Arthritis Care & Research

Aims and Scope

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

Volume 73, No. 6, June 2021

Arthritis Care & Research

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

Arthritis Care & Research

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

VOLUME 73 • June 2021 • NO. 6

Special Articles

Osteoarthritis

Systemic Lupus Erythematosus

Pediatrics

Sjögren's Syndrome

Cover image: The image on the cover (from Baer et al, page 849) shows ultrasound-guided core needle biopsy of the parotid gland during which (left) the biopsy needle is directed from posterior to anterior with the patient in (right) right side down decubitus position.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 765–766
DOI 10.1002/acr.24641

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

Arthritis Care & Research: A Look Back and a View Forward

Marian T. Hannan[1] and Leslie J. Crofford[2]

It is with great joy and a touch of sadness that we complete our term as editors of *Arthritis Care & Research* (AC&R), one of the premier journals in clinical rheumatology. With this issue, the editorial responsibilities will be successfully transferred from our team to Dr. Kelli Allen and colleagues. We are excited for the new team's journey and for the ongoing opportunities for readers and authors of work published in AC&R to engage in the pursuit of new knowledge that advances the care of patients with rheumatic diseases.

AC&R, the official journal of the Association of Rheumatology Professionals (ARP), has a mission covering clinical research relevant to the care of rheumatologic disorders. Important basic and translational sciences are covered by our sister journals, *Arthritis & Rheumatology* and *ACR Open Rheumatology,* along with other articles. AC&R serves a diverse base of investigators and practitioners from different disciplines who converge as one community in the interests of our patients. Articles in AC&R cover a spectrum of essential topics, including articles on treatments, epidemiology, methodologic issues, behavioral sciences, outcomes work, and health economics, to name a few. The interprofessional lens of AC&R provides the conduit for readership and authors to deliver the highest quality understanding and care for patients with rheumatic diseases (1). It is important to underscore the essential mission of AC&R — to address the clinical implementation needs of practicing rheumatologists and health professionals. Thus, our focus is not exclusively research-based but also includes a vital mission to discuss important current clinical issues and to provide a forum for publications focused on rheumatology education and workforce. The success of AC&R has relied upon multidisciplinary scientists, clinicians, and educators submitting their best rheumatologic-based work to our journal.

As we look back on our AC&R experience as editors, 3 key features come to the forefront: 1) the global impact of the journal; 2) the commitment of the authors and readers of AC&R; and 3) our heartfelt thanks for the many people who work tirelessly to bring the best to AC&R in all aspects. Let us provide a bit of detail.

Authors submitting important research to AC&R span the globe with representation from all continents (except Antarctica, thus far). AC&R is viewed as an important worldwide source for scholarship pertaining to clinical issues in rheumatic disease. The themed issues of AC&R are an example of this valuable source (2). In addition to our regular issues, AC&R has covered global scientific insights in the special reserved pages of our themed issues. The topics for the themed issues are often areas of highlighted interest brought to our attention by our readers, often at annual scientific meetings around the world. Themed issues have provided a key concentration of papers across essential topics, such as Muscle and Bone in the Rheumatic Diseases (published in 2012), Fatigue and the Rheumatic Diseases (published in 2016), Rheumatology Registries, Big Data, and the Rheumatic Diseases (published in 2020), and, under current consideration, Rehabilitation Sciences and the Rheumatic Diseases (to be published in 2022). Our CPCs (Clinicopathologic Conferences) are exceptional medical educational tools used worldwide by many fellowships (and first presented in Boston at Harvard Medical School

[1]Marian T. Hannan, DSc, MPH: Hinda and Arthur Marcus Institute for Aging Research, Hebrew SeniorLife, Beth Israel Deaconess Medical Center, and Harvard Medical School, Boston, Massachusetts; [2]Leslie J. Crofford, MD: Vanderbilt University Medical Center, Nashville, Tennessee.

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

Address correspondence to Marian T. Hannan, DSc, MPH, Marcus Institute for Aging Research, Hebrew SeniorLife, Harvard Medical School, 1200 Centre Street, Boston MA, 02131. Email: hannan@hsl.harvard.edu.

