opioid-related costs increased by 82% when the annual lost productivity cost was applied to weak opioids. The direct medical opioid-related costs increased by 18% when the annual pharmaceutical cost of weak and strong opioids was derived using NADAC data and decreased by 32% when the cost was derived using CMS data. The diversion costs

increased and decreased by 22% when the number of diverted opioid prescriptions was raised and lowered to the 95% confidence interval bounds. When we halved the criminal justice cost due to OUD and overdose, the estimated total lifetime opioid-related cost decreased by $200 million.

DISCUSSION

Although current guidelines do not recommend opioid use for knee OA pain management (15,16), we estimate that the US knee OA population generates $14.0 billion in opioid-related costs over their lifetime, approaching $0.5 billion annually. This amount represents 4% of all SKOA-related costs. The lost productivity, criminal justice, and diversion costs of opioid use accounted for almost half ($6.52 billion) of the total opioid-related population cost. Diverted opioid prescriptions alone comprised 35% of the lifetime opioid-related cost for each SKOA opioid user. As utilization of strong opioids among the SKOA population decreased from 11.6% to 1.8% between 2013 and 2020, the total lifetime opioid-related cost decreased from $28.5 billion to $6.1 billion.

Though opioids are used frequently for noncancer pain and knee OA is a prevalent condition in the US population (11,44), there is a paucity of literature examining the cost of prolonged opioid use in the knee OA population. Use of opioids has been associated with higher health care costs in a general OA population (22). The accumulating evidence on the lack of cost-effectiveness of opioids has prompted the question of whether to continue prescribing opioids for knee OA (21). In addition to being costly, prolonged opioid use has been associated with worse outcomes of TKA, an effective treatment for SKOA (19,20). Current guidelines for knee OA treatment do not recommend opioid use (15,16). Our report furthers evidence of the substantial societal cost generated by prescription opioid use and misuse and details the lifetime and annual costs of the knee OA population while incorporating the lost productivity, diversion, and criminal justice costs.

The results of our analysis should be interpreted within the context of its limitations. To portray multiple aspects of knee OA natural history, we used multiple data sources, as no single data source captured all data needed for the model. In the OAPol model, subjects proceed sequentially through treatments, although this pattern may not reflect clinical reality. Our opioid regimen input data were derived from short-term studies; thus, we assumed the durability of weak opioids to be similar to that of NSAIDs and strong opioids to be twice that of NSAIDs. Similarly, we made assumptions about the long-term rates of opioid-related adverse events (see section 3.4 of Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24581). We estimated the percentage of SKOA patients who use strong opioids and the rates of OUD and overdose among strong opioid users from the MCBS, cross-sectional data including only SKOA patients age >65 years; we extended these data to SKOA patients age <65 years. Our estimated rate of OUD does not include undiagnosed cases. Our estimated rate of overdose is not specific to the knee OA population. We do not model a scenario in which OUD increases the risk of mortality in the absence of overdose. We assume that those who overdose will cease to receive opioid prescriptions, although this assumption may not reflect reality (45). The lost productivity, criminal justice, and diversion costs of OUD and overdose were derived from the published cost of opioid misuse in the entire US. Our estimates of lost productivity costs are likely conservative, as we did not consider other knee OA–related productivity losses, such as those attributable to relatives who assist knee OA patients with activities of daily living.

Our results provide additional evidence of the substantial economic burden of opioid use for knee OA pain management and the potential savings from preventing opioid use, supporting guidelines recommending against such use. Reducing opioid use may lead to lower opioid-related and overall knee OA costs for knee OA opioid users and the US SKOA population. We found that opioid-related costs decrease when there are fewer diverted opioid prescriptions and the pharmaceutical cost of opioids is lower. If there is a higher number of undiagnosed OUD cases or if weak opioids carry a lost productivity cost, the cost for the knee OA population will be larger. These parameters were varied in our report to better understand the uncertainty of our estimates but also may be areas to target in reducing the opioid-related cost in the US.

Taking into consideration the ongoing opioid crisis when examining opioid use in the US knee OA population is important; drug overdoses are currently the largest cause of accidental death in the US (4,5). Opioid prescriptions for musculoskeletal pain have substantially contributed to opioid misuse in the US (46). Encouragingly, the prevalence of opioid use among those with knee OA has decreased each year from 2013 to 2016 (23), resulting in an estimated $20 billion decrease in lifetime knee opioid-related costs. These data document the substantial impact that can be achieved by providing alternatives to opiates for knee OA pain management. Several centrally acting regimens, such as duloxetine and gabapentin, as well as biologics such as tanezumab, have been examined as alternatives to opiates in the treatment of knee osteoarthritis (47,48). Further research, which may include review of electronic medical records and registries of opioids prescription, should be performed to identify the best alternatives, and we should work to educate health care providers and knee OA patients on the use of opiates 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. Losina 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. Huizinga, Stanley, Sullivan, Edwards, Losina.

Acquisition of data. Huizinga, Stanley, Song, Losina.

Analysis and interpretation of data. Huizinga, Stanley, Hunter, Paltiel, Neogi, Edwards, Katz, Losina.

ROLE OF THE STUDY SPONSOR

Pfizer had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Pfizer.

REFERENCES

1. Senate Caucus on International Narcotics Control hearing on America's addiction to opioids: heroin and prescription drug abuse, 113th Congress, 2nd Session (2014). Testimony of Nora D. Volkow, Maryland. URL: https://archives.drugabuse.gov/testimonies/2014/americas-addiction-to-opioids-heroin-prescription-drug-abuse.
2. Centers for Disease Control and Prevention. Vital signs: overdoses of prescription opioid pain relievers: United States, 1999–2008. MMWR Morb Mortal Wkly Rep 2011;60:1487–92.
3. Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: results from the 2018 National Survey on Drug Use and Health. Rockville (MD): Center for Behavioral Health Statistics and Quality; 2019.
4. Heron M. Deaths: leading causes for 2017. Natl Vital Stat Rep 2019; 68:1–77.
5. Hedegaard H, Minino AM, Warner M. Drug overdose deaths in the United States, 1999-2018. NCHS data brief, no 356. Hyattsville (MD): National Center for Health Statistics; 2020.
6. Florence CS, Zhou C, Luo F, Xu L. The economic burden of prescription opioid overdose, abuse, and dependence in the United States, 2013. Med Care 2016;54:901–6.
7. Fleming MF, Balousek SL, Klessig CL, Mundt MP, Brown DD. Substance use disorders in a primary care sample receiving daily opioid therapy. J Pain 2007;8:573–82.
8. Inciardi JA, Surratt HL, Lugo Y, Cicero TJ. The diversion of prescription opioid analgesics. Law Enforc Exec Forum 2007;7:127–41.
9. Stokes A, Berry KM, Hempstead K, Lundberg DJ, Neogi T. Trends in prescription analgesic use among adults with musculoskeletal conditions in the United States, 1999-2016. JAMA Netw Open 2019;2: e1917228.
10. Hootman JM, Cisternas M, Murphy L, Losby J. Prevalence and trends in prescribed opioid use among US adults with arthritis, 2008–2013, Medicare Expenditure Panel Survey [abstract]. Annual Scientific Meeting of the American College of Rheumatology, Washington, DC. November 2016.
11. 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.
12. DeMik DE, Bedard NA, Dowdle SB, Burnett RA, McHugh MA, Callaghan JJ. Are we still prescribing opioids for osteoarthritis? J Arthroplasty 2017;32:3578–82.
13. US Census Bureau, Population Division. Annual estimates of the resident population by sex, age, race, and Hispanic origin for the United States and states: April 1, 2010 to July 1, 2018. URL: https://www.census.gov/data/tables/time-series/demo/popest/2010s-national-detail.html.
14. Centers for Medicare and Medicaid Services. Medicare current beneficiary survey (MCBS) 2016. Baltimore, MD. URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS.
15. 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.
16. Kolasinski SL, Neogi T, Hochberg MC, Oatis C, Guyatt G, Block J, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis Rheumatol 2020;72:149–62.
17. Smith SR, Deshpande BR, Collins JE, Katz JN, Losina E. Comparative pain reduction of oral non-steroidal anti-inflammatory drugs and opioids for knee osteoarthritis: systematic analytic review. Osteoarthritis Cartilage 2016;24:962–72.
18. Solomon DH, Rassen JA, Glynn RJ, Lee J, Levin R, Schneeweiss S. The comparative safety of analgesics in older adults with arthritis. Arch Intern Med 2010;170:1968–76.
19. Smith SR, Bido J, Collins JE, Yang H, Katz JN, Losina E. Impact of preoperative opioid use on total knee arthroplasty outcomes. J Bone Joint Surg Am 2017;99:803–8.
20. Kim SC, Jin Y, Lee YC, Lii J, Franklin PD, Solomon DH, et al. Association of preoperative opioid use with mortality and short-term safety outcomes after total knee replacement. JAMA Netw Open 2019;2: e198061.
21. Smith SR, Katz JN, Collins JE, Solomon DH, Jordan JM, Suter LG, et al. Cost-effectiveness of tramadol and oxycodone in the treatment of knee osteoarthritis. Arthritis Care Res (Hoboken) 2017;69:234–42.
22. Zhao X, Shah D, Gandhi K, Wei W, Dwibedi N, Webster L, et al. The association of pain interference and opioid use with healthcare utilization and costs, and wage loss among adults with osteoarthritis in the United States. J Med Econ 2019;22:1192–201.
23. Losina E, Song S, Katz J. Opioids use among medicare beneficiaries with knee osteoarthritis: prevalence and correlates of chronic use [abstract]. Annual Scientific Meeting of the American College of Rheumatology, Atlanta, GA. November 2019.
24. Silva GS, Sullivan JK, Katz JN, Messier SP, Hunter DJ, Losina E. Long-term clinical and economic outcomes of a short-term physical activity program in knee osteoarthritis patients. Osteoarthritis Cartilage 2020;28:735–43.
25. Losina E, Paltiel AD, Weinstein AM, Yelin E, Hunter DJ, Chen SP, et al. Lifetime medical costs of knee osteoarthritis management in the United States: impact of extending indications for total knee arthroplasty. Arthritis Care Res (Hoboken) 2015;67:203–15.
26. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
27. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
28. Losina E, Walensky RP, Reichmann WM, Holt HL, Gerlovin H, Solomon DH, et al. Impact of obesity and knee osteoarthritis on morbidity and mortality in older Americans. Ann Intern Med 2011;154: 217–26.
29. Red Book Online. Truven Health Analytics Inc. 2019. URL: https://www.ibm.com/products/micromedex-red-book.
30. US Department of Veterans Affairs. Federal Supply Schedule. 2019. URL: https://www.fss.va.gov/index.asp.
31. Coalition for Community Pharmacy Action. The cost of dispensing study: an independent comparative analysis of U.S. prescription

dispensing costs. National Community Pharmacists Association, 2015. URL: https://www.nacds.org/pdfs/pharmacy/2020/NACDS-NASP-NCPA-COD-Report-01-31-2020-Final.pdf.

32. Bronnenberg BJ, Dubé JP, Gentzkow M, Shapiro JM. Do pharmacists buy Bayer? Informed shoppers and the brand premium. Quarterly J Econ 2015;130:1669–726.
33. Scott DL, Berry H, Capell H, Coppock J, Daymond T, Doyle DV, et al. The long-term effects of non-steroidal anti-inflammatory drugs in osteoarthritis of the knee: a randomized placebo-controlled trial. Rheumatology (Oxford) 2000;39:1095–101.
34. Centers for Medicare and Medicaid Services. Medicare fee schedules. URL: https://www.cms.gov/Medicare/Medicare.html.
35. Zeng C, Dubreuil M, LaRochelle MR, Lu N, Wei J, Choi HK, et al. Association of tramadol with all-cause mortality among patients with osteoarthritis. JAMA 2019;321:969–82.
36. Moore RA, McQuay HJ. Prevalence of opioid adverse events in chronic non-malignant pain: systematic review of randomised trials of oral opioids. Arthritis Res Ther 2005;7:R1046–51.
37. Substance Abuse and Mental Health Services Administration. Results from the 2013 National Survey on Drug Use and Health: summary of national findings. National Survey on Drug Use and Health Series H-48, HHS Publication No. (SMA) 14-4863. Rockville (MD): Substance Abuse and Mental Health Services Administration; 2014.
38. National Center for Health Statistics. National Health and Nutrition Examination Survey 2007-2008. Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, Hyattsville, MD. URL: https://wwwn.cdc.gov/nchs/nhanes/Default.aspx.
39. Centers for Medicare and Medicaid Services. National average drug acquisition cost, March 2020. URL: https://data.medicaid.gov/dataset/c933dc16-7de9-52b6-8971-4b75992673e0.
40. 46brooklyn Research. Shred the gnarly Medicare part D pricing slopes. URL: https://www.46brooklyn.com/part-d-ski-slope.
41. Kirson NY, Shei A, Rice JB, Enloe CJ, Bodnar K, Birnbaum HG, et al. The burden of undiagnosed opioid abuse among commercially insured individuals. Pain Med 2015;16:1325–32.
42. Birnbaum HG, White AG, Schiller M, Waldman T, Cleveland JM, Roland CL. Societal costs of prescription opioid abuse, dependence, and misuse in the United States. Pain Med 2011;12:657–67.
43. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum 2008;58:26–35.
44. Anastassopoulos KP, Chow W, Tapia CI, Baik R, Ackerman SJ, Biondi D, et al. Economic study on the impact of side effects in patients taking oxycodone controlled-release for noncancer pain. J Manag Care Pharm 2012;18:615–26.
45. LaRochelle MR, Liebschutz JM, Zhang F, Ross-Degnan D, Wharam JF. Opioid prescribing after nonfatal overdose and association with repeated overdose. Ann Internal Med 2016;164:1–9.
46. Bonnie RJ, Schumacher MA, Clark JD, Kesselheim AS. Pain management and opioid regulation: continuing public health challenges. Am J Public Health 2019;109:31–4.
47. Enteshari-Moghaddam A, Azami A, Isazadehfar K, Mohebbi H, Habibzadeh A, Jahanpanah P. Efficacy of duloxetine and gabapentin in pain reduction in patients with knee osteoarthritis. Clin Rheumatol 2019;38:2873–80.
48. Lane NE, Schnitzer TJ, Birbara CA, Mokhtarani M, Shelton DL, Smith MD, et al. Tanezumab for the treatment of pain from osteoarthritis of the knee. N Engl J Med 2010;363:1521–31.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1359–1368
DOI 10.1002/acr.24564

American College of Rheumatology
Empowering Rheumatology Professionals

Application of Heterogeneity of Treatment-Effects Methods: Exploratory Analyses of a Trial of Exercise-Based Interventions for Knee Osteoarthritis

Cynthia J. Coffman,[1] Liubov Arbeeva,[2] Todd A. Schwartz,[3] Leigh F. Callahan,[2] Yvonne M. Golightly,[2] Adam P. Goode,[4] Kim M. Huffman,[5] and Kelli D. Allen[6]

Objective. To evaluate heterogeneity of treatment effects in a trial of exercise-based interventions for knee osteoarthritis (OA).

Methods. Participants (n = 350) were randomized to standard physical therapy (PT; n = 140), internet-based exercise training (IBET; n = 142), or wait list (WL; n = 68) control. We applied qualitative interaction trees (QUINT), a sequential partitioning method, and generalized unbiased interaction detection and estimation (GUIDE), a regression tree approach, to identify subgroups with greater improvements in Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) score over 4 months. Predictors included 24 demographic, clinical, and psychosocial characteristics. We conducted internal validation to estimate optimism (bias) in the range of mean outcome differences among arms.

Results. Both QUINT and GUIDE indicated that for participants with lower body mass index (BMI), IBET was better than PT (improvements of WOMAC ranged from 6.3 to 9.1 points lower), and for those with higher BMI and a longer duration of knee OA, PT was better than IBET (WOMAC improvement was 6.3 points). In GUIDE analyses comparing PT or IBET to WL, participants not employed had improvements in WOMAC ranging from 1.8 to 6.8 points lower with PT or IBT versus WL. From internal validation, there were large corrections to the mean outcome differences among arms; however, after correction, some differences remained in the clinically meaningful range.

Conclusion. Results suggest there may be subgroups who experience greater improvement in symptoms from PT or IBET, and this finding could guide referrals and future trials. However, uncertainty persists for specific treatment-effects size estimates and how they apply beyond this study sample.

INTRODUCTION

Physical therapy (PT) and exercise-based interventions are core components of knee osteoarthritis (OA) treatment (1,2). However, overall effects of these interventions tend to be modest, with substantial variability across patients (3–5). Patients with OA differ substantially from one another in clinical, biomechanical, and psychosocial characteristics that can impact the effectiveness of exercise-based interventions (6). In addition, there are many different types of exercise-based interventions that vary in terms of intensity, duration, delivery mode, amount of supervision, exercise types, and physiologic targets (3). There is little

ClinicalTrials.gov identifier: NCT02312713.

The statements and opinions presented in this article are solely the responsibility of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute, its Board of Governors, or the Methodology Committee.

Supported by a Patient-Centered Outcomes Research Institute Award (CER-1306-02043). The work of Drs. Coffman and Allen was supported by the Center of Innovation to Accelerate Discovery and Practice Transformation (CIN 13-410) at the Durham Veterans Administration Health Care System. The work of Drs. Arbeeva, Schwartz, Callahan, Golightly, Goode, and Allen was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases Core Center for Clinical Research (grant P30-AR-072580). Dr. Allen's work was also supported by a Veterans Administration Health Services Research and Development Research Career Scientist Award (grant 19-332).

[1]Cynthia J. Coffman, PhD: Durham Veterans Administration Healthcare System and Duke University Medical Center, Durham, North Carolina; [2]Liubov Arbeeva, MS, Leigh F. Callahan, PhD, Yvonne M. Golightly, PT, MS, PhD: University of North Carolina at Chapel Hill; [3]Todd A. Schwartz, DrPH: Gillings School of Global Public Health, University of North Carolina at Chapel Hill; [4]Adam P. Goode, PT, DPT, PhD: Duke University School of Medicine and Duke University Medical Center, Durham, North Carolina; [5]Kim M. Huffman, MD, PhD: Duke University Medical Center, Durham, North Carolina; [6]Kelli D. Allen, PhD: University of North Carolina at Chapel Hill and Durham Veterans Administration Healthcare System, Durham, North Carolina.

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

Address correspondence to Kelli D. Allen, PhD, University of North Carolina at Chapel Hill, Durham Veterans Administration Healthcare System, CB 7280, UNC, Chapel Hill, NC 27599. Email: kdallen@email.unc.edu.

Submitted for publication May 15, 2020; accepted in revised form January 14, 2021.

SIGNIFICANCE & INNOVATIONS

- This study is the first to use statistical methods of qualitative interaction trees (QUINT) and generalized unbiased interaction detection and estimation (GUIDE) to examine heterogeneity of treatment effects for different exercise-based treatments among individuals with knee osteoarthritis (OA).
- Both QUINT and GUIDE indicate that for participants with lower body mass index, an internet-based training program (IBET) was better than physical therapy (PT); for those with higher body mass index and longer duration of knee OA symptoms, PT was better than IBET.
- GUIDE analysis indicated that participants who were not employed had greater improvements with PT or IBET, relative to usual care.
- In heterogeneity of treatment-effects analyses with small samples, internal validation provides guidance for interpreting and applying results.

understanding of which types of exercise-based interventions work best for different patients with OA. This gap in knowledge limits our ability to advise patients regarding the exercise-based intervention they may benefit from most, as well as limiting our ability to effectively target interventions in a population-based manner. Consistent with the goals of precision or personalized medicine (7), the OA community needs to develop an understanding of the "right treatment for the right patient at the right time," in the context of exercise-based interventions, to maximize effectiveness.

Exploratory analyses of previous trials provide some evidence that responses to exercise-based interventions for OA may vary according to patient characteristics such as age, sex, pain severity, strength, function, malalignment, radiographic severity, and psychological variables (8–12). However, those analyses have focused on a limited set of potential moderators, since the typical statistical approach of adding interaction terms limits inclusion of a large number of candidate variables. In addition, evaluating potential treatment moderators singly may fail to identify important combinations of variables (13–15). For example, a given exercise-based intervention may be beneficial for older adults who have low strength levels and low-to-moderate pain severity, or effects of specific interventions may differ based on different OA phenotypes (16). New data-driven methods allow exploration of multidimensional subgroups that exhibit heterogeneous treatment effects (15,17–20), and these methods can deepen our understanding of patients' responses to exercise-based interventions.

We recently completed the Physical Therapy versus Internet-Based Exercise Training for Patients with Knee Osteoarthritis (PATH-IN) randomized clinical trial that compared PT and an internet-based exercise training (IBET) program, both relative to a wait list (WL) control group (21,22). We found that the effects of PT and IBET were similar to each other and did not differ significantly from WL for the primary outcome of Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) total score in the overall sample. However, a prespecified aim of the trial was to evaluate heterogeneity of treatment effects to understand whether either PT or IBET may have benefits for subgroups of patients compared to each other or to WL. In this article, we applied 2 different advanced statistical methods to explore heterogeneity of treatment effects in the PATH-IN study.

MATERIALS AND METHODS

The PATH-IN study randomized individuals with knee OA to standard PT (n = 140), IBET (n = 142), or WL (n = 68) in a 2:2:1 allocation (21,22). The study was approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Overview of heterogeneity of treatment-effects methods. Recursive partitioning methods are the underpinning for many data-driven heterogeneity of treatment-effects methods (15,17,23). The basic idea is to create trees that classify patients into subgroups based on independent variables, where treatment-effects sizes are large, are in opposite directions, or meet some difference threshold. The methods are recursive because each subpopulation may be split again until some stopping criterion is reached.

When selecting from among heterogeneity of treatment-effects methods, careful attention must be paid to the specific research question being addressed (15,17–20). The first question we addressed dealt with 2 active treatments (PT and IBET), exploring which treatment worked better for whom. This approach is known as a qualitative subgroup interaction, where one treatment may work better for one subgroup, while another treatment may be better for another (18,24). Qualitative interaction trees (QUINT) is a sequential partitioning method that identifies whether or not qualitative subgroup effects are present, and, if so, partitions the sample into 3 potential subgroups: treatment A is better than B, treatment B is better than A, or neither treatment is better (18,24). The second question we addressed dealt with which subgroups showed the greatest improvement, relative to a control group. Exploring which treatment works better for whom is known as a quantitative subgroup interaction, which occurs when a treatment produces large improvements in outcomes for some patients but little to no improvement for others (19,20). We explored which subgroups showed the greatest WOMAC score improvement in IBET compared to PT. We also explored which subgroups showed the greatest WOMAC score improvement compared to WL, including PT, IBET, and WL in 1 model. Generalized unbiased interaction detection and estimation (GUIDE) is a regression tree approach that identifies whether or not quantitative subgroup effects are present, and,

if so, partitions the sample into subgroups with differential treatment effects (20).

Predictor variable selection. The potential predictors were collected at baseline, prior to randomization. We selected the most relevant variables based on our experience and evidence from previous studies (5,8–12). Following examination of missing data and correlations between predictor variables, 24 candidate variables were selected (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564). Due to missing data in included covariates, we excluded n = 5 participants; the final sample was: PT arm (n = 138), IBET arm (n = 140), and WL control (n = 67).

Heterogeneity of treatment-effects methods. The outcome for our analyses was change from baseline in WOMAC total score (the primary study outcome) at 4 months; a negative change indicates improvement in WOMAC. At 4 months, 45 participants missed follow-up assessment; we used empirical best linear unbiased prediction estimates from linear mixed-effects models as a single imputation for the 4-month outcome and then calculated the change score (25).

We first applied QUINT (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564, for details on the algorithm), which is implemented with the package quint in R software, version 3.5.1 (26). In our analyses, we used the "difference in means" option for the outcome and the default options for partitioning criteria: minimum absolute effect size of 0.3 and equal weighting of effect size difference and cardinality component for determining splits (18). We set the minimum sample size per treatment arm per subgroup at 15 (i.e., total minimum subgroup n = 30), which is close to the default option of 10%. We also ran an analysis in which we increased the minimum sample size per treatment arm per subgroup to 20 (for a total n = 40). Finally, for all analyses, we used the prune.quint function to reduce overfitting and to select the optimal tree with the optimal number of subgroups. The number of bootstrap samples was 25, and as a sensitivity analysis we also set the number of bootstraps to 100; results were similar.

We then applied GUIDE (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564, for details on the algorithm), using the subgroup identification approach (G_i option), which is implemented with the package GUIDE (http://pages.stat.wisc.edu/~loh/guide.html). We set the minimum sample size per treatment arm per subgroup to n = 15 and n = 20 in respective iterations of the algorithm. We used GUIDE in an analysis that included only the 2 active treatments, and also in a second analysis that included all 3 treatment arms. Pruning in GUIDE was applied with cross validation (19).

The conclusion of a QUINT implementation yields values of the baseline variables that define the subgroups, with mean differences and sample sizes for each of the treatments in each subgroup. An implementation of GUIDE yields the values of the baseline variables that define the subgroups, estimated mean differences between treatment groups adjusted for covariates in the model, and sample sizes for each treatment in each subgroup. The mean difference (4-month WOMAC score minus baseline WOMAC score) between treatment groups for both QUINT and GUIDE when only 2 active arms are included is presented as IBET mean difference minus PT mean difference, where a negative value indicates greater improvement for IBET over PT, and a positive value indicates greater improvement for PT over IBET. The mean differences between treatment groups for GUIDE when all 3 arms are included are presented as IBET mean difference minus WL mean difference and as PT mean difference minus WL mean difference, where a negative value indicates greater improvement for IBET or PT compared to WL, and a positive value indicates greater improvement for WL compared to IBET or PT.

We then conducted an internal validation for both QUINT and GUIDE analyses via bootstrap resampling to estimate optimism or bias in the range of mean outcome differences between pairs of arms in the final selected tree (i.e., the apparent range) due to overfitting and to provide a bias-corrected estimate (27). We followed the steps as outlined in Section C.2 of the web appendix of Dusseldorp and Mechelen (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564, for details) (18).