Submitted for publication April 28, 2021; accepted April 29, 2021.

in 1900) (3). Another global product is the AC&R *2020 Outcome Measurements Supplement* on patient-reported outcomes (Dr. Patricia Katz, Guest Editor), which is used worldwide and contains expertise provided by scientists and educators that transcends geographic and academic traditional boundaries. These articles are incredibly useful for clinicians, scientists, and researchers, which is reflected in the high number of article downloads every year (>4,000 downloaded articles in 2020) that occur in many countries. Members of the American College of Rheumatology (ACR) receive AC&R as a benefit of membership, but the impact of our published work is appreciated globally.

Authors and readers of AC&R are our *raison d'etre*. It is an obvious statement. The commitment of authors, readers, and those who conduct peer review have improved scientific communications and enhanced journal processes at every level. The number of authors submitting manuscripts has increased substantially over the 10 years of Dr. Hannan's tenure as Editor. The range of topics and new areas we cover has also expanded. Readership surveys have shown strong appreciation of AC&R content and coverage. Also, readers increasingly download articles of keen interest and comment on social media about AC&R articles. The subset of authors and readers who perform peer review of AC&R articles are owed a special thanks as their thoughts and insights make our published articles better and more pertinent for our readership. Peer review is a cornerstone of research communication and usefulness; our reviewers deserve special mention as the quality of the comments and direction for ultimate impact on rheumatology patients is second-to-none. We appreciate these contributions that solidify AC&R's reputation as a key rheumatology journal in which to publish premiere articles.

Heartfelt thanks are a topic at all of our Editorial and Associate Editor meetings for AC&R. We certainly appreciate the authors, readers, and reviewers. Behind the scenes, many people also worked tirelessly to support AC&R's accomplishments. Our current Managing Editor, Maggie Parry (and previously Nancy Parker), brings her professional skills and knowledge to the forefront of AC&R productions, along with her incredible editorial staff in the AC&R Atlanta home office. The efficient peer review process and excellent communication is due to the efforts and expertise of our Assistant Editor, Margaret L. Graton (and her predecessors Latoya Fladger and Belinda Wong).

A key feature contributing to the continuing success of AC&R is the selection of a team of talented and dedicated associate editors. AC&R has been blessed with superb associate editors. During Dr. Hannan's first tenure as editor-in-chief (4) from 2011 to 2016, the excellent inaugural associate editors included Drs. Hermine Brunner, Karen H. Costenbader, Leslie J. Crofford, Robert F. DeVellis, Agustin Escalante, Monique A. M. Gignac, Sunny Kim, Michael LaValley, Carol A. Oatis, Michael M. Ward, and Allan Gelber. It is important to note that our associate editors reflect the diverse membership of the ACR and the ARP in terms of clinical and academic backgrounds, content areas, and geography.

For the 2016-2021 term, Dr. Hannan was joined by Dr. Leslie J. Crofford as deputy editor with the outstanding team of associate editors, including Drs. Catherine Backman, Bonnie Bermas, Hermine Brunner, Robert F. DeVellis, David I. Daikh, Monique A. M. Gignac, Seoyoung C. Kim, Michael LaValley, Carlo A. Marra, Kaleb Michaud, Pascale Schwab, and Daniel K. White. These associate editors bring their expertise and knowledge to ensure the Journal's content is pertinent and scientifically sound and to make sure the reported work is innovative. We also thank the ACR Committee on Journal Publications, and the ACR Board of Directors for their leadership, encouragement, and support. They have provided the tools and energy to keep AC&R both "on-our-toes" and at the cutting edge of rheumatologic scientific reporting.

During the past 10 years, and especially the last 5, the teams at AC&R have identified and met challenges and opportunities to grow AC&R into a strong publication with worldwide impact. By many metrics (subscribers, number of articles, impact factor, medical and lay press coverage, social media standings, etc.), AC&R is recognized as one of the most influential rheumatology journals, with wide visibility across a broad range of clinical and academic communities.