RESULTS

QUINT results. Descriptive statistics for the 24 predictor variables, overall and by treatment arm, are shown in Table 1. Figure 1 shows the pruned tree (unpruned tree was the same) from QUINT when the minimum subgroup size was n = 40 total participants (i.e., at least n = 20 in each arm), which contains 4 subgroups. For the 2 subgroups that had greater improvement in IBET than PT (red), mean differences in WOMAC scores were 7.2 and 6.3 points lower in IBET compared to PT. The first subgroup (n = 44) was body mass index (BMI) ≤24.31 kg/m^2 and the second (n = 57) was defined by a combination of BMI (>24.31 kg/m^2), duration of OA symptoms (≤9.5) years, and Social Support for Exercise score (≤56.5 points). For the 2 subgroups with greater improvement in PT than IBET (green), mean differences in WOMAC scores were 3.1 points and 6.3 points lower in PT compared to IBET. The first subgroup (n = 50) was defined by a combination of BMI (>24.31 kg/m^2), duration of OA symptoms (≤9.5 years), and Social Support for Exercise score (>56.5 points), and the second subgroup (n = 127) was defined by BMI (>24.31 kg/m^2) and duration of OA symptoms (>9.5 years).

Table 1. Descriptive statistics for baseline patient characteristics and explanatory variables*

Characteristic	Total (n = 345)	PT (n = 138)	IBET (n = 140)	WL control (n = 67)
Age, years	65.3 ± 11	65.7 ± 10.3	65.1 ± 11.4	64.7 ± 11.7
Women, no. (%)	247 (71.6)	99 (71.7)	96 (68.6)	52 (77.6)
White, no. (%)	251 (72.8)	109 (79.0)	93 (66.4)	49 (73.1)
Married/living with partner, no. (%)	213 (61.7)	78 (56.5)	93 (66.4)	42 (62.7)
Bachelor's/postgraduate work, no. (%)	205 (59.4)	84 (60.9)	79 (56.4)	42 (62.7)
Fair or poor health, no. (%)	48 (13.9)	14 (10.1)	22 (15.7)	12 (17.9)
Financial status: comfortable/meet basic needs, a little left over, no. (%)	285 (82.6)	118 (85.5)	112 (80.0)	55 (82.1)
Employed full- or part-time, no. (%)	140 (40.6)	59.0 (42.8)	51 (36.4)	30 (44.8)
Body mass index, kg/m^2	31.3 ± 8.0	31.8 ± 8.6	31.5 ± 7.7	29.8 ± 6.8
No. of joints with OA symptoms	5.3 ± 3.2	5.5 ± 3.0	5.1 ± 3.1	5.4 ± 3.9
Duration of OA symptoms, years	13.1 ± 11.6	13.9 ± 11.5	11.6 ± 11.0	14.4 ± 12.9
History of knee injury, no. (%)	173 (50.1)	71 (51.4)	69 (49.3)	33 (49.3)
Problems learning†	3.7 ± 0.7	3.7 ± 0.6	3.6 ± 0.8	3.8 ± 0.5
Filling out forms‡	0.4 ± 0.7	0.3 ± 0.8	0.4 ± 0.8	0.3 ± 0.6
Internet comfort§	4.1 ± 1.2	4.1 ± 1.2	4.1 ± 1.3	4.1 ± 1.2
Internet frequency¶	1.4 ± 1.2	1.3 ± 1.1	1.5 ± 1.2	1.4 ± 1.3
WOMAC total score (0–96)	31.9 ± 17.8	32.0 ± 17.7	31.3 ± 17.7	33.1 ± 18.8
PHQ-8 score (0–24)	3.8 ± 4.1	4.0 ± 4.5	3.7 ± 4.1	3.5 ± 3.4
PROMIS fatigue score	51.2 ± 8.8	51.9 ± 9.1	50.3 ± 9.0	51.9 ± 7.9
Self-efficacy exercise score	56.2 ± 20.3	57.3 ± 20.8	56.8 ± 19.8	52.8 ± 20.5
Social support exercise score	52.1 ± 18.4	51.9 ± 17.4	51.8 ± 19.6	53.0 ± 18.3
30-second chair stand, no.	9.6 ± 3.9	9.5 ± 4.2	9.6 ± 3.7	9.6 ± 3.6
2-minute march test, no. of steps	50.9 ± 29.6	51.6 ± 31.0	51.5 ± 29.5	48.3 ± 26.8
Unilateral stand test, seconds	7.3 ± 3.6	7.3 ± 3.6	7.4 ± 3.4	6.8 ± 3.7

* Values are the mean ± SD unless indicated otherwise. IBET = internet-based exercise training; OA = osteoarthritis; PHQ-8 = Patient Health Questionnaire 8; PROMIS = Patient-Reported Outcomes Measurement Information System; PT = physical therapy; WL = wait list; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† Problems learning were assessed via questionnaire consisting of 1 question: "How often do you have problems learning about your medical condition because of difficulty understanding written information?" and reported as "always (0)," "often (1)," "sometimes (2)," "occasionally (3)," and "never (4)."
‡ Problems filling out forms were assessed via questionnaire consisting of 1 question: "How confident are you filling out forms by yourself?" and reported as "extremely (0)," "quite a bit (1)," "somewhat (2)," "a little bit (3)," and "not at all (4)."
§ Comfort using internet, Likert scale (1 = not at all; 5 = very).
¶ The frequency of internet use was reported as "every day (1)," "a few times a week (2)," "once a week (3)," "a few times a month (4)," "once a month (5)," "less than once a month (6)," and "not at all (7)."

Figure 2 shows the pruned tree (the unpruned tree had an additional split for subgroup 4) from QUINT when the minimum subgroup size was n = 30 total participants (i.e., at least n = 15 in each arm), which contains 5 subgroups. For the 3 subgroups that had greater improvement with IBET than PT (red), mean differences in WOMAC scores were 9.1, 8.0, and 4.9 points lower in IBET than PT. One subgroup (n = 38) was composed of individuals with lower BMI, with a cutoff similar to the QUINT analysis, with a minimum subgroup size of n = 40 (i.e., 23.94 kg/m^2). The second group (n = 33) had higher BMI (>23.94 kg/m^2) and younger age (≤55.54 years), and the third subgroup (n = 42) included older individuals (age >72 years) with better performance on chair stands (>8.5 stands). For the 2 subgroups that had greater improvement with PT than IBET (green), mean differences in WOMAC scores were 9.6 and 4.7 points lower in PT than IBET. One group (n = 86) included participants with higher BMI (>23.94 kg/m^2) and worse performance on the 30-second chair stand (≤8.5 stands), and the other group (n = 79) had higher BMI (>23.94 kg/m^2), age between 55.5 and 72.0 years, and better performance on the 30-second chair stand (>8.5 stands).

QUINT internal validation. For the QUINT analysis internal validation when the minimum subgroup size was n = 40 total participants (i.e., at least n = 20 in each arm), the apparent range, the difference between the largest negative difference in means between arms in a subgroup (–7.2), and largest positive difference in means between arms in a subgroup (6.3), was –13.5 (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564). When this apparent range (–13.5) was corrected for estimated optimism, it was reduced to –4.5, well below the –8.0 that we would deem as a clinically meaningful difference in WOMAC change (28). Similarly, for the QUINT procedure when the minimum subgroup size was n = 30 participants (at least n = 15 in each arm), we found a large reduction in the apparent range from –18.7 to –9.0 when corrected for estimated optimism (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564). However, this corrected apparent range of –9.0 was above the minimum range expected and was indicative of clinically meaningful difference in WOMAC change.

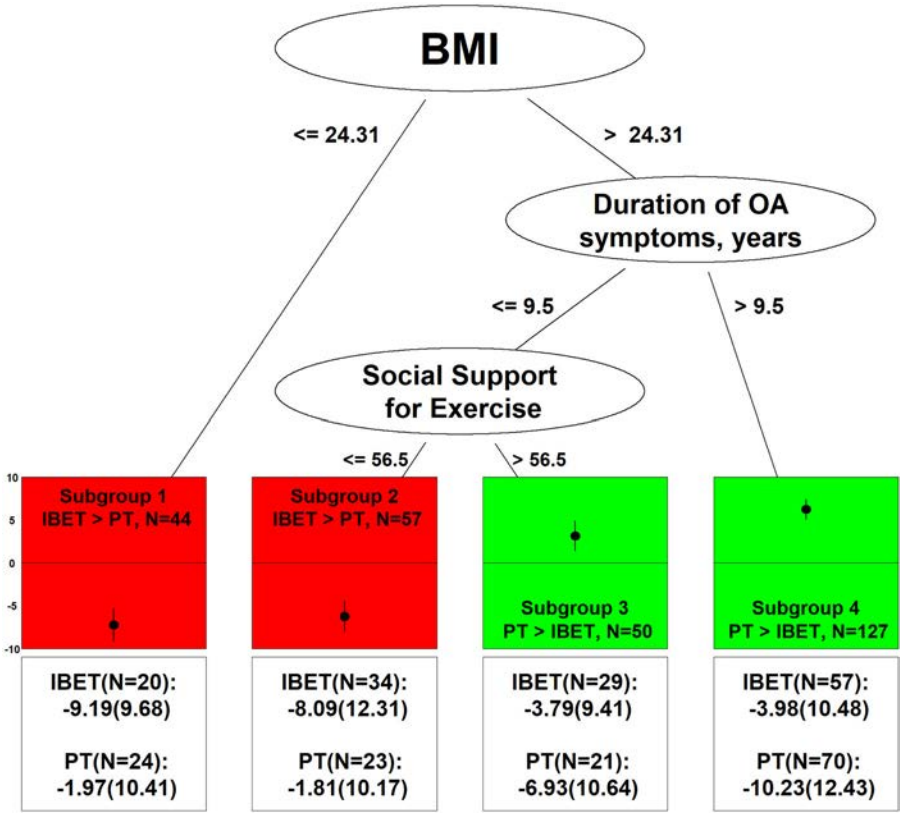


Figure 1. Qualitative interaction trees subgroups with 2 active arms (minimum sample size for subgroup is 40 total participants, i.e., at least 20 in each arm). Subgroup 1: body mass index (BMI) ≤24.31 kg/m^2 (n = 44; mean difference –7.2 points); subgroup 2: BMI >24.31 kg/m^2, osteoarthritis (OA) symptom duration ≤9.5 years, and Social Support for Exercise score ≤56.5 (n = 57; mean difference –6.3 points); subgroup 3: BMI >24.31 kg/m^2, OA symptom duration ≤9.5 years, and Social Support for Exercise score >56.5 (n = 50; mean difference 3.1 points); subgroup 4: BMI >24.31 kg/m^2, OA symptom duration >9.5 years (n = 127; mean difference 6.3 points). Dots indicate the mean difference IBET-PT; error bars show the SE of the difference. IBET = internet-based exercise training; PT = physical therapy. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564/abstract.

GUIDE results. Figure 3 shows the unpruned tree from GUIDE when applied to the 2 active arms when the minimum subgroup size was n = 40 total participants (i.e., at least n = 20 in each arm). The pruned tree was empty. Similar to QUINT, the tree contains 4 subgroups. However, 1 subgroup had greater improvement in WOMAC with IBET compared to PT (red), another subgroup had greater improvement in WOMAC with PT than IBET (green), and 2 subgroups showed no difference between IBET and PT (grey). The subgroup that had a larger improvement with IBET than PT (6.9 points lower, n = 59) was composed of participants with duration of OA symptoms ≤9.5 years and BMI ≤29.45 kg/m^2. The subgroup that had greater improvement with PT than IBET (5.7 points lower, n = 67) was composed of participants with duration of OA symptoms >18.5 years.

Figure 4 shows the unpruned tree from GUIDE when applied to all 3 treatment arms when the minimum subgroup size was n = 60 total participants (i.e., at least n = 20 in each arm). The pruned tree was empty. The tree contains 2 subgroups, where 1 subgroup had greater improvement in IBET and PT compared to WL, and the other subgroup had greater improvement in WL compared to IBET and PT. The subgroup (n = 205) that had lower mean differences in WOMAC for IBET (3.8 points) and PT (6.4 points), compared with WL, was composed of individuals not currently employed. The subgroup (n = 140) for which there were larger mean differences in WOMAC for WL compared to IBET (0.7 points) or PT (1.2 points) was composed of individuals currently employed. Figure 5 shows the unpruned tree from GUIDE including all 3 arms when the minimum subgroup size was n = 45 total participants (i.e., at least n = 15 in each arm); the pruned tree was empty. This unpruned tree contains 4 subgroups, with the first split variable of employment status and then the 2 employment status groups further subdivided to obtain the 4 subgroups. There were 2 subgroups for which IBET and PT both had lower mean differences than WL control by 5.7, 6.8, 1.8, and 5.8 points; 1 subgroup (n = 115) consisted of participants who were not currently employed, with duration of OA symptoms ≤10.5 years, and the other subgroup (n = 90) included participants who were not currently employed and had a duration of symptoms >10.5 years. For 1 subgroup (n = 64), mean differences were lower for IBET than WL control but greater for WL

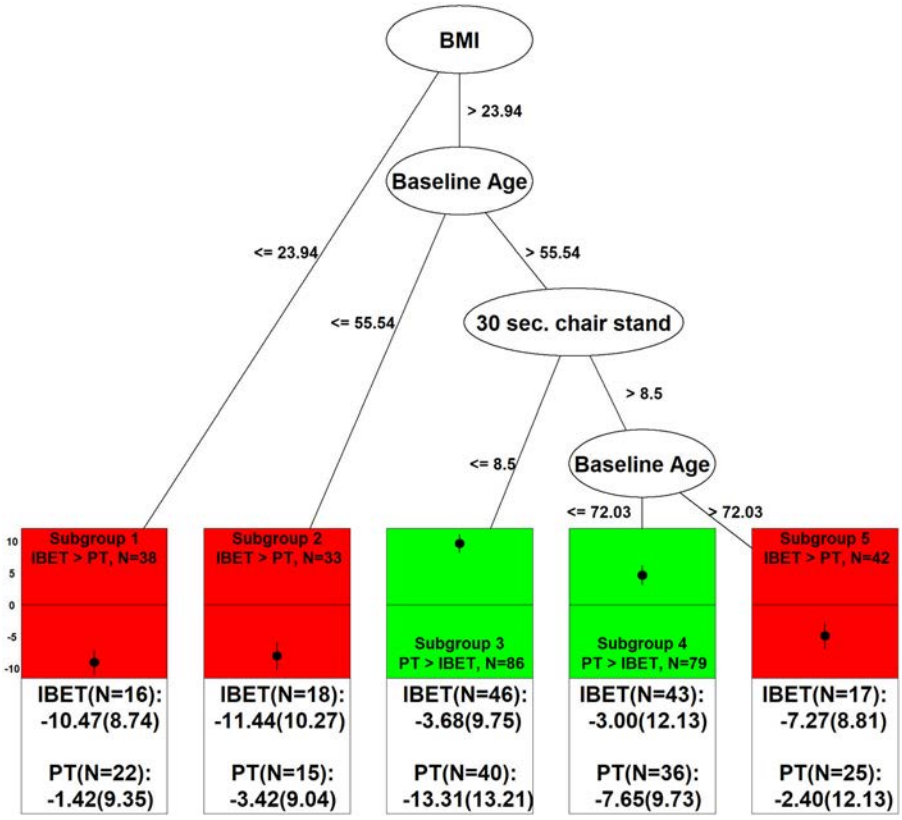


Figure 2. Qualitative interaction trees subgroups with 2 active arms (minimum sample size per subgroup is 30 total participants, i.e., at least 15 per arm). Subgroup 1: body mass index (BMI) ≤23.94 kg/m^2 (n = 38; mean difference –9.1 points); subgroup 2: BMI >23.94 kg/m^2, age ≤55.54 years (n = 33; mean difference –8.0 points); subgroup 3: BMI >23.94 kg/m^2, age >55.54 years, and number of chair stands ≤8.5 (n = 86; mean difference 9.6 points); subgroup 4: BMI >23.94 kg/m^2, age >55.54 and ≤72.03 years, and number of chair stands >8.5 (n = 79; mean difference 4.7 points), and subgroup 5: BMI >23.94 kg/m^2, age >72.03 years, and number of chair stands >8.5 (n = 42; mean difference –4.9 points). Dots indicate the mean difference IBET-PT; error bars show the SE of the difference. IBET = internet-based exercise training; PT = physical therapy. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564/abstract.

than PT, although these differences were small; this group included individuals who were currently employed and had lower scores on the chair stand test (≤9.5 stands). For the last subgroup (n = 76), mean differences were lower for WL than for either IBET (2.6 points) and PT (2.1 points); this subgroup was currently employed and had better chair stand scores (>9.5 stands).

GUIDE internal validation. For the GUIDE internal validation for 2 active arms, the apparent range was –12.6 (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564), and when corrected based on our estimate for optimism, it was reduced to –7.2. For the internal validation of the GUIDE procedure for 3 arms, comparing PT to WL when the minimum subgroup size was set to n = 45 total participants (i.e., at least n = 15 in each arm), the apparent range was –8.9 and with optimism correction reduced to –5.0. In all cases for GUIDE, there were large reductions of the apparent range after correcting for optimism, but many remained on the border of clinically meaningful differences (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564).

DISCUSSION

In this 3-arm study, we addressed several research questions related to heterogeneity of treatment effects for different exercise-based treatments among individuals with knee OA. We used QUINT to address qualitative subgroup interactions in the 2 active treatment arms, exploring which treatment worked better for which subgroups, and GUIDE to explore the more general question of whether some subgroups had larger improvements than others between the 2 active treatments. Based on results involving the 2 active arms, an overall observation was that BMI, age, and disease duration seemed to be important factors regarding whether PT or IBET yielded greater improvement. Although some other factors contributed to subgroup identification, these 3 easily assessed patient characteristics could help to guide referrals in clinical situations. In particular, these results suggest that patients who are older, have higher BMI, and have had

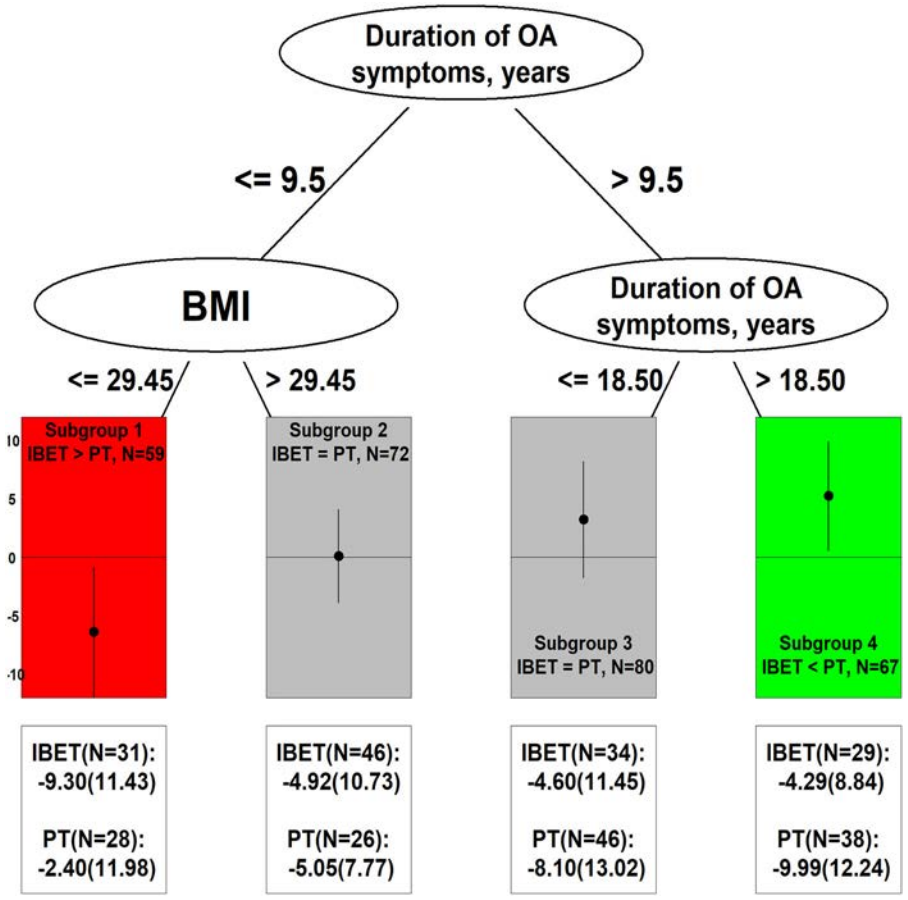


Figure 3. Generalized unbiased interaction detection and estimation subgroups with 2 active arms: internet-based exercise training (IBET) and physical therapy (PT) (n = 278; minimum sample size per subgroup is 40 total participants, i.e., at least 20 in each arm). Subgroup 1: osteoarthritis (OA) symptom duration ≤9.5 years and body mass index (BMI) ≤29.45 kg/m^2 (n = 59; unadjusted mean difference –6.9 points, adjusted mean difference –5.0); subgroup 2: OA symptom duration ≤9.5 years and BMI >29.45 kg/m^2 (n = 72; unadjusted mean difference 0.1 points, adjusted mean difference –0.5); subgroup 3: OA symptom duration >9.5 years and OA symptom duration ≤18.5 years (n = 80; unadjusted mean difference 3.5 points, adjusted mean difference 3.1); subgroup 4: OA symptom duration >9.5 years and >18.5 years (n = 67; unadjusted mean difference 5.7 points, adjusted mean difference 7.5). Dots indicate the mean difference IBET-PT; error bars show the SE of the difference.

knee OA symptoms for a longer period of time may particularly benefit from the personalized support and tailored exercise offered by a physical therapist versus a more self-directed exercise program.

We used data from all 3 study arms to explore, using GUIDE, which subgroups showed the greatest improvement in each of the active treatment arms compared to usual care (WL). For the subgroup of participants who were not employed, both IBET and PT had greater improvements than WL (Figure 4); improvement relative to WL was somewhat larger for PT than IBET. However, for those who were employed, WL was associated with larger improvements in WOMAC than either PT or IBET after adjusting for covariates, including baseline WOMAC score. Notably, the primary driver of magnitude of effects (relative to WL) was employment status, a factor different from those involved in the comparisons of 2 active treatment arms. A likely explanation for the 3-group GUIDE results is that participants who were not employed had more time to engage in the intervention, including adherence to home exercise recommendations. This pattern was observed for the IBET group (though less pronounced than for PT), given that no in-person visits were required for the intervention, and exercises could be completed at participants' convenience. Individuals with knee OA who are still employed may need additional support or strategies to fit regular activity into daily routines. Based on the GUIDE model with a minimum of n = 15 participants per arm, the PT intervention had a particularly strong impact for participants who had OA symptoms for a longer period of time (among those in the not-employed subgroup). These results align with findings of GUIDE analyses of the 2 active groups, in which patients with the longest duration experienced greater benefit from PT (Figure 3).

An important aspect of both the QUINT and GUIDE procedures is pruning of trees to avoid overfitting. For the QUINT analysis, there was little to no difference between pruned and unpruned trees for both analyses. In our GUIDE analysis, when we applied the pruning procedures, all trees were empty. Therefore, results should

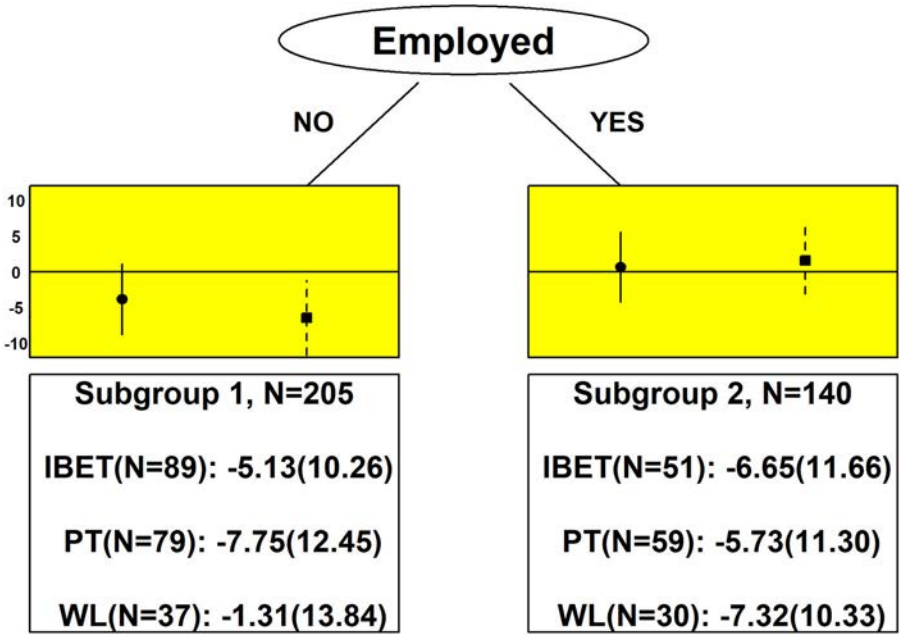


Figure 4. Generalized unbiased interaction detection and estimation subgroups with all 3 arms: internet-based exercise training (IBET), physical therapy (PT), and wait list (WL) control (n = 345; minimum sample size per subgroup is 60 total participants, i.e., at least 20 in each arm). Mean differences for subgroups (nodes) are IBET-usual care followed by PT-WL, with negative values indicating greater improvement in the treatment arm (IBET or PT) compared to WL. Subgroup 1: not employed (n = 205; unadjusted IBET-WL mean difference –3.8 points, adjusted mean difference –5.4; unadjusted PT-WL mean difference –6.4, adjusted mean difference –7.4); subgroup 2: employed (n = 140; unadjusted IBET-WL mean difference 0.7 points, adjusted mean difference 1.2; unadjusted PT-WL mean difference 1.6, adjusted mean difference 2.1). Circle dot indicates the mean difference IBET-WL, and square indicates the mean difference PT-WL; error bars show the SE of the differences. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564/abstract.

be interpreted with caution due to the potential for overfitting and instability of identified subgroups. As these are exploratory analyses in a relatively small sample size for heterogeneity of treatment-effects analyses (but typical for a clinical trial of a behavioral intervention), we presented unpruned trees for GUIDE. However, we applied prepruning procedures by specifying limits on how small subgroups could be, a step to prevent overfitting. Furthermore, we used internal validation to evaluate the bias in our effect-size differences due to overfitting when applying both QUINT and GUIDE.

Another important aspect of our methods was the process of internal validation. In heterogeneity of treatment-effects analyses, validation provides guidance for interpreting and applying results, even when analyses are exploratory and in studies with small samples where overfitting can easily occur. We applied these procedures to estimate optimism or bias in the range of mean differences in outcomes for the final tree (i.e., the apparent range). Based on internal validation results, there were large corrections to all apparent ranges, reflecting potential bias. However, some of the optimism-corrected ranges still fell in the clinically meaningful range, indicating that meaningful differences in subgroups may apply beyond this sample. Specifically, in the QUINT analysis with 2 active treatments, there were large potential biases in the apparent range for both n = 20 and n = 15 per arm for each subgroup, indicating overfitting and instability of results. In the analyses including the 2 active arms only, an optimism-corrected range greater than –8 is indicative of subgroups with clinically meaningful differences that may apply beyond this sample (28). In the analysis with n = 15 per arm, the corrected range was above the threshold of –8, but was well below this threshold in the n = 20 per arm analysis (28). Applying the n = 20 per arm per subgroup is possibly too stringent a criterion, masking smaller subgroups with larger and more stable differences. In GUIDE with all 3 arms, there were large potential biases in the apparent range for both n = 20 and n = 15 per arm per subgroup. In this case, an optimism-corrected range greater than –4 is indicative of subgroups with clinically meaningful differences of treatment versus control that may apply beyond this sample. Similar to the QUINT analysis, the n = 15 per subgroup per arm yielded optimism-corrected ranges greater than this threshold, while the n = 20 per subgroup per arm yielded optimism-corrected ranges below this threshold.

There are some limitations to this study. We have not focused the interpretation of results on uncertainty estimates of treatment differences in subgroups, as standard uncertainty estimates do not account for all the uncertainty due to the data-driven process; methods for uncertainty of estimates accounting for all the uncertainty due to the data-driven process is an area of active research in heterogeneity of treatment-effects analysis. In this pragmatic trial, we did not obtain radiographs, and information

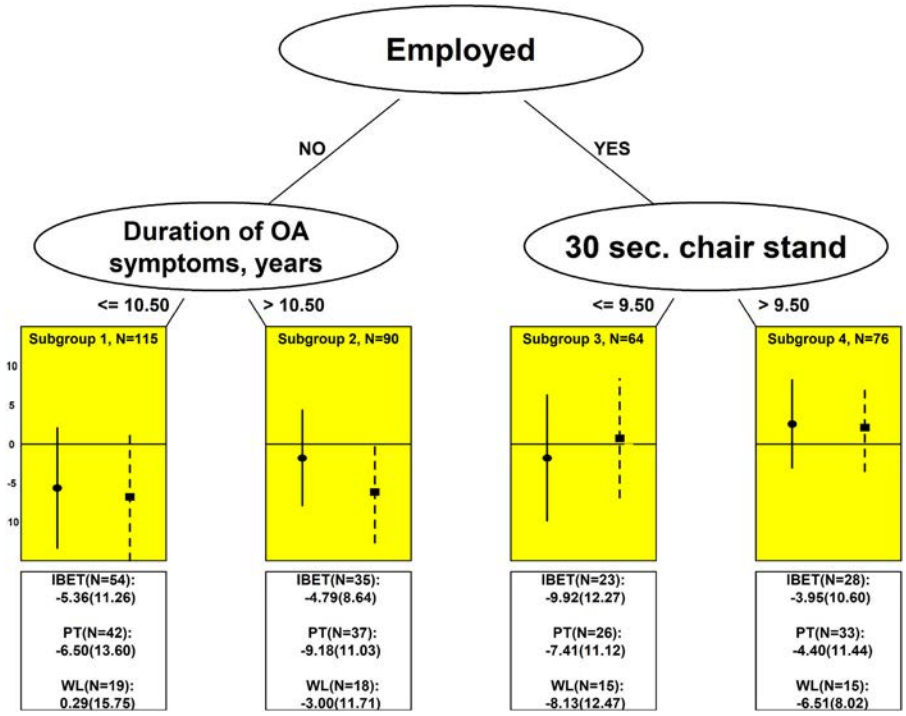


Figure 5. Generalized unbiased interaction detection and estimation subgroups with all 3 arms: internet-based exercise training (IBET), physical therapy (PT), and wait list (WL) control (n = 345; minimum sample size per subgroup is 45 total participants, i.e., at least 15 in each arm). Subgroup 1: not employed and osteoarthritis (OA) symptom duration ≤10.5 years (n = 115; unadjusted IBET-WL mean difference –5.7 points, adjusted mean difference –5.7; unadjusted PT-WL mean difference –6.8, adjusted mean difference –6.5); subgroup 2: not employed and OA symptom duration >10.5 years (n = 90; unadjusted IBET-WL mean difference –1.8 points, adjusted mean difference –5.8; unadjusted PT-WL mean difference –6.2, adjusted mean difference –11.4); subgroup 3: employed and number of chair stands ≤9.5 (n = 64; unadjusted IBET-WL mean difference –1.8 points, adjusted mean difference –1.4; unadjusted PT-WL mean difference 0.7, adjusted mean difference 0.7); subgroup 4: employed and number of chair stands >9.5 (n = 76; unadjusted IBET-WL mean difference 2.6 points, adjusted mean difference 3.7; unadjusted PT-WL mean difference 2.1, adjusted mean difference 2.3). Circle dot indicates the mean difference IBET-WL, and square indicates the mean difference PT-WL; error bars show the SE of the differences. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24564/abstract.

on radiographic severity was not systematically available within the electronic health record. We also did not collect data on joint malalignment. Both radiographic severity and malalignment have some previous evidence for moderating effects of exercise-based interventions for OA and therefore would have been valuable to include in these analyses. However, current radiographic severity and precise measurement of joint alignment are not available at all clinical encounters; therefore, these variables may not be the most practical for use in guiding recommendations for exercise-based interventions.

In summary, these analyses highlight ways to use various data-driven heterogeneity of treatment-effects methods to address different research questions within randomized clinical trials, depending on whether the trial has 2 active treatments (where the QUINT method applies) or usual care and 1 or more treatment arms (where the GUIDE method applies). Problematic areas that need to be considered or addressed when applying these methods are multiple testing implications, the potential for too much complexity, appropriate uncertainty estimation, and reproducibility of subgroups (15). While these analyses were exploratory in nature, they provide some evidence that there are subgroups for whom different exercise-based treatments (IBET versus PT) were more efficacious than for others. Even after correcting for optimism bias, some differences were in the clinically meaningful range. In particular, our results suggest that younger patients with lower BMI may be good candidates for self-guided exercise programs (e.g., our IBET intervention) and that regardless of the type of exercise-based intervention, individuals who are currently employed may need additional supportive strategies. However, additional studies are needed to further explore heterogeneity of treatment effects in the context of exercise-based therapies for OA, including different programs and cohorts.

ACKNOWLEDGMENTS

We thank all of the study participants, without whom this work would not be possible, and we thank the following team members for their contributions to the research: Caroline Nagle, Kimberlea Grimm, Ashley Gwyn, Bernadette Benas, Alex Gunn, Leah Schrubbe, and Quinn Williams. We also express gratitude to the stakeholder panel for this project: Ms. Sandy Walker, LPN (Chapel Hill Children's Clinic), Ms. Susan

Pedersen, RN, BSN, Ms. Sally Langdon Thomas, Mr. Ralph B. Brown, Ms. Frances Talton, CDA, RHS Retired, Dr. Katrina Donahue, MD, MPH (Department of Family Medicine at the University of North Carolina at Chapel Hill), Dr. Alison Brooks, MD, MPH (Department of Orthopedics and Rehabilitation at the University of Wisconsin-Madison), Dr. Anita Bemis-Dougherty, PT, DPT, MAS (American Physical Therapy Association), Dr. Teresa J. Brady, PhD (Centers for Disease Control and Prevention), Ms. Laura Marrow (Arthritis Foundation National Office), Ms. Megan Simmons Skidmore (American Institute of Healthcare and Fitness), and Dr. Maura Daly Iversen, PT, DPT, SD, MPH, FNAP, FAPTA (Department of Physical Therapy, Movement, and Rehabilitation Sciences Northeastern University). The study team thanks study physical therapists and physical therapy assistants: Jennifer Cooke, PT, DPT, Jyotsna Gupta, PT, PhD, and Carla Hill, PT, DPT, OCS, Cert MDT (Division of Physical Therapy, University of North Carolina at Chapel Hill), Bruce Buley, Andrew Genova, and Ami Pathak (Comprehensive Physical Therapy, Chapel Hill, North Carolina), Chris Gridley and Aaron Kline (Pivot Physical Therapy, Smithfield, North Carolina).

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. Coffman 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. Coffman, Arbeeva, Schwartz, Callahan, Golightly, Goode, Huffman, Allen.

Acquisition of data. Arbeeva, Schwartz, Callahan, Allen.

Analysis and interpretation of data. Coffman, Arbeeva, Schwartz, Allen.

REFERENCES

1. Kolasinski SL, Neogi T, Hochberg MC, Oatis C, Guyatt G, Block J,et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2020;72:149–62.
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. Fransen M, McConnell S, Harmer AR, van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee: a Cochrane systematic review. Br J Sports Med 2015;49:1554–7.
4. Pignato M, Arbeeva L, Schwartz TA, Callahan LF, Cooke J, Golightly YM, et al. Level of participation in physical therapy or an internet-based exercise training program: associations with outcomes for patients with knee osteoarthritis. BMC Musculoskelet Disord 2018; 19:238.
5. Holden MA, Burke DL, Runhaar J, van Der Windt D, Riley RD, Dziedzic K, et al. Subgrouping and TargetEd Exercise pRogrammes for knee and hip OsteoArthritis (STEER OA): a systematic review update and individual participant data meta-analysis protocol. BMJ Open 2017; 7:e018971.
6. Ross R, Goodpaster BH, Koch LG, Sarzynski MA, Kohrt WM, Lohannsen NM, et al. Precision exercise medicine: understanding exercise response variability. Br J Sports Med 2019;53:1141–53.
7. National Library of Medicine, MedlinePlus. What is precision medicine? URL: https://ghr.nlm.nih.gov/primer/precisionmedicine/definition.
8. Wright AA, Cook CE, Flynn TW, Baxter GD, Abbott JH. Predictors of response to physical therapy intervention in patients with primary hip osteoarthritis. Phys Ther 2011;91:510–24.
9. French HP, Galvin R, Cusack T, McCarthy GM. Predictors of short-term outcome to exercise and manual therapy for people with hip osteoarthritis. Phys Ther 2014;94:31–9.
10. Bennell KL, Dobson F, Roos EM, Skou ST, Hodges P, Wrigley TV, et al. Influence of biomechanical characteristics on pain and function outcomes from exercise in medial knee osteoarthritis and varus malalignment: exploratory analyses from a randomized controlled trial. Arthritis Care Res (Hoboken) 2015;67:1281–8.
11. Kudo M, Watanabe K, Otsubo H, Kamiya T, Kaneko F, Katayose M, et al. Analysis of effectiveness of therapeutic exercise for knee osteoarthritis and possible factors affecting outcome. J Orthop Sci 2013; 18:932–9.
12. Knoop J, Dekker J, van der Leeden M, van der Esch M, Klein JP, Hunter DJ, et al. Is the severity of knee osteoarthritis on magnetic resonance imaging associated with outcome of exercise therapy? Arthritis Care Res (Hoboken) 2014;66:63–8.
13. Rothwell PM. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. Lancet 2005;365:176–86.
14. Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. BMJ 2010;340:c117.
15. Lipkovich I, Dmitrienko A, D'Agostino RB. Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. Stat Med 2017;36:136–96.
16. Deveza LA, Nelson AE, Loeser RF. Phenotypes of osteoarthritis: current state and future implications. Clin Exp Rheumatol 2019;37 Suppl 120:64–72.
17. Seibold H, Zeileis A, Hothorn T. Model-based recursive partitioning for subgroup analyses. Int J Biostat 2016;12:45–63.
18. Dusseldorp E, van Mechelen I. Qualitative interaction trees: a tool to identify qualitative treatment-subgroup interactions. Stat Med 2014; 33:219–37.
19. Loh WY, Fu H, Man M, Champion V, Yu M. Identification of subgroups with differential treatment effects for longitudinal and multiresponse variables. Stat Med 2016;35:4837–55.
20. Loh WY, He X, Man M. A regression tree approach to identifying subgroups with differential treatment effects. Stat Med 2015;34:1818–33.
21. Williams QI, Gunn AH, Beaulieu JE, Benas BC, Buley B, Callahan LF, et al. Physical therapy vs. internet-based exercise training (PATH-IN) for patients with knee osteoarthritis: study protocol of a randomized controlled trial. BMC Musculoskelet Disord 2015;16:264.
22. Allen KD, Arbeeva L, Callahan L, Golightly YM, Goode AP, Heiderscheit BC, et al. Physical therapy vs. internet-based exercise training for patients with knee osteoarthritis: results of a randomized controlled trial. Osteoarthritis Cartilage 2018;26:383–96.
23. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods 2009;14: 323–48.
24. Doove LL, van Deun K, Dusseldorp E, van Mechelen I. QUINT: A tool to detect qualitative treatment-subgroup interactions in randomized controlled trials. Psychother Res 2016;26:612–22.
25. Littell R, Milliken G, Stroup W, Wolfinger R, Schabenberger O. SAS for mixed models. 2nd ed. Cary (NC): SAS Press; 2006.
26. Dusseldorp E, Doove L, Mechelen I. Quint: An R package for the identification of subgroups of clients who differ in which treatment alternative is best for them. Behav Res Methods 2016;48:650–63.
27. Harrell FE. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. 2nd ed. New York: Springer; 2015.
28. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum 2001;45:384–91.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1369–1373
DOI 10.1002/acr.24582

AMERICAN COLLEGE of RHEUMATOLOGY
Empowering Rheumatology Professionals

BRIEF REPORT

Identification of Radiographic Foot Osteoarthritis: Sensitivity of Views and Features Using the La Trobe Radiographic Atlas

Hylton B. Menz,[1] Shannon E. Munteanu,[2] Michelle Marshall,[3] Martin J. Thomas,[4] Trishna Rathod-Mistry,[3] George M. Peat,[3] and Edward Roddy[4]

Objective. To compare the sensitivity of alternative case finding approaches for the identification of foot osteoarthritis (OA) based on the La Trobe radiographic atlas.

Methods. This was a cross-sectional study of 533 adults age ≥50 years with foot pain in the past year. Weightbearing dorsoplantar (DP) and lateral radiographs were taken of both feet. The La Trobe radiographic atlas was used to document the presence of osteophytes (OPs) and joint space narrowing (JSN). The prevalence of OA in each joint was documented using both views and features in combination (as recommended in the original atlas), and by using a single view (DP or lateral only) and a single feature (OP or JSN only).

Results. Compared to the recommended case definition based on OPs and JSN using both views, a DP-only view identified between 15% and 77% of OA cases, while a lateral-only view identified between 28% and 97% of OA cases. Compared to the recommended case definition of using both features, using only OPs identified between 46% and 94% of OA cases, while using only JSN identified between 19% and 76% of OA cases.

Conclusion. Applying the La Trobe radiographic atlas but using only 1 radiograph view (DP or lateral) or 1 feature (OP or JSN) in isolation misses a substantial number of OA cases, and the sensitivity of these approaches varies considerably between different foot joints. These findings indicate that, where possible, the atlas should be administered according to the original description to avoid under-ascertainment of radiographic foot OA.

INTRODUCTION

Osteoarthritis (OA) is a leading cause of pain and disability and most commonly affects the knees, hips, hands, and feet (1). Although OA affecting the knees, hips, and hands has received considerable research attention, foot OA has been largely ignored until relatively recently (2), despite being highly prevalent (3), disabling (4), and accounting for a substantial number of primary care consultations (5). A key barrier to progress with foot OA research has been the absence of a standardized case definition, with previous studies assessing different combinations of foot joints and using a range of radiographic classification criteria (6). As a consequence of this inconsistency, prevalence estimates of radiographic foot OA have varied widely (6).

To address this issue, a foot-specific atlas (the La Trobe Radiographic Atlas of Foot Osteoarthritis) was developed in 2007 (7). The atlas enables the documentation of radiographic OA in 5 foot joints according to the presence of osteophytes (OPs) and joint space narrowing (JSN) from dorsoplantar (DP) and lateral views, and has since been adopted for use in several population-based studies (8–10). Due to the substantial

The views expressed in this article are those of the authors and not necessarily those of the NHS, the NIHR, Health Education England, or the Department of Health and Social Care.

Supported by the Arthritis Research UK Programme (grant 18174) and by the West Midlands North Connected Learning Research Network. Dr. Menz's work was supported by a National Health and Medical Research Council of Australia Senior Research Fellow award (1135995). Dr. Thomas' work was supported by an Integrated Clinical Academic Programme Clinical Lectureship from the NIHR and Health Education England (ICA-CL-2016-02-014) and by an NIHR Development and Skills Enhancement Award (NIHR300818).

[1]Hylton B. Menz, DSc: School of Medicine, Keele University, Keele, Staffordshire, UK, and La Trobe University, Melbourne, Victoria, Australia; [2]Shannon E. Munteanu, PhD: La Trobe University, Melbourne, Victoria, Australia; [3]Michelle Marshall, PhD, Trishna Rathod-Mistry, MSc, George M. Peat, PhD: School of Medicine, Keele University, Keele, Staffordshire, UK; [4]Martin J. Thomas, PhD, Edward Roddy, DM: School of Medicine, Keele University, Keele, and Midlands Partnership NHS Foundation Trust, Haywood Hospital, Burslem, Staffordshire, UK.

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

Address correspondence to Hylton B. Menz, DSc, La Trobe University, Melbourne, Victoria, Australia. Email: h.menz@latrobe.edu.au.

Submitted for publication October 25, 2020; accepted in revised form February 12, 2021.

SIGNIFICANCE & INNOVATIONS

- Applying the La Trobe Radiographic Atlas using only 1 radiographic view (dorsoplantar or lateral) misses a substantial number of osteoarthritis (OA) cases.
- Applying the La Trobe Radiographic Atlas using only 1 radiographic feature (osteophytes or joint space narrowing) misses a substantial number of OA cases.
- The atlas should be administered according to the original description to avoid under-ascertainment of foot OA.

variability in the bony morphology of foot joints, the authors of the original atlas recommended using both radiographic views and features in combination to identify foot OA, and in a subsequent publication demonstrated that using only 1 view or feature in isolation missed a substantial number of cases (11). However, because this analysis was undertaken by the developers of the atlas on a convenience sample of older adults with a high prevalence of foot OA, we consider replicating this finding to be important, using an independent assessor to determine whether this low sensitivity would also be reflected in a more representative, population-based sample.

If a single radiographic view and/or feature could identify a similar number of cases to a combination of both views and features, foot OA research could potentially be conducted more efficiently. Therefore, the objective of this study was to compare the sensitivity of alternative case definitions for the identification of foot OA based on the La Trobe atlas using data from the Clinical Assessment Study of the Foot, a large, population-based study conducted in the UK (8).

MATERIALS AND METHODS

Study design. Data were collected via a population-based health survey and research assessment clinic as part of the Clinical Assessment Study of the Foot (8). Adults age ≥50 years registered with 4 general practices were invited to take part in the study, irrespective of consultation for foot pain or problems. Ethical approval was obtained from the Coventry Research Ethics Committee (#10/H1210/5). All eligible participants were mailed a Health Survey questionnaire that gathered information on demographic and social characteristics and general health. Participants who reported pain in and around the foot in the past 12 months and provided written consent to further contact were invited to attend a research clinic where radiographs were obtained.

Radiographic assessment of foot OA. Bilateral weight-bearing plain film radiographs were taken according to standardized protocols (8). The participant stood in a relaxed position with their weight distributed equally across both feet. For the DP projection, the radiograph tube was angled 15° cranially with a vertical central ray centered at the base of the third metatarsal. For the lateral projection, the radiograph tube was angled at 90° with a horizontal central ray centered on the base of the first metatarsal (7). The presence of OPs and JSN was assessed in 5 joints: the first metatarsophalangeal (MTP) joint, the first cuneometatarsal (CM) joint, the second CM joint, the navicular–first cuneiform (N-1st C) joint, and the talonavicular (TN) joint. For each joint, the presence of OPs and JSN was graded from 0 to 3 on both DP and lateral views, with the exception of TN joint OPs, where only the lateral view was used, as OPs most commonly develop on the dorsal aspect of this joint, which is difficult to visualize from a DP view (11).

The presence of radiographic OA in each joint was documented using 5 different case definitions: 1) a score of ≥2 for either OPs or JSN from either the DP or lateral view (as recommended in the original atlas), 2) a score of ≥2 for either OPs or JSN from the DP view only, 3) a score of ≥2 for either OPs or JSN from the lateral view only, 4) a score of ≥2 for OPs only from either the DP or lateral view, and 5) a score of ≥2 for JSN only from either the DP or lateral view.

All radiographs were initially graded using the original case definition by a single reader (MM) with previously documented intraexaminer reliability (3). To establish the inter- and intraexaminer reliability of the different case definitions based on individual views and features, HBM and MM independently scored radiographs from 60 randomly selected participants (n = 120 feet).

Statistical analysis. Analyses were conducted using SPSS Statistics, version 25, and Stata SE, version 14.2. The number of OA cases in each joint identified according to the case definitions using individual views and features were expressed as a percentage of cases defined using the original atlas description. The inter- and intraexaminer reliability of the different case definitions were calculated using Gwet's AC1 kappa (12) and percentage agreement statistics.

RESULTS

Study population. As previously reported, a total of 5,109 completed Health Survey questionnaires were received (adjusted response 56%) (3). Of these, 1,635 individuals who reported pain in and around the foot in the past 12 months and who provided written consent were invited to the research assessment clinic and 560 attended. Individuals with inflammatory arthritis (n = 24) were excluded from this analysis, and foot radiographs were unavailable for 3 participants, leaving a total of 533 eligible participants (1,066 feet) (235 men and 298 women with a mean ± SD age of 65 ± 8 years).

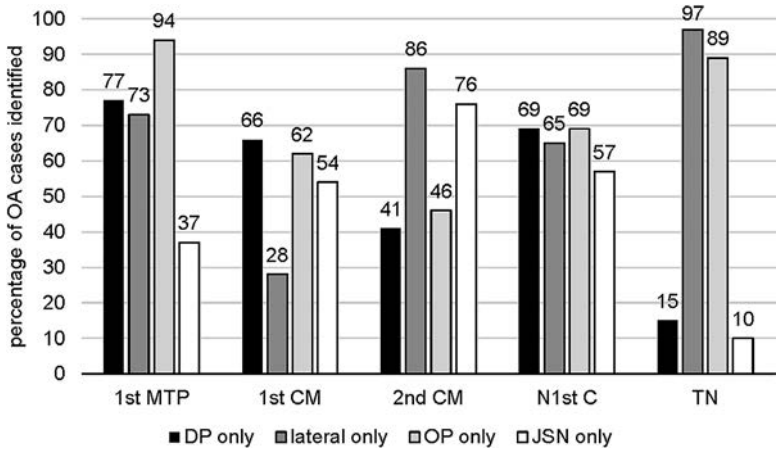


Figure 1. Foot OA cases identified according to individual radiographic views and features as a percentage of cases defined using the original atlas description (n = 1,066 feet). CM = cuneometatarsal; DP = dorsoplantar; JSN = joint space narrowing; MTP = metatarsophalangeal; N1st C = navicular–1st cuneiform; OA = osteoarthritis; OP = osteophytes; TN = talonavicular.

Foot OA cases identified with different case definitions. The prevalence of radiographic OA in each joint according to the original atlas case definition was as follows: 1st MTP joint (n = 294, 27.6%), 1st CM joint (n = 50, 4.7%), 2nd CM joint (n = 50, 4.7%), N-1st C joint (n = 86, 8.1%), and TN joint (n = 158, 14.8%). Figure 1 shows OA cases for each joint identified according to individual views and features as a percentage of cases defined using the original atlas description. Compared to the recommended case definition based on OPs and JSN using both views, a DP-only view identified between 14.5% and 77.2% of OA cases. The highest sensitivity was for the 1st MTP joint and the lowest was for the TN joint. Using a lateral-only view identified between 28% and 96.8% of OA cases. The highest sensitivity was for the TN joint and the lowest was for the 1st CM joint. Compared to the recommended case definition of using both features, using only OPs identified between 45.7% and 94.2% of OA cases. The highest sensitivity was for the 1st MTP joint and the lowest was for the 2nd CM joint. Using only JSN identified between 19.0% and 76.1% of OA cases. The highest sensitivity was for the 2nd CM joint and the lowest was for the TN joint. Figure 2 shows the relative frequency of radiographic features classifying joints as having OA using the original atlas description.

Reliability of different case definitions. Tables 1 and 2 show the intra- and interexaminer reliability of foot OA assessment using the different case definitions. Reliability was similarly high across different combinations of views and features (κ ranging from 0.923 to 1.000 for intraexaminer reliability and 0.705 to 1.000 for interexaminer reliability).