As editors of AC&R, we took great joy in our responsibility to move the field forward, aid in communicating the major research findings by multidisciplinary rheumatology clinical researchers, and provide important education on rheumatology issues to our readers. The Journal shows many signs of continued growth and impact. We expect that the firm foundation that has been left by the previous editors of the Journal since 1988 and the current standings will bring much success as the future of AC&R unfolds. We are elated that Dr. Kelli Allen and her deputy editors, Drs. Sam Lim and Todd Schwartz, will bring a great influx of energy to the journal, and we are excited about their vision for enhancements to the journal. AC&R is a strong and essential presence in clinical rheumatology and rehabilitation sciences across the world. We leave our editor positions with confidence that AC&R will continue to thrive in the years to come.

It has been a great honor to serve as editors for AC&R and to see the Journal's growth and influence multiply over the years.

REFERENCES

1. Reeves S, Pelong F, Harrison R, Goldman J, Zwarestein M. Interprofessional collaboration to improve professional practice and healthcare outcomes. Cochrane Database Syst Rev 2017. URL: https://doi.org/10.2001/14651858.CD00007.pub3.
2. Hannan MT. A quarter-century of excellence and still growing [editorial]. Arthritis Care Res (Hoboken) 2013;65:1–3.
3. Hajar R. The clinicopathologic conference. Heart Views 2015; 16:170–3.
4. Hannan MT. Arthritis Care & Research: Continued success and evolution [editorial]. Arthritis Care Res (Hoboken) 2011;63:925–6.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 767–769
DOI 10.1002/acr.24181

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

Foot Osteoarthritis: Addressing an Overlooked Global Public Health Problem

Yvonne M. Golightly[1] and Lucy S. Gates[2]

Foot pain affects at least 1 in 3 adults over the age of 45 years (1), and 1 in 6 adults ages ≥50 years with foot pain also have radiographic evidence of foot osteoarthritis (OA) (2). Unfortunately, our knowledge of foot OA and its burden substantially lags behind that of other joint sites (i.e., knee, hip, and hand). Given the link of foot pain and disorders with disability and reduced quality of life (3), a greater understanding of foot OA and its pathogenesis and risk factors is needed, especially since foot OA is a serious disease with a significant public health burden.

In this issue of *Arthritis Care & Research* (4), Arnold et al advance the field of foot OA with their novel study on associations of foot and leg muscle strength with symptomatic midfoot OA, a frequent type of foot OA. This study builds on prior research demonstrating the associations of muscle weakness with OA at the hip, knee, and hand with pain and impaired physical function. In this cross-sectional study, 52 participants with midfoot OA demonstrated less strength in all foot and leg muscle groups (i.e., ankle plantarflexors, dorsiflexors, invertors, and evertors; hallux and lesser digit plantarflexors) compared to 36 asymptomatic individuals without radiographic midfoot OA. Additionally, greater ankle invertor strength was associated with less foot pain among individuals with midfoot OA. Based on the study design, the direction of these associations is not known, but these results provide initial data to support a future investigation of foot and leg muscle strengthening as a potential treatment for symptomatic midfoot OA.

The authors call attention to limitations of their study, including the criteria for defining absence of midfoot OA and the smaller sample size of the asymptomatic participant group. They also acknowledge the necessity for larger samples and longitudinal studies to examine the role of foot and leg muscle weakness with pain and structural outcomes in the midfoot, along with a need to investigate foot and leg muscle strengthening as a treatment for midfoot OA. These significant research needs, along with limitations observed in other studies of foot OA, suggest multiple research areas required for progressing our understanding of the etiology of foot OA and its management. The purpose of this editorial is to discuss important considerations for improving future foot OA research across populations globally. Establishing standard definitions and measures of foot OA is essential for longitudinal analyses to identify risk factors for and subgroups of foot OA, as well as for clinical trials of treatments for this highly prevalent and disabling condition.

The current knowledge gaps in foot OA research are in part due to a lack of longitudinal population-based data to determine the number of people who may require care and to estimate the burden of foot OA among the general population. For the foot, there is a lack of agreed clinical definitions for many measures; for example, there is currently no agreement on how to capture OA-related foot pain. Only a handful of international cohorts collect foot data, but unlike the knee and hip, measures of foot OA, pain, and physical function often are not comparable across cohorts, therefore limiting comparative estimates across geographical regions.

Previously, defining radiographic foot OA was limited by the Kellgren/Lawrence grading system, which depended greatly on the presence of osteophytes. This scoring system may not be reliable at the foot, where joint space narrowing could occur alone or prior to osteophyte formation. Until recently, previous studies have focused mainly on first metatarsophalangeal (MTP) joint radiographic OA, likely due to difficulties with radiographic interpretation at the smaller more complex joints. An important advancement in the field of foot OA was the development of the La Trobe Foot Atlas (5), which individually scores radiographic presentation of osteophytes and joint space narrowing at the first MTP joint and 4 midfoot joints: first and second cuneometatarsal joints, navicular-first cuneiform joint, and talonavicular joint. This atlas, which was used by Arnold et al (4), has improved our ability to estimate the prevalence of foot OA in populations. With the recent emergence of longitudinal radiographic foot data, questions

[1]Yvonne M. Golightly, PT, PhD: University of North Carolina at Chapel Hill; [2]Lucy S. Gates, PhD: University of Southampton, Southampton, UK and University of Oxford, Oxford, UK.

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

Address correspondence to Yvonne M. Golightly, PT, PhD, Thurston Arthritis Research Center, University of North Carolina, 3300 Thurston Building, Campus Box 7280, Chapel Hill, NC 27599-7280. Email: golight@email.unc.edu.

Submitted for publication February 4, 2020; accepted in revised form March 3, 2020.

have arisen over how to define incident and progressive disease, and unfortunately, the La Trobe Foot Atlas was not designed to address these issues. For example, would we consider a foot joint with OA to have a clinically important worsening of OA if there is an increase solely in the osteophyte score, or should scores of both the osteophyte and joint space narrowing grades progress? Do these features have equal weighting at each of the 5 joints? Should radiographic OA in 1 joint in the foot be considered radiographic foot OA, or should this definition depend on multiple joint involvement? If the presence of radiographic features is required in multiple joints, how many and which joints?

Evidence for phenotypes of foot OA has begun to emerge, with 3 distinct classes of radiographic foot OA, including no/minimal foot OA, isolated first MTP joint OA, and polyarticular foot OA (6). This work is a useful start in the understanding of patterns of involvement across different joints in the feet, but further investigation across large cohorts is needed.

Questions over the radiographic definition of incidence and progression previously have arisen at the knee and the hand, with solutions suggested based on years of previously published work in the field (7–10). Because of the potential multiple joint involvement in foot OA, definitions used for incidence and progression of hand OA may serve as a model (8,9). Perhaps, the definition of radiographic foot OA would vary based on which joint or joints are involved and their potential association with pain and disability, in which case we should consider which features correspond to such clinical factors.

Defining symptomatic OA at the foot is necessary for producing estimates that are meaningful to clinicians and patients. To date, large cross-sectional cohort studies of symptomatic foot OA have been limited by general definitions of foot pain (e.g., presence of pain anywhere in the foot) (2,11). A notable strength of the study by Arnold et al was the use of a foot manikin and palpation of midfoot joints to determine that foot pain was present in the same region as the joints of interest, providing an example of how other studies could define symptomatic OA in a particular region of the foot. Definitions of foot pain also should consider the duration, quality (e.g., sharp, ache, stiffness), and severity of symptoms. Defining worsening of symptomatic foot OA is considerably more complicated than radiographic foot OA. For instance, would a worsening of symptoms in the presence of stable radiographic features at the foot be considered worsening of symptomatic OA, or should radiographic features progress as well? Likewise, would worsening in structural severity of either or both features in the presence of stable symptoms be considered worsening symptomatic OA?

Other possible ways of defining foot OA that could be developed include the use of images other than radiography, such as magnetic resonance imaging and ultrasound. A clinical definition that could be broadly used to help clinicians diagnose foot OA without imaging would be valuable. Similar to the American College of Rheumatology Diagnostic Guidelines for knee, hip, and hand OA, a clinical definition of foot OA could be developed based on patient age and the presence of signs and symptoms.