DISCUSSION

The objective of this study was to compare the sensitivity of alternative case-finding approaches to the identification of foot

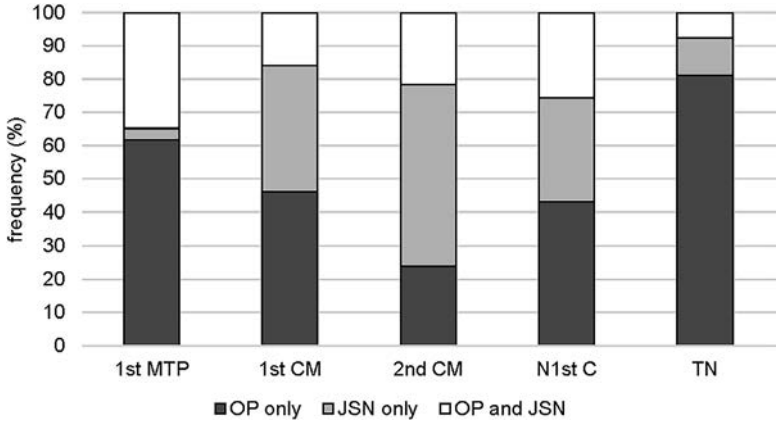


Figure 2. Relative frequency of radiographic features classifying joints as having osteoarthritis using the original atlas description (both views, n = 1,066 feet). CM = cuneometatarsal; JSN = joint space narrowing; MTP = metatarsophalangeal; N1st C = navicular–1st cuneiform; OP = osteophytes; TN = talonavicular.

Table 1. Intra- and interexaminer reliability of foot osteoarthritis assessment using different case definitions according to radiographic view (n = 120 feet)*

	Intraexaminer reliability			Interexaminer reliability		
Joint	Both DP and lateral	DP only	Lateral only	Both DP and lateral	DP only	Lateral only
1st MTP	0.923 (96)	0.860 (91)	0.911 (94)	0.705 (81)	0.868 (90)	0.915 (93)
1st CM	0.960 (97)	0.971 (98)	0.991 (99)	1.000 (100)	0.992 (99)	1.000 (100)
2nd CM	0.942 (96)	1.000 (100)	0.935 (95)	0.802 (84)	0.974 (98)	0.843 (87)
N-1st C	0.979 (98)	0.991 (99)	0.991 (99)	0.916 (93)	0.956 (96)	0.992 (99)
TN	0.950 (97)	0.982 (98)	0.950 (97)	0.923 (93)	1.000 (100)	0.964 (97)
Mean κ	0.951	0.961	0.956	0.869	0.958	0.942

* Values are the kappa (% agreement), unless indicated otherwise. Kappa value is Gwet's AC1 kappa. CM = cuneometatarsal; DP = dorsoplantar; MTP = metatarsophalangeal joint; N-1st C = navicular–first cuneiform; TN = talonavicular.

OA based on the La Trobe atlas (7). We found that compared to the recommended case definition based on identifying OPs and JSN from DP and lateral views, using only 1 feature or view in isolation missed a substantial number of OA cases, and the impact of this varied considerably between joints. These findings suggest that where possible, the atlas should be administered according to the original description to avoid under-ascertainment of radiographic foot OA.

During the development of the original atlas, the inclusion of 2 radiographic projections was justified on the basis that due to differences in bony morphology, the DP view would provide the greatest clarity for some joints, whereas the lateral view would be more suitable for others (7). This is clearly demonstrated in our findings. For example, using the DP view in isolation demonstrated moderate sensitivity for the 1st MTP joint (77%) but very low sensitivity for the TN joint (15%). In contrast, using the lateral view in isolation demonstrated high sensitivity for the TN joint (97%) but low sensitivity for the 1st CM joint (28%). Of the 5 joints evaluated, only the 1st MTP joint demonstrated similar sensitivity when either view was used, as OPs, the most dominant feature of 1st MTP joint OA, are often visible on both the dorsal and mediolateral aspects of the joint.

The inclusion of 2 features, OPs and JSN, appears to be necessary for assessing foot OA due to variation in how OA manifests in individual joints. For example, OA in the 1st MTP joint is characterized by the formation of large OPs, whereas the 2nd CM joint, possibly due to its more proximal location in the foot and limited range of motion, is more likely to develop JSN. If the atlas was applied using OPs in isolation, most cases of 1st MTP joint (96%) and TN joint (89%) OA would be identified, but a substantial number of cases in the remaining joints would be missed. Similarly, using JSN in isolation would provide moderate sensitivity for the 2nd CM joint (76%), but unacceptably low sensitivity (10–57%) for the remaining joints.

Despite substantial differences in sample characteristics and the prevalence of radiographic OA in each foot joint, our findings in relation to the sensitivity of views and features are consistent with those reported in the original atlas (11). The atlas was developed using a convenience sample of people ages 62–94 years (mean age 76 years) and reported a higher prevalence of radiographic OA in individual joints (ranging from 22% for the 1st CM joint to 60% for the 2nd CM joint) than our population-based sample of people age ≥50 years. However, the relative proportion of OA cases identified using limited views or features was similar, as was the overall representation of OPs and JSN across the different joints. A notable difference was the sensitivity of identifying 1st MTP joint OA using the DP view only, which was higher in the original atlas study than in the current study (95% compared to 77%). This finding suggests that the using the DP view alone may be less sensitive in identifying 1st MTP joint OA in a younger population.

Table 2. Intra- and interexaminer reliability of foot osteoarthritis assessment using different case definitions according to radiographic feature (n = 120 feet)*

	Intraexaminer reliability			Interexaminer reliability		
Joint	Both OP and JSN	OP only	JSN only	Both OP and JSN	OP only	JSN only
1st MTP	0.923 (96)	0.923 (96)	0.981 (98)	0.705 (81)	0.772 (84)	0.959 (97)
1st CM	0.960 (97)	0.991 (99)	0.972 (98)	1.000 (100)	1.000 (100)	0.992 (99)
2nd CM	0.942 (96)	0.981 (98)	0.957 (97)	0.802 (84)	0.966 (97)	0.964 (97)
N-1st C	0.979 (98)	0.991 (99)	0.981 (98)	0.916 (93)	0.966 (97)	0.992 (99)
TN	0.950 (97)	0.952 (97)	1.000 (100)	0.923 (93)	1.000 (100)	0.982 (98)
Mean κ	0.951	0.968	0.978	0.869	0.941	0.978

* Values are the kappa (% agreement), unless indicated otherwise. Kappa value is Gwet's AC1 kappa. CM = cuneometatarsal; JSN = joint space narrowing; MTP = metatarsophalangeal joint; N-1st C = navicular–first cuneiform; OP = osteophyte; TN = talonavicular.

Our findings provide further evidence to support the application of the La Trobe atlas as originally described. However, there are several inherent limitations of the atlas that warrant consideration. First, the atlas is limited to 5 foot joints. These joints were selected based on their suspected susceptibility to the development of OA, but also due to their ease of visualization using DP and lateral radiographs (7). Joints not represented in the atlas (including the subtalar joint, lateral tarsal joints, and interphalangeal joints) are also known to develop OA (6), but additional radiographic views would be required to adequately identify changes in these joints. Second, as with all radiographic atlases, there is some degree of subjectivity involved (13), although reliability has repeatedly been demonstrated to be acceptable both within and between examiners (3,7). Third, the atlas is limited to observations of OPs and JSN, and does not include other frequently observed features of OA such as subchondral sclerosis and cysts (14). Finally, all study participants had current/recent foot pain.

In summary, this study has shown that when applying the La Trobe atlas to identify foot OA, using only 1 radiographic view or 1 feature in isolation misses a substantial number of OA cases, and the sensitivity of these approaches varies considerably between different foot joints. These findings indicate that, where possible, the atlas should be administered according to the original description to avoid under-ascertainment of radiographic foot OA.

ACKNOWLEDGMENTS

We would like to thank the administrative, health informatics, and research nurse teams of Keele University's Arthritis Research UK Primary Care Centre, the staff of the participating general practices, and the Haywood Hospital, particularly Dr. Jackie Saklatvala, Carole Jackson, and the radiographers at the Department of Radiology. We would like to acknowledge the contributions of Linda Hargreaves, Gillian Levey, Liz Mason, Dr. Jennifer Pearson, Julie Taylor, and Dr. Laurence Wood to data collection.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Menz, Munteanu, Marshall, Thomas, Peat, Roddy.

Acquisition of data. Menz, Marshall, Thomas, Roddy.

Analysis and interpretation of data. Menz, Munteanu, Marshall, Thomas, Rathod-Mistry, Peat, Roddy.

REFERENCES

1. Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. Br Med Bull 2013;105:185–99.
2. Roddy E, Menz HB. Foot osteoarthritis: latest evidence and developments. Ther Adv Musculoskelet Dis 2018;10:91–103.
3. Roddy E, Thomas MJ, Marshall M, Rathod T, Myers H, Menz HB, et al. The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: the Clinical Assessment Study of the Foot. Ann Rheum Dis 2015;74:156–63.
4. Thomas MJ, Moore A, Roddy E, Peat G. "Somebody to say 'come on we can sort this'": a qualitative study of primary care consultation among older adults with symptomatic foot osteoarthritis. Arthritis Care Res (Hoboken) 2013;65:2051–5.
5. Paterson KL, Harrison C, Britt H, Hinman RS, Bennell KL. Management of foot/ankle osteoarthritis by Australian general practitioners: an analysis of national patient-encounter records. Osteoarthritis Cartilage 2018;26:888–94.
6. Trivedi B, Marshall M, Belcher J, Roddy E. A systematic review of radiographic definitions of foot osteoarthritis in population-based studies. Osteoarthritis Cartilage 2010;18:1027–35.
7. Menz HB, Munteanu SE, Landorf KB, Zammit GV, Cicuttini FM. Radiographic classification of osteoarthritis in commonly affected joints of the foot. Osteoarthritis Cartilage 2007;15:1333–8.
8. Roddy E, Myers H, Thomas MJ, Marshall M, D'Cruz D, Menz HB, et al. The clinical assessment study of the foot (CASF): study protocol for a prospective observational study of foot pain and foot osteoarthritis in the general population. J Foot Ankle Res 2011;4:22.
9. Bowen C, Gates L, McQueen P, Daniels M, Delmestri A, Drechsler W, et al. Natural history of radiographic first metatarsophalangeal joint osteoarthritis: a nineteen-year population-based cohort study. Arthritis Care Res (Hoboken) 2020;72:1224–30.
10. Flowers P, Nelson AE, Hannan MT, Hillstrom HJ, Renner JB, Jordan JM, et al. Foot osteoarthritis frequency and associated factors in a community-based cross-sectional study of White and African American adults. Arthritis Care Res (Hoboken) 2021;73:1784–8.
11. Menz HB, Munteanu SE, Landorf KB, Zammit GV, Cicuttini FM. Radiographic evaluation of foot osteoarthritis: sensitivity of radiographic variables and relationship to symptoms. Osteoarthritis Cartilage 2009;17:298–303.
12. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 2013;13:61.
13. McQueen P, Gates L, Marshall M, Doherty M, Arden N, Bowen C. The effect of variation in interpretation of the La Trobe radiographic foot atlas on the prevalence of foot osteoarthritis in older women: the Chingford general population cohort. J Foot Ankle Res 2017;10:54.
14. Gupta KB, Duryea J, Weissman BN. Radiographic evaluation of osteoarthritis. Radiol Clin North Am 2004;42:11–41.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1374–1383
DOI 10.1002/acr.24562

Relationship Between Patient-Reported Readiness for Total Knee Arthroplasty and Likelihood of a Good Outcome at One-Year Follow-Up

Gillian A. Hawker,[1] Barbara L. Conner-Spady,[2] Eric Bohm,[3] Michael J. Dunbar,[4] C. Allyson Jones,[5] Bheeshma Ravi,[1] Tom Noseworthy,[2] Linda J. Woodhouse,[6] Peter Faris,[2] Donald Dick,[5] James Powell,[2] Paulose Paul,[5] and Deborah A. Marshall,[2] on behalf of the BEST-Knee Study Team

Objective. To determine the relationship between patients' preoperative readiness for total knee arthroplasty (TKA) and surgical outcome at 1 year post-TKA.

Methods. This prospective cohort study recruited patients with knee osteoarthritis (OA) who were ≥30 years and were referred for TKA at 2 hip/knee surgery centers in Alberta, Canada. Those who underwent primary unilateral TKA completed questionnaires prior to TKA to assess pain using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), physical disability using the Knee Injury and Osteoarthritis Outcome Score physical function short form, perceived arthritis coping efficacy, general self-efficacy, depressed mood using the Patient Health Questionnaire 8, body mass index, comorbidities, and TKA readiness (patient acceptable symptom state; willingness to undergo TKA); these same individuals also completed the above questionnaires 1 year post-TKA to assess surgical outcomes. A good TKA outcome was defined as an individual having improved knee symptoms, measured using the Osteoarthritis Research Society International–Outcome Measures in Rheumatology responder criteria, and overall satisfaction with results of the TKA. Poisson regression with robust error estimation was used to estimate the relative risk (RR) of a good outcome for exposures, before and after controlling for covariates.

Results. Of 1,272 TKA recipients assessed at 1 year post-TKA, 1,053 with data for the outcome assessed in the study were included (mean ± SD age 66.9 ± 8.8 years; 58.6% female). Most patients (87.8%) were definitely willing to undergo TKA and had "unacceptable" knee symptoms (79.7%). Among patients who underwent TKA, 78.1% achieved a good outcome. Controlling for pre-TKA OA-related disability, arthritis coping efficacy, comorbid hip symptoms, and depressed mood, definite willingness to undergo TKA and unacceptable knee symptoms were associated with a greater likelihood of a good TKA outcome, with adjusted RRs of 1.18 (95% confidence interval [95% CI] 1.04–1.35) and 1.14 (95% CI 1.02–1.27), respectively.

Conclusion. Among patients who underwent TKA for knee OA, patients' psychological readiness for TKA and willingness to undergo TKA were associated with a greater likelihood of a good outcome. Incorporation of these factors in TKA decision-making may enhance patient outcomes and appropriate the use of TKA.

INTRODUCTION

Total knee arthroplasty (TKA) is one of the most cost-effective surgical interventions (1,2). Over 1,200,000 TKAs are performed annually worldwide (3), with ~95% performed to treat knee osteoarthritis (4), OA. While TKA is highly effective on average, studies consistently show that 15–30% of TKA recipients report little or no improvement in symptoms and/or dissatisfaction

Supported by Canadian Institutes of Health Research operating grant MOP-312807. Dr. Hawker has received research support as the Sir John and Lady Eaton Professor and Chair of Medicine at the University of Toronto. Dr. Marshall's work has received support through a Canada Research Chair in Health Systems and Services Research (2008–2018) and is currently supported by the Arthur J. E. Child Chair in Rheumatology.

[1]Gillian A. Hawker, MD, MSc, Bheeshma Ravi, MD, PhD: University of Toronto, Toronto, Ontario, Canada; [2]Barbara L. Conner-Spady, PhD, Tom Noseworthy, MD, MSc, MPH, Peter Faris, PhD, James Powell, MD, Deborah A. Marshall, PhD: University of Calgary, Calgary, Alberta, Canada; [3]Eric Bohm, MD, MSc: University of Manitoba, Winnipeg, Manitoba, Canada; [4]Michael J. Dunbar, MD, MPH: Dalhousie University and Nova Scotia Health Authority, Halifax, Nova Scotia, Canada; [5]C. Allyson Jones, PhD, Donald Dick, MD, Paulose Paul, MD: University of Alberta, Edmonton, Alberta, Canada; [6]Linda J. Woodhouse, PT, PhD: University of Alberta, Edmonton, Alberta, Canada, and Tufts University, Phoenix, Arizona.

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

Address correspondence to Gillian A. Hawker, MD, MSc, Sir John and Lady Eaton Professor and Chair, Department of Medicine, University of Toronto, 6 Queen's Park Crescent W, Third Floor, Toronto, Ontario M5S 3H2, Canada. Email: g.hawker@utoronto.ca.

Submitted for publication August 18, 2020; accepted in revised form January 12, 2021.

SIGNIFICANCE & INNOVATIONS

- This prospective cohort study assessed measures of total knee arthroplasty (TKA) readiness among patients with knee osteoarthritis (OA) pre-TKA and surgical outcomes at 1 year post-TKA. Of 1,053 study participants, 78% achieved a good TKA outcome (improved symptoms and satisfaction with postoperative results).
- Controlling for other factors, the presence of unacceptable knee symptoms and a preoperative definite willingness to undergo TKA were associated with a good TKA outcome at 1 year post-TKA.
- Consideration of patients' psychological readiness for TKA in surgical decision-making may enhance outcomes of surgery and appropriate TKA use.

with postoperative results (5–9). Appropriate care is broadly defined as that which provides a net "benefit" to the patient (10). Greater understanding of the preoperative factors that increase the likelihood of a net benefit to the patient from TKA will improve outcomes, and thus, appropriate use of this procedure.

In prior qualitative research, we found that patients with knee OA equated "appropriateness" for TKA with candidacy for the procedure (11–14). Pain intensity, the ability to cope with the pain, and how the pain affected their quality of life were seen by patients as the most important factors determining surgical candidacy (12). Patients also stressed the importance of psychological readiness and a positive attitude, which they perceived as critical to the achievement of a good TKA outcome (12,14).

Psychological factors are well known to influence treatment outcomes. In many clinical contexts (e.g., cancer care and cardiac surgery), an optimistic attitude, absence of depressed mood, and greater self-efficacy have been shown to predict better health outcomes (15–18). In the setting of joint replacement, depressed mood has been consistently associated with less improvement in OA symptoms following surgery (19), whereas greater self-efficacy has been variably linked to better adherence to postoperative rehabilitation and greater symptom improvement postoperatively (15,17,20). The willingness of a patient to consider total joint replacement has been shown to reflect their perceptions regarding candidacy for the procedure, the relative benefits versus potential risks of surgery, and the acceptability of these risks (21). Greater willingness to undergo TKA has been linked to higher rates of referral for, and performance of, TKA (21,22). However, to date, the influence of measures of psychological readiness for TKA have not been examined as potentially modifiable determinants of a good outcome from surgery. To address this knowledge gap, the present study examined the relationship between preoperative measures of psychological readiness to pursue TKA (assessed as the willingness to consider TKA and acceptability of knee OA symptoms) to a composite measure of a "good TKA outcome" at 1 year post-TKA, controlling for depressed mood, measures of self-efficacy, and other potential confounders.

PATIENTS AND METHODS

Study setting and design. This prospective cohort study recruited patient participants from a group of individuals with knee OA who were referred to a surgeon for elective TKA at 2 orthopedic hip and knee centralized intake clinics in Alberta, Canada, between October 2014 and September 2016. Individuals who were ages ≥30 years, able to read and comprehend English, and confirmed to have knee OA on physical examination and imaging were eligible for study participation. Individuals with inflammatory arthritis were excluded. To ensure representation of the most rapidly increasing group of TKA recipients (individuals <60 years [23,24]), we continued recruitment for the current study until there were at least 200 men and women in each of 3 age groups (30–59 years, 60–69 years, and ≥70 years). All 45 surgeons at the clinics that patient participants were recruited from agreed to participate in the study as well and provided written consent. See Appendix A for a list of the BEST-Knee Study Team members.

The present study was approved by the Research Ethics Boards of the University of Alberta (PRO-00051108), University of Calgary (REB 14-1294), and the Women's College Hospital (REB 2014-0092) at the University of Toronto. Both patient and surgeon participants provided written informed consent.

Assessments. After providing written consent and prior to surgeon consultation, patients completed a standardized questionnaire assessing sociodemographic characteristics (education, income, living situation, employment status), history of prior joint replacement of a hip or the contralateral knee, and preferences for TKA based on their current understanding of the risks and benefits of knee replacement and the severity of their arthritis (5-point scale, measured from "definitely willing to consider surgery now" to "definitely not willing to consider surgery now").

Following a mandatory education session prior to TKA, a second questionnaire was administered. Knee OA symptom severity was assessed with the Western Ontario and McMaster Universities Osteoarthritis (WOMAC) pain subscale (25,26) and Knee Injury and Osteoarthritis Outcome Score physical function (KOOS-PS) short form scale (27). Patients' perceived ability to cope with their knee OA pain was assessed with the 4-item Arthritis Coping Efficacy scale (28,29). Items on this scale included the following: "I am successfully coping with the pain of my arthritis" and "I am successfully coping with the emotional aspects of my arthritis," with responses measured on a 5-point scale as "strongly disagree" to "strongly agree." Item scores are summed to produce a score of 4–20, with higher scores indicating greater perceived arthritis coping efficacy. To assess overall self-efficacy,

participants also completed the General Self-Efficacy Scale, which asks respondents to indicate their level of agreement with responses of 1, 2, 3, and 4 indicating "Not at all true," "Hardly true," "Moderately true," and "Exactly true," respectively, for each of 10 statements, such as, "I can always manage to solve difficult problems if I try hard enough," and "If someone opposes me, I can find the means and ways to get what I want." Item scores are summed to create a total score of 10–40, with higher scores indicating greater general self-efficacy (30).

To assess health status, we evaluated the following in study participants: height and weight of participants to calculate body mass index (BMI), incidence of depressed mood using the Patient Health Questionnaire 8-item depression scale (PHQ-8) (31), physician-diagnosed conditions for which they were receiving treatment, and musculoskeletal comorbidity (pain or stiffness of the hips, contralateral knee, or low back). We also assessed the acceptability of their knee symptoms, which was evaluated with responses of "acceptable" or "unacceptable" on the Patient Acceptable Symptom State (PASS) questionnaire as an additional measure of psychological readiness for TKA (32). On the PASS questionnaire, the following prompt is given, "Think about all the ways your knee OA has affected you during the last 48 hours. If you were to remain in the next few months as you were the last 48 hours, would this be acceptable or unacceptable to you?" This prompt thus determines the subject's overall state of well-being (feeling good) in the context of their knee symptoms. The age and sex of participants were obtained from clinic records.

One year postsurgery, participants completed a questionnaire (WOMAC pain and KOOS-PS) to reassess knee symptoms and assess for surgical complications (open-ended text), patient global assessment of disease activity (PtGA) of change in knee pain and function prior to TKA compared to knee pain and function following TKA (possible responses on the PtGA were as follows: "much better," "somewhat better," "about the same," "somewhat worse," and "much worse"), and overall satisfaction with TKA results (4-point Likert scale, with responses including "very dissatisfied," "somewhat dissatisfied," "somewhat satisfied, and "very satisfied") (33). Participant-reported complications were verified against the participants' electronic health record. To optimize response rates, questionnaires were completed online, by interview, or on paper as desired by the study participant.

Exposures of interest and covariates. Exposures of interest were measures of psychological readiness for TKA, which included willingness of the patient to undergo TKA at surgeon consultation and responses on the patient-reported PASS questionnaire, which assessed if knee symptoms were "acceptable" or "unacceptable." We assumed that the majority of individuals scheduled to undergo TKA for knee OA would indicate they were definitely willing to have the surgery and hypothesized that those who were "probably willing," "unsure," or "unwilling" might be "less ready" and thus potentially less likely to experience a good outcome. Thus, willingness was dichotomized as "definitely willing" at surgeon consultation (yes/no). Covariates were depressed mood (measured on the PHQ-8) and additional preoperative factors that have been associated prospectively with TKA outcome (age, sex, WOMAC pain, KOOS-PS, Perceived Arthritis Coping Efficacy, General Self-Efficacy scale, BMI, number of nonmusculoskeletal comorbid conditions, musculoskeletal comorbidity [pain/stiffness in hips, contralateral knee, or low back], and occurrence of a TKA-related complication, which was defined as any revision, manipulation, infection, or patellar resurfacing [yes/no]).

Primary outcome. In prior qualitative work, surgeons and patients defined a TKA as "worthwhile" and thus appropriate if it resulted in symptom improvement and patient satisfaction with surgical results (11–14). Thus, in the present study, we used a composite dichotomous outcome to assess the net benefit from surgery, which was characterized by a patient meeting the Outcome Measures in Rheumatology (OMERACT)–Osteoarthritis Research Society International (OARSI) responder criteria and a patient's reported overall satisfaction (somewhat or very) with TKA results (yes/no). The OARSI-OMERACT criteria require the following: 1) knee pain described as "much improved" or "somewhat improved" on the PtGA, 2) an absolute improvement in pain on the WOMAC scale and function on the KOOS-PS of >20%, 3) and improvements in pain on the WOMAC scale and function on the KOOS-PS of >1/20 and 10/100, respectively (34).

Statistical analysis. After assessing distributions for normality, all variables were calculated using proportions, means, and medians, as appropriate. The number of comorbid conditions was summed and categorized as 0–1, 2, and ≥3 conditions. KOOS scores were reverse-coded so that higher scores indicated worse status. Preoperative characteristics and occurrence of a TKA complication were compared by achievement of a good outcome using chi-square test, Student's *t*-test, and Wilcoxon's rank sum test as appropriate. Collinearity of independent variables was assessed using a variance inflation factor of >4 and tolerance value of <0.25 (35). The relative risks (RRs) of a good TKA outcome associated with each of our exposures of interest were estimated using Poisson regression with robust error estimation (36), before and after controlling for covariates. Variable selection for multivariable modeling was based on a univariate screen; those variables associated with a good TKA outcome at a *P* value of ≤0.25 were included in the multivariable model. Occurrence of a TKA complication (any/none) was then added to the model to assess for any effects on our exposures of interest. A sensitivity analysis was performed to examine the relationship between "probable willingness" (yes/no) to undergo TKA and achievement of a good outcome, with and without adjustment for the variables in the original multivariable model. Modeling was then repeated using multivariable logistic regression to assess goodness of model fit and to estimate the predicted probabilities of a good TKA outcome for combinations

of patient willingness (yes/no response for "definitely willing") and responses on the PASS questionnaire (acceptable/unacceptable), while holding all other variables constant. *P* values of <0.05 were considered statistically significant. All statistical analyses were performed in SAS, version 9.4.

RESULTS

Description of the cohort. Of 1,374 consenting and eligible patients who completed preoperative assessments and underwent TKA, 1,272 (92.6%) completed the one-year follow-up assessment. Of these participants, 1,053 had complete data for the outcome measured in the present study and were included in analyses. There were no significant differences ($P < 0.05$) in sociodemographic characteristics, pre-TKA OA severity, health status, or measures of TKA readiness between TKA recipients who had data for the primary outcome and those without (data not shown).

Cohort characteristics overall and by TKA outcome. The mean ± SD age of study cohort was 66.9 ± 8.8 years, and 58.6% of participants were female. More than half (56.4%) of study participants had received postsecondary education, 46.2% had an annual income of more than $60,000, and 33.1% were employed (Table 1). The median score for general self-efficacy was 32 (interquartile range [IQR] 29–37). Mean ± SD scores on the WOMAC pain subscale and KOOS-PS questionnaire were 11.4 ± 3.5 (on a 1–20 scale) and 52.8 ± SD 17.1 (on a 0–100-point scale), respectively, which indicated moderate-to-severe knee OA symptoms. The mean ± SD score for perceived arthritis coping efficacy was 13.4 ± 3.8. Among the study participants, the mean ± SD BMI was 32.5 ± 6.3 kg/m^2, with 13.0% having a BMI of ≥40 kg/m^2. Most participants (73.7%) had ≥1 nonmusculoskeletal comorbid conditions. Median score on the PHQ-8 was 5.0 (IQR 2–10). One-half of the study participants (50.6%) reported pain or stiffness in the contralateral knee, 23.1% in one or both hips, and 25.4% in the low back; 15.9% had

Table 1. Characteristics of TKA recipients overall and by achievement of a good TKA outcome*

Characteristic	Total sample (n = 1,053)	Primary outcome not achieved (n = 231)	Primary outcome achieved (n = 822)
Sociodemographic characteristics			
Age, mean ± SD years	66.9 ± 8.8	66.8 ± 8.4	66.9 ± 8.9
Female sex	617 (58.6)	124 (53.7)	493 (59.9)
Postsecondary education	584/1,036 (56.4)	138/228 (60.5	446/808 (55.2)
White	947/1,036 (91.4)	206/228 (90.4)	741/808 (91.7)
Annual household income of >$60,000	420/910 (46.2)	88/196 (44.9)	332/714 (46.5)
Employed for pay	344/1,040 (33.1)	72/230 (31.3)	272/810 (33.6)
Living with others	835/1,040 (80.3)	181/229 (79.0)	654/811 (80.6)
Prior joint replacement hip or knee	164/1,030 (15.9)	40/224 (17.9)	124/806 (15.4)
General Self-Efficacy scale, median (IQR) score	32.0 (29.0–30.7)	32.0 (30.0–36.0)	32.0 (29.0–37.0)
Joint symptoms†			
WOMAC pain, mean ± SD score (0–20 scale)	11.4 ± 3.5	10.2 ± 3.5	11.8 ± 3.4‡
KOOS-PS short form, mean ± SD score (0–100 scale)	52.8 ± 17.1	45.4 ± 17.0	54.9 ± 16.5‡
Perceived Arthritis Coping Efficacy, mean ± SD score (4–40 scale)	13.4 ± 3.8	13.9 ± 3.6	13.2 ± 3.8‡
General health status			
BMI, mean ± SD kg/m^2	32.5 ± 6.3	32.3 ± 5.7	32.5 ± 6.4
Comorbid conditions			
0	272/1,034 (26.3)	72/229 (31.4)	200/805 (24.8)
1	363 (35.1)	74 (32.3)	289 (35.9)
2	244 (23.6)	49 (21.4)	195 (24.2)
+3	155 (15.0)	34 (14.8)	121 (15.0)
Psychological readiness and willingness			
PASS questionnaire response ("Knee symptoms Unacceptable")	836/1,049 (79.7)	156/231 (67.5)	680/818 (83.1)‡
Willingness to undergo TKA			
Definitely willing	905/1,031 (87.8)	180/225 (80.0)	725/806 (90.0)‡
Probably willing	96 (9.3)	38 (16.9)	58 (7.2)
Unsure	24 (2.3)	6 (2.7)	18 (2.2)
Definitely/probably unwilling	6 (0.6)	1 (0.4)	5 (0.6)
Occurrence of TKA complication at 1-year follow-up§	23 (2.2)	18 (7.8)	5 (0.6)‡

* Except where indicated otherwise, values are the number (%) of recipients. Data are reported as the mean ± SD for normally distributed continuous variables, the median (interquartile range [IQR]) for non-normally distributed continuous variables, and frequency (%) for categorical variables. Denominator is shown when response is <100%. *P* values are shown for univariate testing (Student's *t*-test, the Wilcoxon rank sum test, or chi-square test). BMI = body mass index; KOOS-PS = Knee Injury and Osteoarthritis Outcome Score physical function; PASS = Patient Acceptable Symptom State; TKA = total knee arthroplasty; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† Higher score indicates worse symptoms.
‡ $P < 0.05$.
§ Includes 11 revisions, 7 manipulations under anesthesia, 3 patellar resurfacing, and 2 infections.

undergone a prior joint replacement of the other knee or hip. Of the 1,053 study participants, 1,031 had reported their willingness to consider TKA at surgeon consultation, 905 (87.8%) reported they were definitely willing to undergo TKA, 96 (9.3%) indicated they were probably willing, 24 (2.3%) were unsure, and 6 (0.6%) were probably or definitely unwilling to undergo TKA. Most participants (79.7%) reported their knee OA symptoms as being "unacceptable."

Patient-reported outcomes at 1 year post-TKA are shown in Table 2. A total of 23 (2.2%) had experienced a TKA complication (11 revisions, 7 manipulations under anesthesia, 3 patellar resurfacings, and 2 infections). Including these individuals who experienced TKA complications, 78.1% met our composite criteria for a good TKA outcome (79.3% without a complication versus 21.7% with a complication; $P < 0.0001$). Compared to TKA recipients who did not experience a good TKA outcome, those who did have a good outcome had greater knee pain (measured on the WOMAC pain subscale) and disability (measured on the KOOS-PS questionnaire), lower scores for perceived arthritis coping efficacy, no hip complaints, but contralateral knee symptoms were more likely to have unacceptable knee symptoms and to be definitely willing to undergo TKA.

Relationship between measures of pre-TKA readiness and achievement of a good TKA outcome at 1 year. RRs and 95% confidence intervals (95% CIs) for exposures of interest and covariates are shown in Table 3. In univariate analyses, the following factors were eligible for inclusion in the multivariable model: sex, WOMAC pain, KOOS-PS, comorbid hip and knee complaints, and depressed mood (PHQ-8). In the multivariable model, individuals with greater pre-TKA physical function (adjusted RR of 1.05 [95% CI 1.03–1.08] per 10-unit increase on the KOOS-PS questionnaire), individuals with unacceptable knee symptoms (adjusted RR 1.14 [95% CI 1.02–1.27]) compared to those with acceptable knee symptoms, and individuals who indicated definite willingness to undergo TKA (adjusted RR 1.18 [95% CI 1.04–1.35]) compared to those who were probably willing or unsure/unwilling to undergo TKA) were at higher likelihood of having a good TKA outcome whereas those with symptoms of depression (adjusted RR of 0.93 [95% CI 0.875–0.985] per 10-unit increase on the PHQ-8) or concomitant hip complaints (adjusted RR 0.91 [95% CI 0.84–0.99]) were at a lower likelihood of having a positive surgical outcome. No independent effects were found for sex, WOMAC pain, arthritis coping efficacy, or general self-efficacy.

In secondary analysis, adding occurrence of a complication to the multivariable model, we found that those with complications were at a lower likelihood of experiencing a good TKA outcome (adjusted RR 0.32 [95% CI 0.12–0.83]), but the estimated RRs for our exposures of interest were unchanged. Using logistic regression, model fit was good before inclusion of complications and after inclusion of complications in the model (P values of 0.75 and 0.27, respectively, which were assessed with the Hosmer Lemeshow goodness-of-fit test).

The predicted probability of a good TKA outcome for combinations of patient willingness and acceptability of knee pain as measured on the PASS questionnaire are shown in Table 4. Models were run controlling for KOOS-PS score, PHQ-8 depression score, and comorbid hip complaints; KOOS-PS and PHQ-8 scores were held at the median value for the cohort (52.8 and 5.0, respectively), and we assumed no hip complaints. The predicted probability of a good TKA outcome ranged from 66.8% (95% CI 54.7%–77.1%) for those who were not definitely willing to undergo surgery and who had acceptable symptoms (3.5% of participants) to 86.1% (95% CI 83.0–88.7) for those who were definitely willing to undergo surgery and had unacceptable symptoms (71.1% of participants). For 17% of study participants who were definitely willing to undergo TKA but who had acceptable knee symptoms, the estimated probability of a good TKA outcome was 78.8% (95% CI 72.2–84.2).

In sensitivity analysis, the unadjusted and adjusted RRs for a good outcome among individuals who were probably willing to

Table 2. Patient-reported outcomes at 1 year post–total knee arthroplasty*

Outcome	Number (%)
Modified OMERACT–OARSI responder criteria	
a) Knee pain on the PtGA much or somewhat improved	984 (93.45)
b) ≥20% improvement in pain on the WOMAC	991 (94.1)
c) ≥20% improvement in KOOS-PS short form scores	898 (85.3)
d) Absolute change in pain of ≥1/10 on the WOMAC	1009 (95.8)
e) Absolute change of ≥10/100 on the KOOS-PS short form	877 (83.3)
Met responder criteria	837 (79.5)†
Overall satisfaction with surgical results	
Very satisfied	787 (74.7)
Somewhat satisfied	176 (16.7)
Achieved a good surgical outcome (OMERACT–OARSI score + overall satisfaction)	822 (78.1)†

* KOOS-PS = Knee Injury and Osteoarthritis Outcome Score physical function; OMERACT–OARSI = Outcome Measures in Rheumatology–Osteoarthritis Research Society International; PASS = Patient Acceptable Symptom State; PtGA = patient global assessment of disease activity; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† Statistically significant.

Table 3. Relationship between exposures of interest and achievement of a good TKA outcome (n = 992)*

Independent variables	Dependent variable (primary outcome achieved [yes versus no])	
	Unadjusted RR (95% CI)	Adjusted RR (95% CI)
Covariates		
Age (per decade)	1.00 (0.97–1.04)	
Female sex	0.96 (0.90–1.03)†	0.98 (0.915–1.04)
General Self-Efficacy (per point)	1.00 (0.99–1.01)	
WOMAC pain (per point)	1.03 (1.02–1.04)†	1.01 (1.00–1.02)
KOOS-PS (per point)	1.01 (1.005–1.01)†	1.05 (1.03–1.08)†
Perceived Arthritis Coping (per point)	0.99 (0.98–1.00)†	1.00 (0.99–1.01)
BMI (per unit)	1.00 (1.00–1.01)	
Nonmusculoskeletal comorbid conditions (Ref. 0–1)		
2 conditions	1.02 (0.92–1.34)	
+3 conditions	1.06 (0.91–1.24)	
PHQ-8 (per point)	1.005 (1.00–1.01)†	0.93 (0.875–0.985)†
Symptomatic hips (yes)	0.92 (0.84–1.00)†	0.91 (0.84–0.99)†
Symptomatic contralateral knee (yes)	1.09 (1.02–1.16)†	1.06 (1.00–1.13)
Measures of psychological readiness for TKA		
PASS questionnaire response ("Knee symptoms Unacceptable")	1.26 (1.135–1.40)†	1.14 (1.02–1.27)†
Definitely willing to undergo TKA (yes versus no)	1.25 (1.09–1.43)†	1.18 (1.04–1.35)†
TKA complication (any versus none)	0.32 (0.12–0.84)†	0.32 (0.12–0.83)†

* A good total knee arthroplasty (TKA) outcome was defined by the following: 1) TKA recipient having met the Outcome Measures in Rheumatology–Osteoarthritis Research Society International responder criteria and 2) TKA recipient indicating they were somewhat or very satisfied with their overall TKA results. 95% CI = 95% confidence interval; BMI = body mass index; KOOS-PS = Knee Injury and Osteoarthritis Outcome Score physical function; PASS PHQ-8 = Patient Health Questionnaire 8; PtGA = patient global assessment of disease activity; RR = relative risk; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† Statistically significant.

undergo TKA were 0.75 (95% CI 0.64–0.88) and 0.78 (95% CI 0.66–0.92), respectively, compared to an unadjusted RR of 1.25 (95% CI 1.09–1.43) and an adjusted RR of 1.18 (95% CI 1.03–1.34) for those who were definitely willing to undergo TKA.

DISCUSSION

In a large cohort of recipients of primary elective TKA for knee OA, we examined the relationship between preoperative measures of psychological readiness for TKA with subsequent achievement of a good TKA outcome 1 year post-TKA. Including individuals who had experienced a TKA complication, 78% of study participants experienced a good outcome based on meaningful improvement in their knee symptoms and overall satisfaction with results. Controlling for pre-TKA OA-related disability (KOOS-PS), comorbid hip symptoms and depressed mood, definite willingness to undergo TKA, and unacceptable knee symptoms were associated with a greater likelihood of achieving a good TKA outcome. Given that ensuring a patient's psychological and physical readiness for surgery is a critical component of obtaining informed consent for any operative procedure, we believe these findings are clinically important.

We conducted the current research with the view of improving appropriate provision of TKA—the proportion of TKA recipients who experience a "net benefit" from the surgery, as measured by achievement of a good outcome. Based on our prior work, we defined a "good TKA outcome" as a dichotomous composite measure that incorporated both the "journey" (measured improvements in pain and function scores) and the "destination" (perceived improvement in pain and function and satisfaction with surgical results) (37). Identifying preoperative patient factors that accurately discriminate those who will benefit

Table 4. Predicted probability of a good TKA outcome by measures of TKA readiness*

Scenario†	Number (%) of study participants	Definitely willing	PASS	Estimated probability of a good TKA outcome (95% CI)
1 (best case)	879 (71)	Yes	Unacceptable	86.1 (83.0–88.7)
2	209 (17)	Yes	Acceptable	78.8 (72.2–84.2)
3	106 (8.5)	No	Unacceptable	77.1 (68.6–83.9)
4 (worst case)	44 (3.5)	No	Acceptable	66.8 (54.7–77.1)

* 95% CI = 95% = confidence interval; PASS = Patient Acceptable Symptom State; TKA = total knee arthroplasty.
† Models were run controlling for score on the Knee injury and Osteoarthritis Outcome Score physical function (KOOS-PS) short form, score on the Patient Health Questionnaire 8 (PHQ-8) depression subscale, and the presence of comorbid hip complaints. KOOS-PS and PHQ-8 depression scores were held at the median value for the cohort (52.8 and 5.0, respectively), and we assumed no comorbid hip complaints.

from TKA versus those who will not, based on achievement of a good outcome, is essential to improving the performance of this common procedure.

In our prior work, patients with knee OA described having a "good attitude" about TKA as critical to achieving a good TKA outcome (12,14). In the present study, we assessed participants' attitudes about TKA based on their reported willingness to undergo TKA at surgeon consultation. The likelihood of achieving a good TKA outcome in the present study was 18% higher in those who were definitely willing to undergo TKA versus those who were not. In sensitivity analysis, we also showed that those who were "probably willing" to undergo TKA were less likely than those who were "definitely willing" to undergo TKA to experience a good TKA outcome. We and others have shown that, in the context of joint replacement for hip/knee OA, willingness to undergo surgery reflects patients' perceptions and beliefs about the benefits and risks of surgery and their acceptability (21,38). A potential explanation for our current findings is that individuals with a greater willingness to undergo TKA may have more realistic expectations for the postoperative disease course and be more adherent to postoperative treatment recommendations (39,40). Alternatively, greater willingness to undergo surgery may reflect greater optimism about one's health overall, which has been linked to better physical health outcomes (16,17). To our knowledge, this is the first study of TKA recipients to demonstrate a clear relationship between an individual's willingness to consider TKA and the achievement of a "net benefit" from surgery.

In prior work (11–14), patients and surgeons have agreed that TKA was appropriate for patients if their knee OA symptoms were no longer acceptable. Thus, we were also interested in the relationship of symptom acceptability, measured using the PASS questionnaire, and TKA outcome. Controlling for other factors, we found that TKA recipients with unacceptable symptoms pre-TKA were at significantly higher odds of having a good TKA outcome. Interestingly, 17% of our TKA recipients indicated that they were definitely willing to undergo TKA despite their knee symptoms being acceptable. The motivation for TKA in these individuals is unclear but may reflect secular and lifestyle trends in perceptions of TKA from one of "managing disability" to a more proactive approach of "disability prevention" (41). Although the difference between these groups did not achieve statistical significance, the estimated probability of a good TKA outcome for this subset of the cohort was 7.3% lower than it was for participants who were definitely willing to undergo TKA and had unacceptable symptoms. These findings are consistent with the notion that patient readiness for TKA is an important contributor to achieving a "net benefit" from surgery. Further research is warranted to confirm or refute the results of the present study.

Our aim was to identify novel patient factors that might help explain why, despite surgical advances, 15–30% of individuals who undergo TKA for knee OA do not achieve a good outcome. Thus, we recruited individuals receiving TKA for knee OA. As would be expected for people seeking surgeon consultation regarding TKA, the majority were definitely willing to undergo TKA at consultation. We did not reassess willingness pre-TKA as we assumed that those who consented to undergo surgery were definitely willing to have it. However, our finding of a relationship between willingness at consultation and TKA outcome suggests that not all patients who consent to surgery are convinced they should have it. Others have also found that some patients who decided to have surgery and were on the surgery waiting list remained ambivalent or uncertain about their decision (14). The observed uncertainty of some of the study participants may reflect lack of information about alternatives to surgery and their consequences or the emotional distress associated with making a choice involving risk and uncertain outcomes (14). Regardless, the findings of the present study support the need for informed and shared decision-making between patient and surgeon that incorporates assessment of underlying beliefs about the indications for, benefits of, and risks associated with primary TKA so that these factors can be weighed during shared decision-making.

To date, the strongest determinant of postoperative pain and function is preoperative pain and function (6,42,43). Controlling for preoperative knee symptoms (44–46), mental health status and comorbidity (19,47), including the presence of other musculoskeletal problems, have been consistently shown to contribute independently to pain and function outcomes (43). However, as noted earlier, this research has largely examined predictors of the magnitude of change in symptoms and/or the level of pain and disability at follow-up rather than the achievement of a good outcome. The current study found that greater preoperative physical function, decreased incidence of depressed mood, and absence of comorbid hip complaints were associated with a greater likelihood of achieving a good patient-reported TKA outcome. These findings may be useful in patient–physician discussions concerning TKA.

Greater preoperative self-efficacy has been variably linked to better arthroplasty outcomes (48). The present study found no significant relationship between preoperative general self-efficacy and achievement of a good TKA outcome. Although lower perceived arthritis coping efficacy was associated with a greater likelihood of a good TKA outcome, this relationship was attenuated and became nonsignificant after controlling for other factors, including knee OA symptom severity and willingness to undergo surgery. A potential explanation for the findings of the present study is that the effect of individuals' perceived arthritis coping efficacy on TKA outcome is mediated by willingness to consider surgery.

Strengths of our study include the large sample size, which was likely representative of the patient population in that it was inclusive of younger TKA recipients and 45 arthroplasty surgeons,

who together perform ~60% of knee replacements in the province (49). The large sample allowed for assessment of multiple measures of readiness and to control for other preoperative factors that have been linked to TKA outcome.

There are also potential study limitations. First, in Canada, consultation with an orthopedic surgeon requires referral from a primary care or other physician. This may explain the high proportion of patients with definite willingness to consider surgery prior to seeing the surgeon in consultation but should not affect the relationship between willingness and surgical outcomes. Second, due to our focus on patient appropriateness for surgery, we did not control for perioperative and postoperative factors, such as the type and adequacy of pain management or the hospital experience (50), which may also contribute to a good TKA outcome. Third, despite the size of the study sample, this study was conducted in a single Canadian province and participants predominantly self-identified as White, thus findings may not be generalizable to other settings. Finally, we did not consider the influence of knee OA severity on imaging or clinical examination as we had previously found that surgeons felt this information was important for surgical planning and approach, but not relevant in determining patient appropriateness. Worse knee OA on examination and imaging (e.g., flexion deformity or severe malalignment) may influence patients' readiness to undergo surgery and may contribute to TKA outcome. However, the direction of these relationships is unclear. Individuals with advanced changes may be less willing to undergo surgery due to fear of suboptimal outcome or complications; alternatively, these individuals may be more willing due to higher burden of pain and functional limitations.

In conclusion, controlling for preoperative knee function, comorbid hip complaints, and depressed mood, our study found that unacceptable pre-TKA knee OA symptoms and definite willingness to undergo TKA were associated with a greater likelihood of achieving a composite measure of a good TKA outcome. These findings underscore the need for enhanced, shared patient–surgeon decision-making in TKA, which considers the mental and physical readiness of patients and their willingness to undergo surgery based on an informed understanding of the risks and potential benefits in the context of their knee OA experience. Fostering patient–physician discussion about the likelihood of benefit from TKA, and thus appropriateness for surgery, has potential to improve TKA outcomes and use of limited health care resources.

ACKNOWLEDGMENTS

We would like to thank the participants from the Edmonton Bone and Joint Center (Edmonton, Alberta, Canada) who contributed to the present study, including Dr. Gordon Arnett, Dr. Robert Balyk, Dr. Jeffery Bury, Dr. John Cinats, Dr. Donald Dick, Dr. D'Arcy Durand, Dr. Lee Ekert, Dr. Don Glasgow, Dr. Robert Glasgow Sr., Dr. Gordon Goplen, Dr. Ben Herman, Dr. Catherine Hui, Dr. Larry Hunka, Dr. Hongxing Jiang, Dr. William C. Johnson, Dr. Frank Kortbeek, Dr. Guy Lavoie, Dr. Mitch Lavoie, Dr. Paul K. Leung, Dr. James Mahood, Dr. Edward Masson, Dr. Richard McLeod, Dr. James McMillan, Dr. Greg O'Connor, Dr. David Otto, Dr. Carlo Panaro, Dr. Paulose Paul, Dr. Gordon Russell, Dr. Don Weber, Dr. Colleen Weeks, Dr. Andrea Woo, Jane Squire Howden, Candace Kenyon, Anne-Marie Adachi, Jessica Beatty, Shakib Rahman, and Braden Woodhouse. We also thank the participants from the Alberta Hip & Knee Clinic (Calgary, Alberta, Canada) who contributed to the present study, including Dr. Greg Abelseth, Dr. Kelley De Souza, Dr. John Donaghy, Dr. Paul Duffy, Dr. Kelly Johnston, Dr. Robert Korley, Dr. Raul Kuchinad, Dr. Michael Monument, Dr. Maureen O'Brien, Dr. James Powell, Dr. Shannon Puloski, Dr. Ed Rendall, Dr. Alex Rezansoff, Dr. Raj Sharma, Dr. James Stewart, Dr. Scott Timmerman, Dr. Jason Werle, Tanya Reczek, and Jeffrey Depew. We also thank Bukky Dada (Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada) and Ian Stanaitis (Women's College Hospital/University of Toronto, Toronto, Ontario, Canada).

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

Study conception and design. Hawker, Conner-Spady, Bohm, Dunbar, Jones, Ravi, Noseworthy, Woodhouse, Faris, Dick, Powell, Paul, Marshall.

Acquisition of data. Hawker, Faris, Dick, Powell, Paul, Marshall.

Analysis and interpretation of data. Hawker, Faris.

REFERENCES

1. Losina E, Walensky RP, Kessler CL, Emrani PS, Reichmann WM, Wright EA, et al. Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume. Arch Intern Med 2009;169:1113–21.
2. Hawker GA, Badley EM, Croxford R, Coyte PC, Glazier RH, Guan J, et al. A population-based nested case-control study of the costs of hip and knee replacement surgery. Med Care 2009;47:732–41.
3. Kurtz SM, Ong KL, Lau E, Bozic KJ. Impact of the economic downturn on total joint replacement demand in the United States: updated projections to 2021. J Bone Joint Surg Am 2014;96:624–30.
4. Canadian Institute for Health Information. Hip and knee replacements in Canada, 2014–2015: Canadian Joint Replacement Registry annual report. URL: https://secure.cihi.ca/free_products/cjrr-annual-report-2016-en.pdf.
5. Jones CA, Beaupre LA, Johnston DW, Suarez-Almazor ME. Total joint arthroplasties: current concepts of patient outcomes after surgery. Rheum Dis Clin North Am 2007;33:71–86.
6. Brander VA, Stulberg SD, Adams AD, Harden RN, Bruehl S, Stanos SP, et al. Predicting total knee replacement pain: a prospective, observational study. Clin Orthop Relat Res 2003:27–36.
7. Dickstein R, Heffes Y, Shabtai EI, Markowitz E. Total knee arthroplasty in the elderly: patients' self-appraisal 6 and 12 months postoperatively. Gerontology 1998;44:204–10.
8. Mancuso CA, Salvati EA, Johanson NA, Peterson MG, Charlson ME. Patients' expectations and satisfaction with total hip arthroplasty. J Arthroplasty 1997;12:387–96.
9. Woolhead G, Donovan J, Dieppe P. Patient expectations and total joint arthroplasty. J Rheumatol 2003:1656–7.
10. Sanmartin C, Murphy K, Choptain N, Conner-Spady B, McLaren L, Bohm E, et al. Appropriateness of healthcare interventions: concepts

and scoping of the published literature. Int J Technol Assess Health Care 2008;24:342–9.

11. Hawker G, Bohm ER, Conner-Spady B, de Coster C, Dunbar M, Hennigar A, et al. Perspectives of Canadian stakeholders on criteria for appropriateness for total joint arthroplasty in patients with hip and knee osteoarthritis. Arthritis Rheumatol 2015;67:1806–15.
12. Frankel L, Sanmartin C, Conner-Spady B, Marshall DA, Freeman-Collins L, Wall A, et al. Osteoarthritis patients' perceptions of "appropriateness" for total joint replacement surgery. Osteoarthritis Cartilage 2012;20:967–73.
13. Frankel L, Sanmartin C, Hawker G, de Coster C, Dunbar M, Bohm E, et al. Perspectives of orthopaedic surgeons on patients' appropriateness for total joint arthroplasty: a qualitative study. J Eval Clin Pract 2016;22:164–70.
14. Conner-Spady BL, Marshall DA, Hawker GA, Bohm E, Dunbar MJ, Frank C, et al. You'll know when you're ready: a qualitative study exploring how patients decide when the time is right for joint replacement surgery. BMC Health Serv Res 2014;14:454.
15. Wylde V, Dixon S, Blom AW. The role of preoperative self-efficacy in predicting outcome after total knee replacement. Musculoskeletal Care 2012;10:110–8.
16. Rasmussen HN, Scheier MF, Greenhouse JB. Optimism and physical health: a meta-analytic review. Ann Behav Med 2009;37:239–56.
17. Waldrop D, Lightsey OR Jr, Ethington CA, Woemmel CA, Coke AL. Self-efficacy, optimism, health competence, and recovery from orthopedic surgery. J Counsel Psychol 2001;48:233–8
18. Scheier MF, Matthews KA, Owens JF, Schulz R, Bridges MW, Magovern GJ, et al. Optimism and rehospitalization after coronary artery bypass graft surgery. Arch Intern Med 1999;159:829835.
19. Singh JA, Lewallen DG. Medical and psychological comorbidity predicts poor pain outcomes after total knee arthroplasty. Rheumatology (Oxford) 2013;52:916–23.
20. Balck F, Lippmann M, Jeszenszky C, Gunther KP, Kirschner S. The influence of optimism on functionality after total hip replacement surgery. J Health Psychol 2016;21:1758–67.
21. Hawker GA, Wright JG, Badley EM, Coyte PC. Perceptions of, and willingness to consider, total joint arthroplasty in a population-based cohort of individuals with disabling hip and knee arthritis. Arthritis Rheum 2004;51:635–41.
22. Hawker GA, Guan J, Croxford R, Coyte PC, Glazier RH, Harvey BJ, et al. A prospective population-based study of the predictors of undergoing total joint arthroplasty. Arthritis Rheum 2006;54:3212–20.
23. Ravi B, Croxford R, Reichmann WM, Losina E, Katz JN, Hawker GA. The changing demographics of total joint arthroplasty recipients in the United States and Ontario from 2001 to 2007. Best Pract Res Clin Rheumatol 2012;26:637–47.
24. Leskinen J, Eskelinen A, Huhtala H, Paavolainen P, Remes V. The incidence of knee arthroplasty for primary osteoarthritis grows rapidly among baby boomers: a population-based study in Finland. Arthritis Rheum 2012;64:423–8.
25. Bellamy N. Pain assessment in osteoarthritis: experience with the WOMAC osteoarthritis index. Semin Arthritis Rheum 1989;18:14–7.
26. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
27. Perruccio AV, Lohmander SL, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS): an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16:542–50.
28. Gignac MA, Cott C, Badley EM. Adaptation to chronic illness and disability and its relationship to perceptions of independence and dependence. J Gerontol B Psychol Sci Soc Sci 2000;55:P362–72.
29. Gignac MA. An evaluation of a psychotherapeutic group intervention for persons having difficulty coping with musculoskeletal disorders. Soc Work Health Care 2000;32:57–75.
30. Schwartzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in health psychology: a user's portfolio. Causal and Control Beliefs. Windsor, England: Nfer-Nelson; 1995.
31. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009;114:163–73.
32. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann Rheum Dis 2005;64:34–7.
33. Mahomed N, Gandhi R, Daltroy L, Katz JN. The self-administered patient satisfaction scale for primary hip and knee arthroplasty. Arthritis 2011;2011:591253.
34. Pham T, van der Heijde D, Altman RD, Anderson JJ, Bellamy N, Hochberg M, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. Osteoarthritis Cartilage 2004;12:389–99.
35. Freund R, Wilson W. Regression analysis: statistical modeling of a response variable. 1st ed. Academic Press; 1998.
36. Zou K, Wong J, Abdullah N, Abdullah N, Chen X, Smith T, et al. Examination of overall treatment effect and the proportion attributable to contextual effect in osteoarthritis: meta-analysis of randomised controlled trials. Ann Rheum Dis 2016;75:1964–70.
37. Katz JN. Appropriateness of total knee arthroplasty [editorial]. Arthritis Rheumatol 2014;66:1979–81.
38. Allen KD, Golightly YM, Callahan LF, Helmick CG, Ibrahim SA, Kwoh CK, et al. Race and sex differences in willingness to undergo total joint replacement: the Johnston County Osteoarthritis Project. Arthritis Care Res (Hoboken) 2014;66:1193–202.
39. Leedham B, Meyerowitz BE, Muirhead J, Frist WH. Positive expectations predict health after heart transplantation. Health Psychol 1995; 14:74–9.
40. Mondloch MV, Cole DC, Frank JW. Does how you do depend on how you think you'll do? A systematic review of the evidence for a relation between patients' recovery expectations and health outcomes. CMAJ 2001;165:174–9.
41. Losina E, Thornhill TS, Rome BN, Wright J, Katz JN. The dramatic increase in total knee replacement utilization rates in the United States cannot be fully explained by growth in population size and the obesity epidemic. J Bone Joint Surg Am 2012;94:201–7.
42. Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. J Bone Joint Surg Am 2004;86-a:2179–86.
43. Hawker GA, Badley EM, Borkhoff CM, Croxford R, Davis AM, Dunn S, et al. Which patients are most likely to benefit from total joint arthroplasty? Arthritis Rheum 2013;65:1243–52.
44. Judge A, Arden NK, Cooper C, Javaid MK, Carr AJ, Field RE, et al. Predictors of outcomes of total knee replacement surgery. Rheumatology (Oxford) 2012;51:1804–13.
45. Baker PN, van der Meulen JH, Lewsey J, Gregg PJ. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. J Bone Joint Surg Br 2007;89:893–900.
46. Clement ND, Macdonald D, Burnett R. Predicting patient satisfaction using the Oxford knee score: where do we draw the line? Arch Orthop Trauma Surg 2013;133:689–94.

47. Perruccio AV, Power JD, Evans HM, Mahomed SR, Gandhi R, Mahomed NN, et al. Multiple joint involvement in total knee replacement for osteoarthritis: effects on patient-reported outcomes. Arthritis Care Res (Hoboken) 2012;64:838–46.
48. Magklara E, Burton CR, Morrison V. Does self-efficacy influence recovery and well-being in osteoarthritis patients undergoing joint replacement? A systematic review. Clin Rehabil 2014;28: 835–46.
49. Alberta Bone & Joint Health Institute. URL: https://albertaboneandjoint.com/.
50. Hamilton DF, Lane JV, Gaston P, Patton JT, Macdonald D, Simpson AH, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. BMJ Open 2013;3:e002525.

APPENDIX 1: MEMBERS OF THE BEST-KNEE STUDY TEAM

Members of the BEST-Knee Study Team are as follows: Dr. Gillian A. Hawker (University of Toronto, Toronto, Ontario), Dr. Deborah A. Marshall (University of Calgary, Calgary, Alberta), Dr. Eric Bohm (University of Manitoba, Winnipeg, Manitoba), Dr. Michael J. Dunbar (Dalhousie University, Halifax, Nova Scotia), Dr. Peter Faris (University of Calgary, Calgary, Alberta), Dr. C. Allyson Jones (University of Alberta, Edmonton, Alberta), Dr. Tom Noseworthy (University of Calgary, Calgary, Alberta), Dr. Bheeshma Ravi (University of Toronto, Toronto, Ontario), and Dr. Linda Woodhouse (University of Alberta, Edmonton, Alberta, and Tufts University, Phoenix, Arizona).

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1384–1390
DOI 10.1002/acr.24588

Association Between Baseline Meniscal Symptoms and Outcomes of Operative and Nonoperative Treatment of Meniscal Tear in Patients With Osteoarthritis

Lindsey A. MacFarlane,[1] Heidi Yang,[2] Jamie E. Collins,[1] Robert H. Brophy,[3] Brian J. Cole,[4] Kurt P. Spindler,[5] Ali Guermazi,[6] Morgan H. Jones,[5] Lisa A. Mandl,[7] Scott Martin,[8] Robert G. Marx,[7] Bruce A. Levy,[9] Michael Stuart,[9] Clare Safran-Norton,[2] John Wright,[10] Rick W. Wright,[3] Elena Losina,[1] and Jeffrey N. Katz[1]

Objective. Patients with meniscal tears reporting meniscal symptoms such as catching or locking have traditionally undergone arthroscopy. The present study was undertaken to investigate whether patients with meniscal tears who report meniscal symptoms have greater improvement with arthroscopic partial meniscectomy (APM) than physical therapy (PT).

Methods. We used data from the Meniscal Tear in Osteoarthritis Research (MeTeOR) trial, which randomized participants with knee osteoarthritis (OA) and meniscal tear to APM or PT. The frequency of each meniscal symptom (clicking, catching, popping, intermittent locking, giving way, swelling) was measured at baseline and 6 months. We used linear regression models to determine whether the difference in improvement in Knee Injury and Osteoarthritis Outcome Score (KOOS) pain score at 6 months between patients treated with APM versus PT was modified by the presence of each meniscal symptom. We also determined the percentage of participants with resolution of meniscal symptoms by treatment group.

Results. We included 287 participants. The presence (versus absence) of any of the meniscal symptoms did not modify the improvement in KOOS pain score between APM versus PT by >0.5 SD (all *P* interaction >0.05). APM led to greater resolution of intermittent locking and clicking than PT (locking 70% versus 46%, clicking 41% versus 25%). No difference in resolution of the other meniscal symptoms was observed.

Conclusion. Meniscal symptoms were not associated with improved pain relief. Although symptoms of clicking and intermittent locking had a greater reduction in the APM group, the presence of meniscal symptoms in isolation should not inform clinical decisions surrounding APM versus PT in patients with meniscal tear and knee OA.

INTRODUCTION

Symptomatic knee osteoarthritis (OA) affects an estimated 14 million individuals in the US, with up to 91% of patients with knee OA demonstrating a meniscal tear on magnetic resonance imaging (MRI) (1,2). Knee symptoms such as catching, popping, or locking elicited in young persons with acute injuries have been considered mechanical symptoms. Historically, these mechanical

ClinicalTrials.gov identifier: NCT00597012.

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grants R01-AR-05557, T32-AR-055885, K24-AR-057827, and P30-AR-072577) and the Rheumatology Research Foundation (Scientific Development award).

[1]Lindsey A. MacFarlane, MD, MPH, Jamie E. Collins, PhD, Elena Losina, PhD, Jeffrey N. Katz, MD, MSc: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [2]Heidi Yang, MS, MPH, Clare Safran-Norton, PT, PhD: Brigham and Women's Hospital, Boston, Massachusetts; [3]Robert H. Brophy, MD, Rick W. Wright, MD: Washington University School of Medicine, Saint Louis, Missouri; [4]Brian J. Cole, MD, MBA: Rush University, Chicago, Illinois; [5]Kurt P. Spindler, MD, Morgan H. Jones, MD: Cleveland Clinic, Cleveland, Ohio; [6]Ali Guermazi, MD, PhD, MPH: VA Boston Healthcare System, Boston University School of Medicine, Boston, Massachusetts; [7]Lisa A. Mandl, MD, MPH, Robert G. Marx, MD: Hospital for Special Surgery, Weill Cornell Medicine, New York, New York; [8]Scott Martin, MD: Massachusetts General Hospital, Boston, Massachusetts; [9]Bruce A. Levy, MD, Michael Stuart, MD: Mayo Clinic, Rochester, Minnesota; [10]John Wright, MD: Johnson & Johnson, Raynham, Massachusetts.

Dr. MacFarlane has received research support from Samumed. Dr. Collins has received consulting fees from Boston Imaging Core Lab (less than $10,000). Dr. Cole has received consulting fees, speaking fees, and/or honoraria from Aesculap, Smith and Nephew, Bandgrip, Acumed, Encore Medical, GE Healthcare, Merck Sharp & Dohme, SportsTek Medical, Vericel Corporation, Elsevier, Ossio, Regentis (less than $10,000 each), and Arthrex (more than $10,000). Dr. Spindler has received consulting fees from Mitek, Flexion Therapeutics, Samumed, and Novopeds (less than $10,000 each). Dr. Guermazi has received consulting fees, speaking fees, and/or honoraria from Boston Imaging Core Lab, AstraZeneca, Roche, Galapagos (less than $10,000 each), TissueGene, Pfizer, and MerckSerono (more than $10,000 each). Dr. Jones has received consulting fees, speaking fees, and/or honoraria from Samumed and Regeneron (less than $10,000 each) and research support from Flexion Therapeutics. Dr. Mandl has received royalties from UpToDate and research support from Regeneron. Dr. Marx has received consulting fees, speaking fees, and/or honoraria from Mend Nutrition (less than $10,000) and royalties from Springer and Demos Health. Dr. Levy has received consulting fees from Arthrex and Smith and Nephew (less than $10,000 each). Dr. Stuart has received consulting fees and royalties from Arthrex (more than $10,000) and research support from Arthrex and Stryker. Dr. Safran-Norton owns stock or stock options in Merck and Johnson & Johnson. Dr. J. Wright owns stock or stock options in Johnson & Johnson. Dr. R. Wright has received consulting fees and royalties from Responsive

SIGNIFICANCE & INNOVATIONS

- We leveraged data from a randomized control trial to evaluate the association between meniscal symptoms and knee pain after surgery versus physical therapy for meniscal tear.
- The presence or absence of meniscal symptoms was not associated with differential pain outcomes after surgery versus physical therapy.
- Knee symptoms such as clicking and catching have historically been ascribed to meniscal pathology. These data cast further doubt on the ability of meniscal symptoms to help direct management of meniscal tear.

symptoms along with symptoms such as pain with twisting have been grouped together as "meniscal symptoms" and attributed to meniscal tear or to other internal derangements. These patients were often referred to orthopedic surgeons for consideration of arthroscopic diagnosis and management. However, in the current era of advanced imaging, meniscal tear can be visualized on MRI, obviating the need for direct surgical visualization.

Middle-aged and older patients with knee OA frequently report meniscal symptoms; clinicians continue to suspect symptomatic meniscal tear in these patients, even though there is little evidence that meniscal symptoms stem from meniscal pathology in older patients with degenerative (rather than traumatic) meniscal tears (3). In fact, prior evidence suggests that meniscal tears were seen in a similar proportion of asymptomatic and symptomatic knees (4).

Randomized trials comparing surgical treatment versus conservative therapy for patients with degenerative meniscal tears have found that both strategies reduce pain (5–8). Although meniscal symptoms may not be specific to meniscal damage, many clinicians feel that patients with meniscal symptoms may represent a subgroup with a favorable response to arthroscopic partial meniscectomy (APM) (9), as resection of the torn meniscus is thought to aid in restoring smooth joint motion. Therefore, there is considerable interest in whether patients with meniscal symptoms might benefit more from surgery than those without these symptoms.

Our group previously developed a more comprehensive list of commonly considered meniscal symptoms based on input from physicians, physical therapists, and patients. While the original list included several pain parameters, here we focus on the traditional mechanical or meniscal symptoms including clicking, catching, popping, intermittent locking, giving way, and swelling of the knee (10). We sought to evaluate whether patients reporting any of these expanded meniscal symptoms had greater improvement with APM than with physical therapy (PT) using data from the Meniscal Tear in Osteoarthritis Research (MeTeOR) trial, a randomized trial of APM versus PT in patients with knee OA and meniscal tear (6). We evaluated the association between these expanded meniscal symptoms, treatment group (APM, PT), and patient improvement. Here, we test the null hypothesis that in patients with OA, the association between treatment group and 6-month change in pain will not be different for those with versus those without baseline meniscal symptoms.

PATIENTS AND METHODS

Study sample. We used data from participants in the MeTeOR trial; details of this trial have been previously published (6,11). Three hundred fifty-one subjects were recruited from 7 academic centers from 2008 through 2011. Participants included males and females age ≥45 years who had at least 4 weeks of knee pain and an MRI with evidence of a meniscal tear extending to the meniscal surface in at least 2 consecutive slices. Included participants also had evidence of mild-to-moderate osteoarthritic change (Kellgren/Lawrence [K/L] grade ≤3) as determined by osteophyte and/or joint space narrowing on plain radiographs, or full-thickness articular cartilage defect on at least 1 tibial or femoral surface on MRI. We excluded patients with a chronically locked knee (e.g., subject unable to flex or extend knee on examination), inflammatory arthritis, prior surgery on the index knee, K/L grade 4 OA, and contraindication to MRI. Participants reporting locking but able to flex and extend the knee on examination were included and are designated as "intermittent locking" in our analyses. Participants were randomized either to PT or to APM followed by the PT regimen. The surgical intervention was APM with resection of the damaged meniscus back to a stable rim. Meniscal repairs were not permitted as part of the trial. All participants provided consent, and the study was approved by the Partners HealthCare Human Research Committee (2005P000440). This trial is registered at ClinicalTrials.gov (NCT00597012).

Data elements. We collected data on age, sex, and body mass index (BMI, kg/m^2) at baseline. The frequency of patient-reported meniscal symptoms was obtained at baseline and 6-month follow-up. Meniscal symptoms included clicking, catching, popping, intermittent locking, giving way, and knee swelling. Questionnaires assessed frequency of each meniscal symptom as follows: none; once/week; 2–6 times/week; 1–2 times/day; and several times/day. Based on the distribution of the

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

Address correspondence to Lindsey A. Macfarlane, MD, MPH, Brigham and Women's Hospital, 60 Fenwood Road, Boston, MA 02115. Email: lmacfarlane@bwh.harvard.edu.

Submitted for publication July 21, 2020; accepted in revised form February 25, 2021.

categorical responses, all meniscal symptoms were dichotomized to "none" versus "any." Radiographic severity of OA was measured at baseline using the K/L grade (12).

Outcome and assessment. The primary outcome of interest was change in patient-reported pain from baseline to 6-month follow-up assessed with the Knee Injury and Osteoarthritis Outcome Score (KOOS) pain scale (13). We transformed KOOS pain score to a 0–100 scale, with 0 being least amount of pain, and 100 the greatest, with negative change indicative of improvement.

Statistical analysis. We described baseline characteristics of the cohort using means and percentages. For the primary analysis, we excluded the participants crossing over from PT to APM prior to 6 months, as these participants could be early in the recovery process at the 6-month assessment, as well as participants randomized to APM who did not undergo the surgery. The participants crossing over from PT to APM after 6 months were included in the PT arm. Participants missing either baseline or 6-month KOOS pain scores were excluded. We built separate multivariable linear regressions for each meniscal symptom (clicking, catching, popping, intermittent locking, giving way, and swelling), with the dichotomous symptom variable as the independent variable, and change in KOOS pain score from baseline to 6 months as the dependent variable. We examined the interaction between each meniscal symptom and treatment type (APM, PT) on change in KOOS pain score. All models were adjusted for age, sex, BMI, and baseline KOOS pain score.

In a second set of models, we also adjusted for K/L grade to account for radiographic OA severity. In another sensitivity analysis, we dichotomized meniscal symptoms as "less than daily" (none; once/week; 2–6 times/week) versus "daily" (1–2 times/day and several times/day) to allow investigation of participants with more frequent meniscal symptoms. The original MeTeOR trial was not powered to detect these interactions; thus, these analyses are intended to be hypothesis generating.

To address potential bias due to the exclusion of crossovers from the primary analysis, we assessed whether including crossovers in the APM arm or the PT arm altered results through 2 sensitivity analyses. In the first, we used an intent-to-treat approach, in which we included participants crossing over from PT to APM prior to 6 months and participants crossing over from PT to APM after 6 months in the PT arm. The second analysis used an as-treated approach including participants crossing over from APM to PT prior to 6 months in the APM arm. (The participants crossing over after 6 months were kept in the PT arm, as the primary outcome was at 6 months.)

As a secondary analysis, we investigated resolution of meniscal symptoms from baseline to 6-month follow-up. From the subset of participants reporting any meniscal symptoms at baseline, we defined "resolution" as those participants reporting "none" at follow-up. We investigated differences in this outcome across each treatment category, APM and PT. In this analysis, we included participants crossing over from PT to APM after 6 months in the PT arm, and excluded patients crossing over between arms during the first 6 months. Participants with missing 6-month meniscal symptom data were considered "non-resolvers" rather than omitted, as this was felt to be the most conservative analytic approach. We used contingency tables and the chi-square test to assess for statistically significant differences in percent improvement among the treatment groups. For all analyses *P* values less than 0.05 were considered statistically significant. All analyses were performed using SAS statistical software, version 9.4.

RESULTS

Of the 351 participants,164 participants (47%) were randomized to and received APM. One hundred nine (31%) were randomized to PT and did not crossover; 14 participants (4%) were randomized to PT but crossed over after 6 months and were therefore included in the PT arm. Ten participants (3%) were randomized to APM but did not have the procedure, and 54 (15%) were randomized to PT but received APM within 6 months and were excluded from analysis. The primary analysis included the 287 participants (82%) who were randomized to and received APM or were randomized to and received PT in the first 6 months. Mean age and BMI were similar among the treatment groups. Clicking, catching, popping, and giving way were present in 48–67% of participants at baseline. Twenty-seven to 31% of participants reported intermittent locking, and 71–75% reported swelling. The percentage of participants reporting each meniscal symptom by treatment group at baseline is outlined in Table 1.

Primary analysis. In the primary analysis, participants randomized to and receiving APM were considered in the APM group (n = 164), and those randomized to PT who remained in the PT group at least until 6 months were considered in the PT group (n = 123). Six-month change in KOOS pain score was missing in 23 participants in the APM group and 20 participants in the PT group; thus, the final analysis included 141 in the APM group and 103 in the PT group. Overall, regardless of meniscal symptoms at baseline, those undergoing APM had slightly greater improvement in KOOS pain scores at 6 months compared to PT. In the individual models for meniscal symptoms after adjustment for age, sex, BMI, and baseline KOOS pain score, participants without clicking, catching, popping, or locking and with giving way and swelling had a small but greater improvement in KOOS pain score after APM than PT. Assuming that the SD of KOOS pain score is 15 (14), the differences correspond to an effect size of 0.3–0.5. Participants with clicking, catching, popping, or locking and without giving way and swelling had minimal

Table 1. Baseline characteristics by treatment group*

Characteristic	Arthroscopic partial meniscectomy (n = 164)	Physical therapy (n = 123)	*P*
Age, mean ± SD years	59 ± 8	58 ± 6	0.08
BMI, mean ± SD	30 ± 6	30 ± 6	0.99
Female	94 (57)	67 (54)	0.63
KOOS pain score, mean ± SD	46 ± 16	46 ± 16	0.74
Kellgren/Lawrence grade			0.29
0	14 (9)	10 (8)	
1	31 (19)	34 (28)	
2	63 (38)	37 (30)	
3	56 (34)	42 (34)	
Meniscal symptoms			
Clicking	106 (66)	80 (67)	0.81
Catching	81 (51)	62 (52)	0.86
Popping	79 (50)	61 (51)	0.85
Intermittent locking	43 (27)	37 (31)	0.42
Giving way	77 (48)	64 (54)	0.27
Swelling	114 (71)	90 (75)	0.49

* Values are the number (%) unless indicated otherwise. Of the 287 participants, data were missing on 0–6% for each baseline characteristic. BMI = body mass index; KOOS = Knee Injury and Osteoarthritis Outcome Score.

differences in KOOS pain score between APM and PT. All interaction *P* values were > 0.09 (Table 2). Further adjusting models for K/L grade did not alter results. The results of this analysis did not change when meniscal symptoms were considered as daily versus less than daily, aside from swelling, where those with daily and less than daily swelling had a 3- and 5-point greater improvement with APM, respectively (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24588). As in the primary analysis, the effect of each meniscal symptom on change in KOOS pain score over 6 months was not modified by treatment (*P* value for interaction >0.05 for each symptom).

Sensitivity analyses. In the sensitivity analysis, using an intent-to-treat approach, 177 participants were categorized as PT, and 164 as APM. Change in KOOS pain score over 6 months was missing in 26 participants in the PT group, and 23 in the APM group. At baseline, the crossover group had a higher percentage of female participants at 65%, versus 57% for APM and 54% for PT. Mean baseline KOOS pain score was also greater in the crossover participants at 51, versus 46 for both APM and PT (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24588). Results were analogous to the primary analysis, and the presence or absence of each meniscal symptom and

Table 2. Change in Knee Injury and Osteoarthritis Outcome Score (KOOS) pain score for arthroscopic partial meniscectomy (APM) and physical therapy (PT) by presence of meniscal symptom*

	Mean change in KOOS pain score from baseline to 6 months (95% CI)†			Difference (95% CI)‡	*P* for interaction
	PT	APM	*P*		
Clicking	18.9 (14.7, 23.2)	22.2 (18.6, 25.8)	0.24	−3.3 (−8.8, 2.2)	0.58
No clicking	21.9 (16.4, 27.4)	27.8 (22.9, 32.7)	0.11	−5.9 (−13.1, 1.4)	0.58
Catching	20.1 (15.3, 24.8)	21.9 (17.8, 25.9)	0.56	−1.8 (−8.0, 4.3)	0.37
No catching	20.6 (15.9, 25.4)	26.5 (22.2, 30.7)	0.07	−5.8 (−12.1, 0.4)	0.37
Popping	18.9 (13.9, 23.8)	21.2 (17.1, 25.2)	0.47	−2.3 (−8.7, 4.0)	0.41
No popping	21.6 (17.0, 26.1)	27.6 (23.5, 31.6)	0.05	−6.0 (−12.0, 0.04)	0.41
Intermittent locking	21.2 (15.2, 27.1)	21.5 (15.7, 27.3)	0.94	−0.3 (−8.6, 7.9)	0.32
No intermittent locking	19.9 (15.9, 24.0)	25.2 (21.8, 28.5)	0.05	−5.2 (−10.5, −0.02)	0.32
Giving way	17.8 (13.0, 22.5)	25.7 (21.5, 29.8)	0.01	−7.9 (−14.1, −1.7)	0.09
No giving way	22.7 (17.8, 27.5)	22.9 (19.0, 26.9)	0.93	−0.3 (−6.4, 5.9)	0.09
Swelling	19.8 (16.0, 23.7)	25.9 (22.4, 29.3)	0.02	−6.0 (−11.1, −0.9)	0.12
No swelling	22.5 (15.7, 29.3)	20.6 (15.1, 26.1)	0.66	1.9 (−6.7, 10.4)	0.12

* 95% CI = 95% confidence interval.
† Adjusted for age, sex, body mass index, and baseline KOOS pain score.
‡ Negative values favor APM, and positive values favor PT.

Table 3. Proportion of participants with resolution of meniscal symptoms over 6 months by each treatment category*

Resolution	Clicking	Catching	Popping	Locking	Giving way	Swelling
APM	43 (41)†	48 (59)	39 (49)	30 (70)†	44 (57)	43 (38)
PT	20 (25)†	31 (50)	23 (38)	17 (46)†	35 (55)	36 (40)
RR (95% CI)‡	1.62 (1.04, 2.53)	1.19 (0.87, 1.61)	1.31 (0.88, 1.94)	1.52 (1.02, 2.27)	1.04 (0.78, 1.40)	0.94 (0.67, 1.33)

* Values are the number (%) unless indicated otherwise. Resolution reflects participants with any symptom at baseline and none at 6-month follow-up, if data missing at 6 months are regarded as no resolution. APM = arthroscopic partial meniscectomy; PT = physical therapy; RR = relative risk.
† $P < 0.05$ for difference in resolution between APM and PT groups.
‡ RR >1 favors APM.

treatment type did not clinically or statistically significantly modify the change in KOOS pain score at 6 months (see "intention to treat" in Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24588). In the second sensitivity analysis including those crossing over from PT to APM prior to 6 months in the APM arm, 218 were categorized as APM, and 123 as PT. Change in KOOS pain score over 6 months was missing in 29 participants in the APM arm and in 20 participants in the PT arm. Again, the results were similar to the primary analysis (see "as treated" in Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24588).

Secondary outcome. In this analysis, participants crossing over from PT to APM after 6 months were included in the PT arm, while those crossing over before 6 months were excluded. For each meniscal symptom, 14–32 participants did not provide 6-month data, and missingness did not vary between treatment. At 6 months, the percentage of participants with resolution (reporting any meniscal symptom at baseline and none at 6-month follow-up) for clicking, catching, popping, intermittent locking, and giving way was greater in those undergoing APM. Among those undergoing PT, clicking resolved in 25%, catching in 50%, popping in 38%, locking in 46%, and giving way in 55%. Improvement in swelling was greater in the PT group than in those receiving APM (Table 3). The greater extent of resolution in intermittent locking and clicking in the APM group as compared with the PT group was statistically significant ($P < 0.05$).

DISCUSSION

Our study suggests that, in general, individuals with OA and meniscal symptoms do not have greater clinically meaningful improvement in pain after APM compared with PT. The differences in 6-month change KOOS pain score between APM and PT did not exceed 7.9 points; as the minimum clinically important difference for KOOS pain score is 8–15 points (15,16), these differences are unlikely to be clinically meaningful (15). In separate analyses for each symptom, the presence of clicking, catching, popping, intermittent locking, and swelling at baseline did not demonstrate a statistically significant or clinically greater improvement in 6-month pain outcomes with APM than with PT. Only with presence of giving way did the difference between APM and PT reach an effect size of 0.5, indicating a moderate effect. While a greater proportion of participants undergoing APM reported improvement in clicking, catching, popping, giving way, and intermittent locking over 6 months, only intermittent locking and clicking showed statistically significant differences in improvement among the treatment groups. This suggests that APM may offer greater relief of clicking and intermittent locking meniscal symptoms than PT despite not offering greater relief of pain.

Our findings are comparable to those of 2 recent studies. Sihvonen et al (16) analyzed data from the FIDELITY trial, in which participants with meniscal tear without knee OA were randomized to APM versus sham surgery, to evaluate whether participants with meniscal symptoms (sensation of catching or locking) had greater improvement with APM. Results demonstrated no significant difference in the prevalence of meniscal symptoms after APM versus sham surgery at 2, 6, or 12 months (16). Our study differs in that we found that APM was more likely to relieve intermittent locking and clicking than PT. But, like Sihvonen et al, we also found that relief in overall pain was not influenced by meniscal symptoms (17).

Gauffin et al included patients with meniscal tear and Ahlbäck grade 0 knee OA (<50% joint space narrowing) randomized to exercise versus APM. Secondary analyses of this study showed no effect of meniscal symptoms (catching or locking for >2 seconds) or interaction between meniscal symptoms and treatment on change in KOOS pain score at 3-year follow-up (18). Similar results were seen in the main trial with 1-year follow-up (19). However, the 3-year as-treated data also found that participants with meniscal symptoms had less improvement in KOOS pain score with APM. The 5-year follow-up data from this study again demonstrated a statistically significant greater reduction in KOOS pain score for those without meniscal symptoms in the APM group (20). As noted by Gauffin et al, meniscal symptoms may be nonspecific and not necessarily reflect meniscal pathology (18). Regardless, our study adds to this body of literature by evaluating a broader range of meniscal symptoms and again suggests that traditional meniscal symptoms do not clearly relate to meniscal pathology in patients with OA, as assessed by response to partial meniscal resection.

Orthopedic surgeons generally assert that the decision to refer a patient with meniscal tear for surgical evaluation should

not be based on the presence of meniscal symptoms alone but be grounded in the surgeon's clinical judgement and patient preference. We acknowledge the wide range of views on this important topic and encourage additional research, such as ours, to clarify unresolved questions regarding the nature of meniscal symptoms and their role in selecting patients for treatment (19,21–23). Prior work from our group using MeTeOR data has shown that patients with fewer osteoarthritic changes on MRI (bone marrow lesions and cartilage damage) have greater improvement of pain with APM than with PT, while those with more substantial OA changes have similar outcomes regardless of whether they undergo APM or PT (24). Therefore, clinical features such as extent of underlying OA and tear type may be more salient to the initial surgical decision than the presence or frequency of meniscal symptoms (24,25). In our study, adjusting for K/L grade, a radiographic marker of OA severity, did not alter results. However, it is likely K/L grade is not sensitive enough to capture underlying pathology. Overall, studies on the use of APM for treatment of meniscal tear have not found APM to be superior to PT (6,7,8,26), although Gauffin et al (19) found benefit to APM and PT compared with PT alone. Based on the current evidence, there are no widely accepted criteria for identifying patients more likely to improve from APM than from PT.

Our study has several limitations. Thirty-one percent (n = 54) of the participants randomized to PT crossed over to APM over 6 months. To address bias from excluding these participants, we included them in intent-to-treat and as-treated analyses. The results of these analyses were similar to those of the primary analysis. We excluded participants without complete 6-month KOOS pain score data, which may introduce bias. As this study is a secondary analysis of MeTeOR data, we have limited power to detect interactions. We did not correct for multiple comparisons and thus recommend caution in interpretation. The follow-up period was 6 months; therefore, we are unable to assess if these results are durable. Meniscal symptoms, including intermittent locking, fluctuate over time. We cannot rule out that any observed improvement was due to chance or natural disease course instead of treatment, and additional confirmatory studies are warranted. Last, as all patients had OA changes in addition to meniscal tear, we were unable to ascertain whether the etiology of the meniscal symptoms was indeed the meniscus or other sources such as damage to cartilage or surrounding structures. Finally, we cannot use these data to draw conclusions regarding younger patients with traumatic-type tears.

In conclusion, our results suggest that in our patients with mild-to-moderate knee OA and meniscal tear, the presence of self-reported clicking, catching, popping, intermittent locking, or swelling does not identify a subgroup that is more likely to have pain relief following APM. Although symptoms of clicking and intermittent locking had a greater reduction in the APM group, the presence of meniscal symptoms in isolation is not sufficient to make a clinical decision regarding APM versus PT for the reduction of pain in this patient population, and further clinical data points must be considered, including patient characteristics, physical examination results, and imaging findings.

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. MacFarlane 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. MacFarlane, Yang, Collins, Brophy, Cole, Spindler, Guermazi, Jones, Mandl, Martin, Marx, Levy, Stuart, Safran-Norton, Wright, Wright, Losina, Katz.

Acquisition of data. Brophy, Cole, Spindler, Guermazi, Jones, Martin, Marx, Levy, Stuart, Safran-Norton, Wright, Wright, Losina, Katz.

Analysis and interpretation of data. MacFarlane, Yang, Collins, Brophy, Cole, Spindler, Guermazi, Jones, Mandl, Martin, Marx, Levy, Stuart, Safran-Norton, Wright, Wright, Losina, Katz.

ADDITIONAL DISCLOSURES

Author J. Wright is an employee of Johnson & Johnson.

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. Bhattacharyya T, Gale D, Dewire P, Totterman S, Gale ME, McLaughlin S, et al. The clinical importance of meniscal tears demonstrated by magnetic resonance imaging in osteoarthritis of the knee. J Bone Joint Surg Am 2003;85:4–9.
3. Hare KB, Stefan Lohmander L, Kise NJ, Risberg MA, Roos EM. Middle-aged patients with an MRI-verified medial meniscal tear report symptoms commonly associated with knee osteoarthritis. Acta Orthop 2017;88:664–9.
4. Englund M, Guermazi A, Gale D, Hunter DJ, Aliabadi P, Clancy M, et al. Incidental meniscal findings on knee MRI in middle-aged and elderly persons. N Engl J Med 2008;359:1108–15.
5. Herrlin S, Hållander M, Wange P, Weidenhielm L, Werner S. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. Knee Surg Sports Traumatol Arthrosc 2007;15:393–401.
6. Katz JN, Brophy RH, Chaisson CE, de Chaves L, Cole BJ, Dahm DL, et al. Surgery versus physical therapy for a meniscal tear and osteoarthritis. N Engl J Med 2013;368:1675–84.
7. Yim JH, Seon JK, Song EK, Choi JI, Kim MC, Lee KB, et al. A comparative study of meniscectomy and nonoperative treatment for degenerative horizontal tears of the medial meniscus. Am J Sports Med 2013; 41:1565–70.
8. Kise NJ, Risberg MA, Stensrud S, Ranstam J, Engebretsen L, Roos EM. Exercise therapy versus arthroscopic partial meniscectomy for degenerative meniscal tear in middle aged patients: randomised controlled trial with two year follow-up. BMJ 2016;354:i3740.
9. Krych AJ, Carey JL, Marx RG, Dahm DL, Sennett BJ, Stuart MJ, et al. Does arthroscopic knee surgery work? Arthroscopy 2014;30:544–5.
10. Niu NN, Losina E, Martin SD, Wright J, Solomon DH, Katz JN. Development and preliminary validation of a meniscal symptom index. Arthritis Care Res (Hoboken) 2011;63:208–15.

11. Katz JN, Chaisson CE, Cole B, Guermazi A, Hunter DJ, Jones M, et al. The MeTeOR trial (Meniscal Tear in Osteoarthritis Research): rationale and design features. Contemp Clin Trials 2012;33:1189–96.
12. Ball J, Jeffrey MR, Kellgren JH, editors. The epidemiology of chronic rheumatism: atlas of standard radiographs of arthritis. Vol. II. Oxford: Blackwell Scientific Publications; 1963.
13. 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.
14. Knee Injury and Osteoarthritis Outcome Score. KOOS FAQs. URL: http://www.koos.nu/koosfaq.html.
15. Roos EM, Lohmander LS. The Knee Injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes 2003;1:64.
16. Sihvonen R, Englund M, Turkiewicz A, Jarvinen TL. Mechanical symptoms and arthroscopic partial meniscectomy in patients with degenerative meniscus tear: a secondary analysis of a randomized trial. Ann Intern Med 2016;164:449–55.
17. Sihvonen R, Paavola M, Malmivaara A, Itälä A, Joukainen A, Nurmi H, et al. Arthroscopic partial meniscectomy versus placebo surgery for a degenerative meniscus tear: a 2-year follow-up of the randomised controlled trial. Ann Rheum Dis 2018;77:188–95.
18. Gauffin H, Sonesson S, Meunier A, Magnusson H, Kvist J. Knee arthroscopic surgery in middle-aged patients with meniscal symptoms: a 3-year follow-up of a prospective, randomized study. Am J Sports Med 2017;45:2077–84.
19. Gauffin H, Tagesson S, Meunier A, Magnusson H, Kvist J. Knee arthroscopic surgery is beneficial to middle-aged patients with meniscal symptoms: a prospective, randomised, single-blinded study. Osteoarthritis Cartilage 2014;22:1808–16.
20. Sonesson S, Kvist J, Yakob J, Hedevik H, Gauffin H. Knee arthroscopic surgery in middle-aged patients with meniscal symptoms: a 5-year follow-up of a prospective, randomized study. Orthop J Sports Med 2020;8:2325967119893920.
21. Bollen S. An open letter to the editor of the BMJ. Arthroscopy 2018;34:8–11.
22. Thorlund JB, Juhl CB, Roos EM, Lohmander LS. Arthroscopic surgery for degenerative knee: systematic review and meta-analysis of benefits and harms. BMJ 2015;350:h2747.
23. Thorlund JB, Englund M, Christensen R, Nissen N, Pihl K, Jørgensen U, et al. Patient reported outcomes in patients undergoing arthroscopic partial meniscectomy for traumatic or degenerative meniscal tears: comparative prospective cohort study. BMJ 2017;356:j356.
24. MacFarlane LA, Yang H, Collins JE, Guermazi A, Jones MH, Spindler KP, et al. Influence of baseline magnetic resonance imaging features on outcome of arthroscopic meniscectomy and physical therapy treatment of meniscal tears in osteoarthritis. Am J Sports Med 2019;47:612–9.
25. Kamimura M, Umehara J, Takahashi A, Mori Y, Chiba D, Kuwahara Y, et al. Meniscal tear morphology independently affects pain relief following arthroscopic partial meniscectomy in middle-aged patients. Knee Surg Sports Traumatol Arthrosc 2019;27:2460–7.
26. Herrlin SV, Wange PO, Lapidus G, Hallander M, Werner S, Weidenhielm L. Is arthroscopic surgery beneficial in treating non-traumatic, degenerative medial meniscal tears? A five year follow-up. Knee Surg Sports Traumatol Arthrosc 2013;21:358–64.

Arthritis Care & Research
Vol. 74, No. 8, August 2022, pp 1391–1398
DOI 10.1002/acr.24568

American College of Rheumatology
Empowering Rheumatology Professionals

Presence of Magnetic Resonance Imaging–Defined Inflammation Particularly in Overweight and Obese Women Increases Risk of Radiographic Knee Osteoarthritis: The POMA Study

Frank W. Roemer,[1] Ali Guermazi,[2] Michael J. Hannon,[3] Tomoko Fujii,[3] Patrick Omoumi,[4] David J. Hunter,[5] Felix Eckstein,[6] and C. Kent Kwoh[3]

Objective. The present study was undertaken to assess whether the odds for incident radiographic osteoarthritis (OA) differ between men and women in regard to body mass index (BMI) and inflammatory magnetic resonance imaging (MRI) markers 1 and 2 years prior, and whether the presence of inflammation on MRI differs between normal-weight and overweight/obese individuals who develop radiographic OA up to 4 years prior.

Methods. We studied 355 knees from the Osteoarthritis Initiative study that developed incident radiographic OA and 355 matched controls. MRIs were read for effusion-synovitis and Hoffa-synovitis for up to 4 consecutive annual time points. Subjects were classified as normal-weight (BMI <25), overweight (BMI ≥25 and <30), or obese (BMI ≥30). Conditional logistic regression was used to assess odds of incident radiographic OA for effusion-synovitis and Hoffa-synovitis at 1 and 2 years prior to radiographic OA incidence (i.e., "P-1" and "P-2") considering BMI category. Bivariate logistic regression was used to assess odds of inflammation for cases only.

Results. One hundred seventy-eight (25.1%) participants were normal weight, 283 (39.9%) overweight, and 249 (35.1%) obese. At P-2, being overweight with Hoffa-synovitis, which had an odds ratio [OR] of 3.26 (95% confidence interval [95% CI] 1.39–7.65), or effusion-synovitis (OR 3.56 [95% CI 1.45–8.75]) was associated with greater odds of incident radiographic OA in women. For those with incident radiographic OA, there were no increased odds of synovitis in the overweight/obese subgroup for most time points, but increased odds for effusion-synovitis were observed at P-2 (OR 2.21 [95% CI 1.11–4.43]).

Conclusion. Presence of inflammatory markers seems to play a role especially in overweight women, while obese women have increased odds for radiographic OA also in the absence of these markers.

INTRODUCTION

Obesity is one of the key risk factors for the development of knee osteoarthritis (OA) (1). Associations linking OA development to components of the so-called metabolic syndrome beyond obesity have been suggested. These include chronic low-grade inflammation, a feature shared by OA and metabolic disorders that may contribute to the genesis of both (2,3). While studies

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

The image analyses were funded by an NIH contract with the University of Pittsburgh (National Heart, Lung, and Blood Institute grant HHSN268201000021C; Pivotal Osteoarthritis Initiative Magnetic Resonance Imaging Analyses [POMA] contract) and in part by a vendor contract from the Osteoarthritis Initiative coordinating center at University of California, San Francisco (grant N01-AR-2-2258). The statistical data analysis was funded by an NIH contract with the University of Pittsburgh (National Heart, Lung, and Blood Institute grant HHSN268201000021C; POMA contract).

[1]Frank W. Roemer, MD: Boston University School of Medicine, Boston, Massachusetts, and Friedrich-Alexander University Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany; [2]Ali Guermazi, MD, PhD: Boston University School of Medicine, Boston, Massachusetts, and VA Boston Healthcare System, West Roxbury, Massachusetts; [3]Michael J. Hannon, MA (current address: Pinney Associates, Pittsburgh, Pennsylvania), Tomoko Fujii, MS, MPH (current address: The University of Tokyo, Tokyo, Japan), C. Kent Kwoh, MD (current address: University of Arizona College of Medicine, Tucson): University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; [4]Patrick Omoumi, MS, PhD: Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland; [5]David J. Hunter, MBBS, PhD: Royal North Shore Hospital and Institute of Bone and University of Sydney, Sydney, New South Wales, Australia; [6]Felix Eckstein, MD: Paracelsus Medical University Salzburg and Nuremberg, Salzburg, Austria, and Chondrometrics, Ainring, Germany.

SIGNIFICANCE & INNOVATIONS

- In women, being overweight with Hoffa-synovitis and being overweight or obese with effusion-synovitis increases odds for incident radiographic osteoarthritis (OA) 2 years later.
- Presence of effusion-synovitis increases odds for incident radiographic OA in overweight and obese women but not in men.
- For individuals who develop incident radiographic OA, increased odds for effusion-synovitis were observed 2 years prior (odds ratio 2.21 [95% confidence interval 1.11–4.43]).
- Both mechanical load and inflammation seem to have a role in OA incidence for overweight and obese women, while for men the role of inflammation in conjunction with high body mass index seems to be less relevant.

have reported that the metabolic syndrome is clearly associated with increased risk of knee OA (4), a recent meta-analysis suggested that this may only be indirect, and that there was insufficient evidence that the metabolic syndrome was associated with incident knee OA independent of body mass index (BMI) (5).

Beyond proinflammatory systemic factors, local intraarticular adipose tissues such as Hoffa's fat pad produce inflammatory and catabolic mediators that may contribute to OA pathogenesis (6). Further, it is unclear whether women and men show differences regarding the presence of metabolic syndrome and incident knee OA. While one study did not report any sex-specific differences (3), others have highlighted that inflammation and metabolic syndrome may have a larger impact on OA incidence in women compared to men (7). We hypothesize that individuals with high body mass index (BMI) and local inflammation as assessed by magnetic resonance imaging (MRI), considered as surrogates for some of the components of the metabolic syndrome (8), may be at increased odds for incident knee OA, and that overweight and obese individuals are at increased odds for exhibiting signs of local joint inflammation as assessed by MRI up to 4 years prior to the incidence of radiographic OA.

The aims of this study were as follows: to assess whether odds for incident radiographic OA differ between men and women in regard to BMI and inflammatory MRI markers 1 and 2 years prior to radiographic OA incidence using a matched case–control sample of subjects who developed or did not develop incident radiographic OA; and to analyze whether odds of presence of MRI features of inflammation such as effusion-synovitis (effusion) and Hoffa-synovitis (synovitis) differ between normal-weight, and overweight/obese individuals who develop incident radiographic OA over a period of up to 4 years prior.

MATERIALS AND METHODS

The Osteoarthritis Initiative (OAI). The OAI is a longitudinal cohort study designed to identify biomarkers of the onset and/or progression of knee OA. Both knees of 4,796 participants were studied using 3T MRI and fixed-flexion radiography at baseline, 12, 24, 36, and 48 months of follow-up (9). The institutional review boards at each of the sites approved the study, and all participants gave informed consent.

Radiography. OAI knee radiographs were acquired using the posteroanterior fixed-flexion weight-bearing protocol using a positioning frame. Kellgren/Lawrence (K/L) grade was determined by central readings of baseline serial fixed-flexion knee radiographs (10).

Case and control knee selection. Cases were defined as study participants who had at least 1 knee that developed incident radiographic OA during the 4 years of follow-up. Incident radiographic OA was defined as the first occurrence of radiographic findings compatible with OA (K/L grade of ≥2 on the posteroanterior view based on central readings) during the course of study. This time point was called P0, with P-1 being defined as the time point 1 year before radiographic OA was detected, P-2 defined as 2 years prior, P-3 three years prior, and P-4 four years prior to when incident radiographic OA was read. All participants fulfilling the case definition were included. An identical number of control knees were selected from knees that did not develop incident radiographic OA during the study period. The controls were matched to case knees according to K/L grade, sex, age (within 5 years), and contralateral knee OA status (i.e., K/L grade = 0, 1, or 2+ in the other knee). Each case was matched to those who were at risk at the time of case occurrence and those with available images at relevant time points, whether this was at 12, 24, 36, or 48 months of follow-up. Both cases and control

Dr. Roemer owns stock or stock options in Boston Imaging Core Lab. Dr. Guermazi has received consulting fees, speaking fees, and/or honoraria from AstraZeneca, Galapagos, Roche (less than $10,000 each), Pfizer, Merck Serono, and TissuGene (more than $10,000 each), and owns stock or stock options in Boston Imaging Core Lab. Mr. Hannon has received consulting fees from EMD Serono (less than $10,000). Dr. Hunter has received consulting fees, speaking fees, and/or honoraria from Pfizer, Merck Serono, TissueGene, and TLC (less than $10,000 each). Dr. Eckstein has received consulting fees, speaking fees, and/or honoraria from MerckSerono, Synarc, Servier, and Medtronic (less than $10,000 each) and research support from Pfizer, Eli Lilly and Company, Stryker, Novartis, MerckSerono, GlaxoSmithKline, Wyeth, Centocor, AbbVie, Kolon, Synarc, Ampio, and Orthotrophix. Dr. Kwoh has received consulting fees, speaking fees, and/or honoraria from Astellas, Fidia, GlaxoSmithKline, Kolon TissueGene, Regeneron, Regulus, Taiwan Liposome Company, Thusane (less than 10,000 each), EMD Serono, and Express Scripts (more than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Frank W. Roemer, MD, Adjunct Professor of Radiology, Quantitative Imaging Center, Department of Radiology, Boston University School of Medicine, FGH Building, 3rd floor, 820 Harrison Avenue, Boston, MA 02118. Email: froemer@bu.edu.

Submitted for publication June 16, 2020; accepted in revised form January 28, 2021.

knees were either K/L grade 0 or 1 at baseline based on central readings. Only 1 knee per subject was used as a case knee. A flow chart of the inclusion of cases and controls is included as Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24568.

MRI acquisition and assessment. MRI of both knees was performed on identical 3T systems (Siemens Trio) at the 4 OAI clinical sites. The OAI pulse sequence protocol and the sequence parameters have been published in detail (9).

Two musculoskeletal radiologists with 11 (FWR) and 14 (AG) years' experience of semiquantitative assessment of knee OA at the time of reading, blinded to clinical data and case–control status, read the MRIs according to the MRI OA Knee Score system (11). Baseline and follow-up MRIs were read with the chronological order known to the readers. Diffuse hyperintense signal on the sagittal intermediate-weighted fat-suppressed sequence in the intercondylar region of Hoffa's fat pad were scored from 0 to 3 as a surrogate for synovial thickening, termed Hoffa-synovitis (i.e., synovitis). The degree of hyperintensity was assessed according to the following grades: 0 = normal, 1 = mild, 2 = moderate, and 3 = severe. Joint effusion (also called effusion-synovitis, as it is not possible to discern joint fluid from synovial thickening on non–contrast-enhanced MRI) was graded from 0 to 3 in terms of the estimated maximal distention of the synovial cavity (i.e., effusion) as follows: grade 0 = none, grade 1 = small, grade 2 = medium, and grade 3 = large (11,12). Examples of the different grades of Hoffa-synovitis and effusion-synovitis are presented in Figure 1. Detailed reliability data of MRI assessment are presented in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24568.

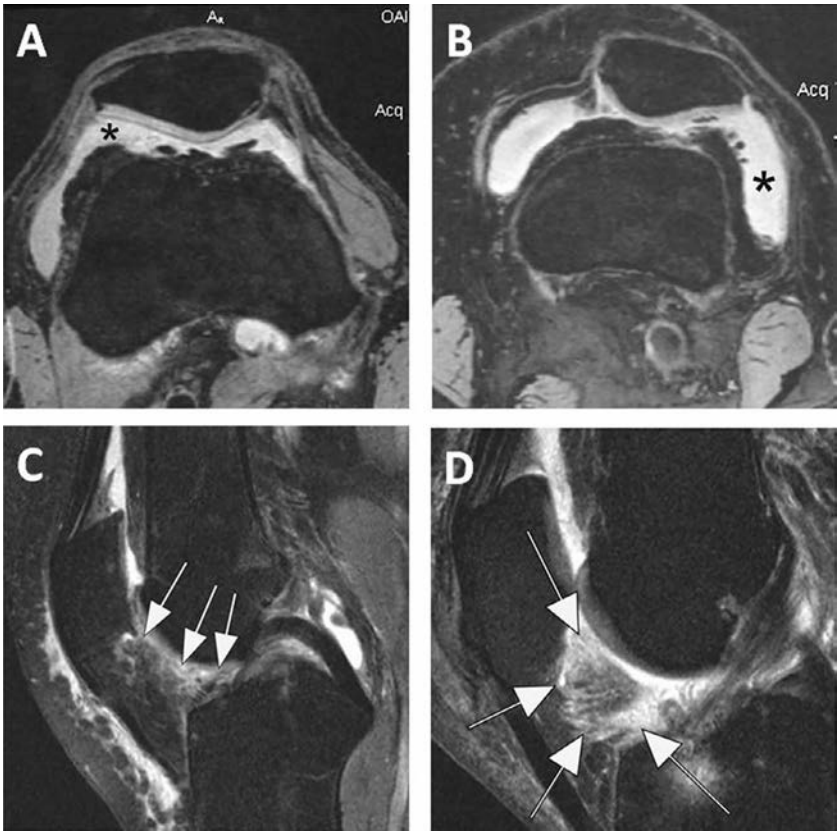


Figure 1. Magnetic resonance imaging (MRI) markers of inflammation in osteoarthritis (OA). Fluid-sensitive sequences are capable of delineating intraarticular joint fluid. However, a distinction between true joint effusion and synovial thickening is not possible, as both are visualized as a hyperintense signal within the joint cavity. For this reason, the term effusion-synovitis was introduced, which in the MRI OA Knee Score system is scored based on the distension of the joint capsule and is graded from 0 to 3 in terms of the estimated maximal distention of the synovial cavity, with 0 = normal, grade 1 = <33% of maximum potential distention, grade 2 = 33–66% of maximum potential distention, and grade 3 = >66% of maximum potential distention. Axial dual-echo steady-state MRI shows grade 2 effusion-synovitis (**asterisk**) (**A**) and grade 3 effusion-synovitis (**asterisk**) (**B**). In addition, signal changes in Hoffa's fat pad are commonly used as a surrogate for synovitis on non–contrast-enhanced MRI. Although synovitis can only be visualized directly on contrast-enhanced sequences, it has been shown that Hoffa's signal changes are a sensitive but nonspecific surrogate of synovitis (**C**). Sagittal intermediate-weighted fat-suppressed MRI shows a discrete ill-defined hyperintense signal alteration in Hoffa's fat pad consistent with grade 2 Hoffa-synovitis (**arrows**) (**D**). Severe, grade 3 signal alterations almost occupying the entire fat pad are seen in this image (**arrows**).

Table 1. Demographic characteristics of the sample*

	Cases (n = 355)	Controls (n = 355)	*P*
Age, mean ± SD years	60.1 ± 8.6	60.0 ± 8.4	NA
BMI, mean ± SD kg/m^2	28.9 ± 4.5	27.7 ± 4.4	0.0003
WOMAC knee pain score, mean ± SD†	2.6 ± 3.3	1.4 ± 2.5	<0.0001
WOMAC functioning score, mean ± SD†	8.4 ± 10.8	4.3 ± 7.8	<0.0001
Sex			NA
Female	237 (66.8)	237 (66.8)	
Male	118 (33.2)	118 (33.2)	
BMI, kg/m^2			0.0032
Normal/underweight	70 (19.7)	108 (30.4)	
Overweight	147 (41.4)	136 (38.3)	
Obese	138 (38.9)	111 (31.3)	
Race			0.2143
White	283 (79.7)	299 (84.2)	
African American	61 (17.2)	47 (13.2)	
Asian	6 (1.7)	2 (0.6)	
Other	5 (1.4)	7 (2)	
K/L grade			NA
0	133 (37.5)	133 (37.5)	
1	222 (62.5)	222 (62.5)	
Knee injury at OAI baseline‡	136 (38.3)	70 (19.7)	<0.0001
Knee surgery at OAI baseline§	54 (15.2)	24 (6.8)	0.0004

* Values are the number (%) unless indicated otherwise. *P* values for differences by Fisher's exact test for categorical variables and *t*-tests for ordinal variables were not calculated for variables used in matching. BMI = body mass index; K/L = Kellgren/Lawrence; NA = not applicable; OAI = Osteoarthritis Initiative; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† WOMAC knee pain is on a scale of 1 to 20, and WOMAC functioning of 1 to 96, with higher values representing more pain/less functioning.
‡ Knee injury defined as 1 inhibiting ability to walk for at least 2 days.
§ Knee surgery includes arthroscopy.

Statistical analysis. Subjects were classified as normal weight (BMI <25 kg/m^2), overweight (BMI ≥25 kg/m^2 and <30 kg/m^2), or obese (BMI ≥30 kg/m^2) at OAI enrollment. In the case–control design part of the study, conditional logistic regression was used to assess the risk of incident radiographic OA stratified by presence of synovitis and effusion focusing on the time points P-1 and P-2 only. Presence of synovitis and effusion was defined as "any," i.e., knees that exhibited grades 1–3 of synovitis or effusion on MRI. The time points P-3 and P-4 were not considered, as low numbers did not allow meaningful interpretation of the interactions (for P-3 only, 59 cases, and for P-4 only, 53 cases were available). For the case–control analysis, stratification by sex was undertaken, and BMI, synovitis, and effusion or the interaction were used as exposure variables. First, the bivariate associations of radiographic OA and the different synovitis and effusion categories and BMI were estimated. After this initial analysis, the risk of radiographic OA for the interaction of BMI and effusion/synovitis was examined. The category of normal weight, especially in men, was sufficiently uncommon that we used the overweight category as the referent for the BMI analysis because it was the norm.

Bivariate logistic regression was used to assess the odds of the presence of synovitis and effusion at time points P-1, P-2, P-3, P-4, and baseline in subjects who developed radiographic OA (i.e., only cases), comparing overweight and obese subjects combined to subjects of normal weight as the reference. We considered a 2-tailed *P* value of less than 0.05 as statistically significant. All statistical calculations were performed using Stata/IC, version 11.2, for Windows and SAS, version 9.3.

RESULTS

A total of 355 case knees and 355 matched control knees were included. Participants had a mean ± SD age of 60.2 ± 8.6 years; 66.8% were female. Cases had a slightly higher BMI compared to controls (28.9 kg/m^2 versus 27.7 kg/m^2; *P* = 0.0003). No significant differences with regard to ethnicity between cases and controls were observed (84% of the subjects were White). The case-defining visit of radiographic OA incidence was 12 months for 119 knees (33.5%), 24 months for 83 knees (23.4%), 36 months for 103 knees (29.0%), and 48 months for 50 knees (14.1%). In total, 178 (25.1% of all study participants; n = 138 [77.5%] women) participants were normal weight, 283 (39.9% of all study participants; n = 166 [58.7%] women) were overweight, and 249 (35.1% of all study participants; n = 170 [68.3%] women) were obese at baseline. Details of demographic characteristics regarding cases and controls are presented in Table 1.

Regarding the interaction of BMI with synovitis and effusion, using overweight women and men without synovitis or effusion

as the reference, obesity without synovitis was associated with greater odds of radiographic OA in women at P-2, with an odds ratio (OR) of 2.87 (95% confidence interval [95% CI] 1.21–6.83), as was being overweight with synovitis (OR 3.26 [95% CI 1.39–7.65]). Being obese with synovitis was not associated with increased odds at P-2. For men, there were no combinations of synovitis and BMI that were associated with increased odds of radiographic OA compared to those being overweight without synovitis at P-2. Furthermore, being overweight with joint effusion at P-2 was associated with increased OA odds in women (OR 3.56 [95% CI 1.45–8.75]), an association also observed in women who were obese (OR 3.46 [95% CI 1.38–8.72]).

At P-1 and combining all BMI categories, having any synovitis or any effusion was associated with increased odds of radiographic OA in both men and women. Further, presence of synovitis was associated with incident radiographic OA in overweight and obese women and men, with the latter association also seen for men of normal weight, which was not the case for women of normal weight. Positive associations of effusion with incident OA were only seen in overweight (OR 3.14 [95% CI 1.55–6.36]) and obese women (OR 3.03 [95% CI 1.50–6.15]) but not women of normal weight or in men. Table 2 gives a detailed overview of these results regarding the interactions between BMI, sex, and severity of inflammation at P-2 and P-1.

For those knees that developed radiographic OA, there were no increased odds of synovitis in the combined overweight/obese (i.e., categories combined) BMI subgroup compared to the normal-weight subgroup at any of the 4 time

Table 2. Odds for developing radiographic osteoarthritis (OA) at Osteoarthritis Initiative (OAI) visits year 2 (P-2) or year 1 (P-1) prior to the case-defining visit in matched cases and controls*

	P-2					P-1				
	All, no. (%)	Men, no. (%)	Men, OR (95% CI) (n = 136)	Women, no. (%)	Women, OR (95% CI) (n = 300)	All, no. (%)	Men, no. (%)	Men, OR (95% CI) (n = 224)	Women, no. (%)	Women, OR (95% CI) (n = 436)
Normal weight	116 (26.6)	21 (15.4)	1.32 (0.42–4.15)	95 (31.7)	0.57 (0.31–1.04)	166 (25.2)	38 (17.0)	0.96 (0.44–2.10)	128 (29.4)	0.52 (0.31–0.87)
Overweight	169 (38.8)	63 (46.3)	Ref.	106 (35.3)	Ref.	264 (40.0)	114 (50.9)	Ref.	150 (34.4)	Ref.
Obese	151 (34.6)	52 (38.24)	0.89 (0.44–1.78)	99 (33.0)	1.3 (0.77–2.43)	230 (34.8)	72 (32.1)	0.92 (0.52–1.64)	158 (36.2)	1.44 (0.89–2.35)
No synovitis	221 (50.7)	59 (43.4)	Ref.	162 (54.0)	Ref.	337 (51.1)	99 (44.4)	Ref.	238 (54.6)	Ref.
Synovitis	215 (49.3)	77 (56.6)	1.79 (0.91–3.50)	138 (46.0)	1.75 (1.05–2.91)†	322 (48.9)	124 (55.6)	3.62 (1.94–6.74)†	198 (45.4)	1.97 (1.29–3.01)†
No effusion	234 (53.7)	72 (52.9)	Ref.	162 (54.0)	Ref.	338 (51.2)	107 (47.8)	Ref.	231 (53.0)	Ref.
Effusion	202 (46.3)	64 (47.1)	0.75 (0.36–1.59)	138 (46.0)	2.88 (1.64–5.03)	322 (48.8)	117 (52.2)	1.88 (1.05–3.37)†	205 (47.0)	2.89 (1.87–4.47)†
No synovitis, BMI normal	66 (15.1)	11 (8.1)	1.45 (0.31–6.77)	55 (18.3)	0.67 (0.29–1.59)	91 (13.8)	18 (8.1)	1.57 (0.42–5.81)	73 (16.7)	0.75 (0.35 1.59)
No synovitis, BMI overweight	84 (19.3)	25 (18.4)	Ref.	59 (19.7)	Ref.	130 (19.7)	47 (21.1)	Ref.	83 (19.0)	Ref.
No synovitis, BMI obese	71 (16.3)	23 (16.9)	1.03 (0.31–3.42)	48 (16.0)	2.87 (1.21–6.83)†	116 (17.6)	34 (15.2)	1.63 (0.58–4.57)	82 (18.8)	2.30 (1.17–4.56)†
Synovitis, BMI normal	50 (11.5)	10 (7.4)	3.25 (0.69–15.29)	40 (13.3)	1.52 (0.61–3.77)	74 (11.2)	19 (8.5)	4.10 (1.46–11.55)†	55 (12.6)	1.23 (0.57–2.64)
Synovitis, BMI overweight	85 (19.5)	38 (27.9)	1.99 (0.66–6.00)	47 (15.7)	3.26 (1.39–7.65)†	134 (20.3)	67 (30.0)	5.69 (2.06–15.67)†	67 (15.4)	3.66 (1.74–7.69)†
Synovitis, BMI obese	80 (18.3)	29 (21.3)	1.63 (0.58–4.60)	51 (17.0)	1.86 (0.80–4.34)	114 (17.3)	38 (17.0)	3.72 (1.39–9.98)†	76 (17.4)	2.71 (1.24–5.92)†
No effusion, BMI normal	69 (15.8)	15 (11.0)	3.86 (0.83 18.05)	54 (18.0)	0.68 (0.28–1.63)	105 (15.9)	26 (11.6)	1.00 (0.35–2.86)	79 (18.1)	0.50 (0.23–1.09)
No effusion, BMI overweight	92 (21.1)	31 (22.8)	Ref.	61 (20.3)	Ref.	125 (18.9)	46 (20.5)	Ref.	79 (18.1)	Ref.
No effusion, BMI obese	73 (16.7)	26 (19.1)	1.41 (0.46 4.33)	47 (15.7)	1.16 (0.50–2.67)	108 (16.4)	35 (15.6)	0.61 (0.25–1.51)	73 (16.7)	1.71 (0.85–3.47)
Effusion, BMI normal	47 (10.8)	6 (4.4)	0.00 (0.00–0.00)‡	41 (13.7)	1.30 (0.51–3.30)	61 (9.2)	12 (5.4)	1.75 (0.46–6.58)	49 (11.2)	1.78 (0.81–3.89)
Effusion, BMI overweight	77 (17.7)	32 (23.5)	1.64 (0.53–5.12)	45 (15.0)	3.56 (1.45–8.75)†	139 (21.1)	68 (30.4)	1.41 (0.60–3.30)	71 (16.3)	3.14 (1.55–6.36)†
Effusion, BMI obese	78 (17.9)	26 (19.1)	1.10 (0.36–3.37)	52 (17.3)	3.46 (1.38–8.72)†	122 (18.5)	37 (16.5)	1.94 (0.77–4.86)	85 (19.5)	3.03 (1.50–6.15)†

* 95% CI = 95% confidence interval; BMI = body mass index; OR = odds ratio; P-1 = OAI visit 1 year prior to the case-defining visit when incident radiographic OA was diagnosed/read; P-2 = OAI visit 2 years prior to the case-defining visit when incident radiographic OA was diagnosed/read; Ref. = reference.
† Statistically significant at $P < 0.05$.
‡ No case knees in this category.

points prior to the case visit or the baseline visit. However, being overweight/obese was associated with an increased odds of effusion at P-2 (OR 2.21 [95% CI 1.11–4.43]). Though not statistically significant, increased odds for effusion were also observed for the visit P-1 (OR 1.68 [95% CI 0.98–2.88]). Table 3 presents details for the case knees and associated odds for synovitis or effusion at several time points prior to the incidence of radiographic OA.

DISCUSSION

The presence of synovitis increased the odds of developing radiographic OA in overweight women at the time point 2 years before radiographic OA was detected, while obese women had an increased risk for radiographic OA also without synovitis. At the time point of 1 year prior to OA incidence, we observed increased odds for incident radiographic OA in overweight and obese women with presence of joint effusion, but not in men. At the same time point, increased odds for radiographic OA incidence were seen in both overweight and obese women and men in the presence of synovitis, but not for normal-weight women with synovitis, suggesting that the presence of effusion seems to play a role particularly in overweight or obese women. In knees that developed radiographic OA, increased odds of effusion were observed for the combined overweight/obese group at P-2 but not for Hoffa-synovitis or any of the other time points, suggesting a possible link between high BMI, presence of joint effusion, and radiographic OA development 2 years later.

While the role of body weight and knee radiographic OA incidence is well established, its interactions with local

Table 3. Risk for Hoffa-synovitis and effusion-synovitis in case knees that developed incident radiographic osteoarthritis (OA) based on baseline body mass index (BMI) status for different Osteoarthritis Initiative (OAI) visits, with normal weight participants as the reference (Ref.)*

OAI visit prior to incident radiographic OA and BMI status	No synovitis/ effusion	Yes synovitis/ effusion	OR (95% CI)	*P*
Synovitis (any)				
P-1				
Normal (n = 63)	28	35	Ref.	
Overweight (n = 266)	106	160	1.21 (0.68–2.15)	0.52
P-2				
Normal (n = 48)	22	26	Ref.	
Overweight (n = 177)	78	99	1.07 (0.54–2.13)	0.84
P-3				
Normal (n = 38)	21	17	Ref.	
Overweight (n = 112)	53	59	1.38 (0.63–3.00)	0.42
P-4				
Normal (n = 13)	6	7	Ref.	
Overweight (n = 37)	18	19	0.90 (0.23–3.51)	0.89
Baseline				
Normal (n = 69)	35	34	Ref.	
Overweight (n = 285)	128	157	1.26 (0.73–2.19)	0.41
Effusion (any)				
P-1				
Normal (n = 64)	33	31	Ref.	
Overweight (n = 266)	103	163	1.68 (0.98–2.88)	0.06
P-2				
Normal (n = 48)	30	18	Ref.	
Overweight (n = 177)	76	101	2.21 (1.11–4.43)†	0.02
P-3				
Normal (n = 38)	25	13	Ref.	
Overweight (n = 112)	55	57	1.99 (0.91–4.38)	0.09
P-4				
Normal (n = 13)	7	6	Ref.	
Overweight (n = 37)	17	20	1.37 (0.36–5.22)	0.64
Baseline				
Normal (n = 70)	42	28	Ref.	
Overweight (n = 285)	139	146	1.58 (0.93–2.68)	0.09

* Values are the number unless indicated otherwise. Ref. is normal weight (BMI <25 kg/m^2). Overweight and obese subgroups are combined. 95% CI = 95% confidence interval; OR = odds ratio; P-1 = OAI visit 1 year prior to the case-defining visit (when radiographic OA incidence was diagnosed/read); P-2 = OAI visit 2 years prior to the case-defining visit (when radiographic OA incidence was diagnosed/read); P-3 = OAI visit 3 years prior to the case-defining visit (when radiographic OA incidence was diagnosed/read); P-4 = OAI visit 4 years prior to the case-defining visit (when radiographic OA incidence was diagnosed/read).
† Statistically significant at $P < 0.05$.

inflammation have been less clear (1). Reported associations between obesity and OA development also for non–weight-bearing joints suggest a more complex interaction beyond increased biomechanical loading. In a population-based cohort study, it has been reported that metabolic syndrome may be prevalent in 59% of patients with knee OA and in 23% of those without (13). Conversely, Niu et al found in a population-based study that among women, abdominal obesity and high blood pressure were associated with incident radiographic OA, but metabolic syndrome was not (3). We have shown previously a strong association between the presence of joint inflammatory markers based on MRI and subsequent radiographic OA incidence, and this current work expands this, taking also into account sex and BMI differences (14). The fact that 2 years prior to radiographic OA incidence, obesity in women without synovitis exhibited increased risk for radiographic OA, as did being overweight with synovitis, but that obesity with synovitis did not, was not an expected finding. We can only speculate that potentially in obese individuals other factors, including direct results of increased loading due to higher BMI resulting in structural changes like bone marrow alterations, cartilage damage, or meniscal lesions and extrusion, may be more relevant than inflammatory manifestations such as effusion or synovitis.

Concerning the second part of our analysis focusing on cases only regarding prevalence of inflammatory markers in the different subgroups, we found that up to 4 years prior to radiographic OA incidence, in general, the combined overweight/obese subgroup did not show significantly increased rates of local inflammation, with the exception of effusion 2 years prior to radiographic OA incidence, while at 1 year prior, the association was close to being significant. A recent study also from the OAI reported a significantly greater prevalence and severity of synovial inflammation imaging biomarkers in knees of overweight and obese participants compared to those that have normal weight (15). In contrast to our study, however, almost 20% of included subjects exhibited radiographic OA grades 2 and 3, and for those without radiographic OA, it is not known how many developed radiographic OA at later time points. Thus, we speculate based on our findings that for case knees only (i.e., for those that developed radiographic OA), other factors beyond obesity, including local structural damage such as meniscal or cartilage lesions, may have additional impact on the presence of synovial inflammation, thus diluting possible impact of increased BMI.

We acknowledge that in this exploratory study we did not analyze subjects with defined metabolic syndrome, as we only analyzed interactions of BMI and MRI markers of inflammation, which limits extrapolation of our findings to patients with metabolic syndrome (3). An additional limitation of our study includes the absence of information on symptomatic OA. We do not know if subjects who developed radiographic OA also developed symptoms, and if subjects developed symptoms prior to the diagnosis of radiographic OA. Furthermore, the OAI study does not include contrast-enhanced MRI sequences, the gold standard for synovitis assessment (16). However, we used an established surrogate for whole-joint synovitis that has been used in multiple studies applying MRI (11). Inter- and intrareader agreement was almost perfect for effusion grading, but only substantial for synovitis assessment, which is a limitation and likely reflects the nonspecificity of non–contrast-enhanced MRI (17).

In conclusion, the presence of MRI-defined Hoffa-synovitis seems to play a role for incident radiographic OA development, especially in overweight women, whereas obese women have increased odds for radiographic OA even in the absence of Hoffa-synovitis. Presence of joint effusion has an impact on radiographic OA development particularly in overweight and obese women but not men. Being overweight/obese increased odds for joint effusion in the knees that developed incident radiographic OA at time points 1 and 2 years prior. These results suggest that both mechanical load and inflammation have a role in OA incidence for overweight and obese women, while for men, the role of inflammation in conjunction with high BMI appears to be less relevant.

ACKNOWLEDGMENTS

The authors thank the OAI participants, OAI investigators, OAI clinical and technical staff, the OAI coordinating center, and the OAI funders for providing this unique public data base.

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. Roemer 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. Roemer, Guermazi, Hannon, Fujii, Hunter, Eckstein, Kwoh.

Acquisition of data. Roemer, Guermazi, Fujii, Kwoh.

Analysis and interpretation of data. Roemer, Guermazi, Hannon, Fujii, Omoumi, Hunter, Eckstein, Kwoh.

ADDITIONAL DISCLOSURES

Author Hannon is currently an employee of Pinney Associates but was employed by the University of Pittsburgh during the time the study was conducted. Author Eckstein is an employee of Chondrometrics.

REFERENCES

1. Niu J, Zhang YQ, Torner J, Nevitt M, Lewis CE, Aliabadi P, et al. Is obesity a risk factor for progressive radiographic knee osteoarthritis? Arthritis Rheum 2009;61:329–35.
2. Griffin TM, Huebner JL, Kraus VB, Yan Z, Guilak F. Induction of osteoarthritis and metabolic inflammation by a very high-fat diet in mice: effects of short-term exercise. Arthritis Rheum 2012;64:443–53.
3. Niu J, Clancy M, Aliabadi P, Vasan R, Felson DT. Metabolic syndrome, its components, and knee osteoarthritis: the Framingham Osteoarthritis Study. Arthritis Rheumatol 2017;69:1194–203.

4. 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.
5. Wang H, Cheng Y, Shao D, Chen J, Sang Y, Gui T, et al. Metabolic syndrome increases the risk for knee osteoarthritis: a meta-analysis. Evid Based Complement Alternat Med 2016;2016: 7242478.
6. Clockaerts S, Bastiaansen-Jenniskens YM, Runhaar J, van Osch GJ, van Offel JF, Verhaar JA, et al. The infrapatellar fat pad should be considered as an active osteoarthritic joint tissue: a narrative review. Osteoarthritis Cartilage 2010;18:876–82.
7. Lee BJ, Yang S, Kwon S, Choi KH, Kim W. Association between metabolic syndrome and knee osteoarthritis: a cross-sectional nationwide survey study. J Rehabil Med 2019;51:464–70.
8. Alberti KG, Zimmet P, Shaw J. Metabolic syndrome: a new world-wide definition. A consensus statement from the International Diabetes Federation. Diabet Med 2006;23:469–80.
9. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis Cartilage 2008;16:1433–41.
10. Osteoarthritis Initiative. Appendix A. Project 15: reader discrepancies and adjudication procedures. OAI central reading of knee x-rays for K-L grade and individual features of knee OA. 2013; p. 8. URL: https://nda.nih.gov/binaries/content/documents/ndacms/resources/oai/oai-image-assessment-overview/oai-image-assessment-overview/ndacms:resource.
11. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage 2011;19:990–1002.
12. Hill CL, Hunter DJ, Niu J, Clancy M, Guermazi A, Genant H, et al. Synovitis detected on magnetic resonance imaging and its relation to pain and cartilage loss in knee osteoarthritis. Ann Rheum Dis 2007;66:1599–603.
13. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Association of knee osteoarthritis with the accumulation of metabolic risk factors such as overweight, hypertension, dyslipidemia, and impaired glucose tolerance in Japanese men and women: the ROAD study. J Rheumatol 2011;38:921–30.
14. Roemer FW, Kwoh CK, Hannon MJ, Hunter DJ, Eckstein F, Fujii T, et al. What comes first? Multitissue involvement leading to radiographic osteoarthritis: magnetic resonance imaging–based trajectory analysis over four years in the Osteoarthritis Initiative. Arthritis Rheumatol 2015;67:2085–96.
15. Kanthawang T, Bodden J, Joseph GB, Lane NE, Nevitt M, McCulloch CE, et al. Obese and overweight individuals have greater knee synovial inflammation and associated structural and cartilage compositional degeneration: data from the osteoarthritis initiative. Skeletal Radiol 2021;50:217–29.
16. Guermazi A, Roemer FW, Hayashi D, Crema MD, Niu J, Zhang Y, et al. Assessment of synovitis with contrast-enhanced MRI using a whole-joint semiquantitative scoring system in people with, or at high risk of, knee osteoarthritis: the MOST study. Ann Rheum Dis 2011;70:805–11.
17. Roemer FW, Jarraya M, Felson DT, Hayashi D, Crema MD, Loeuille D, et al. Magnetic resonance imaging of Hoffa's fat pad and relevance for osteoarthritis research: a narrative review. Osteoarthritis Cartilage 2016;24:383–97.