Due to a lack of longitudinal investigation and consistent definitions of foot outcomes, there is little evidence available regarding the potential risk factors for foot OA, without which it is difficult to develop interventional research. In fact, there are few randomized controlled trials in foot OA. Existing trials provide initial evidence for the effectiveness of pain relief of physical therapy, rocker sole shoes, foot orthoses, and surgical interventions in first MTP joint OA and of prefabricated orthoses in midfoot OA (12). The role of occupational activities, injuries, and physical activity in foot OA are not known; longitudinal investigations of these factors would determine their part in disabling foot OA and inform interventions.

In summary, standardizing definitions of foot OA would help us better understand the disease pathogenesis and risk factors, perhaps delineating different phenotypes of foot OA that require distinctive management approaches. By ensuring that future work related to foot OA can be compared across the globe, the true burden of foot OA can be established. As quality longitudinal data are gathered over time, we can gain a better comprehension of the natural history, treatment response, and economic impact of foot OA. Lessons learned for defining incident and progressive OA at the knee and hand can be applied to the foot. With standard definitions, clinical trials could be implemented to determine the effectiveness of applying existing OA management approaches to foot OA or developing new interventions. We are at a pivotal time for providing definitions that can be used in all future longitudinal studies, which will inform the global public health impact of foot OA and advance our ability to treat this disabling condition.

AUTHOR CONTRIBUTIONS

Both authors drafted the article, revised it critically for important intellectual content, approved the final version to be published, and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Golightly YM, Hannan MT, Shi XA, Helmick CG, Renner JB, Jordan JM. Association of foot symptoms with self-reported and performance-based measures of physical function: the Johnston County osteoarthritis project. Arthritis Care Res (Hoboken) 2011;63:654–9.
2. 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: cross-sectional findings from the clinical assessment study of the foot. Ann Rheum Dis 2015;74:156–63.
3. Thomas MJ, Roddy E, Zhang W, Menz HB, Hannan MT, Peat GM. The population prevalence of foot and ankle pain in middle and old age: a systematic review. Pain 2011;152:2870–80.
4. Arnold JB, Halstead J, Grainger AJ, Keenan AM, Hill CL, Redmond AC. Foot and leg muscle weakness in people with midfoot osteoarthritis. Arthritis Care Res (Hoboken) 2021;73:772–80.
5. 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.

6. Rathod T, Marshall M, Thomas MJ, Menz HB, Myers HL, Thomas E, et al. Investigations of potential phenotypes of foot osteoarthritis: cross-sectional analysis from the clinical assessment study of the foot. Arthritis Care Res (Hoboken) 2016; 68:217–27.
7. Felson DT, Niu J, Guermazi A, Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. Ann Rheum Dis 2011;70:1884–6.
8. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis 2011;70:1581–6.
9. Haugen IK, Magnusson K, Turkiewicz A, Englund M. The prevalence, incidence, and progression of hand osteoarthritis in relation to body mass index, smoking, and alcohol consumption. J Rheumatol 2017;44:1402–9.
10. Zhang Y, Niu J, Felson DT, Choi HK, Nevitt M, Neogi T. Methodologic challenges in studying risk factors for progression of knee osteoarthritis. Arthritis Care Res (Hoboken) 2010;62:1527–32.
11. Flowers P, Nelson AE, Hillstrom HJ, Renner JB, Jordan JM, Golightly YM. Cross-sectional analysis of foot osteoarthritis frequency and associated factors: The Johnston County Osteoarthritis Project. Arthritis Rheumatol 2017;69:S10.
12. Roddy E, Menz HB. Foot osteoarthritis: latest evidence and developments. Ther Adv Musculoskelet Dis 2018;10:91–103.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 770–771
DOI 10.1002/acr.24493

EDITORIAL

Hydroxychloroquine: Not a Heart Breaker!

Julianna Desmarais[1] and Mark S. Link[2]

Prior to 2020, there was a lack of general awareness of the potential for hydroxychloroquine (HCQ) and chloroquine (CQ) to cause cardiac arrhythmias, as rheumatologists and dermatologists have prescribed these medications for decades with no obvious cardiac safety signal. And while HCQ and CQ are listed as drugs with a known risk of torsades de pointes (www.credibleMeds.org) due to their effect on cardiac ion channels, there was little thought of HCQ causing arrhythmias (1). However, with the recent coronavirus disease 2019 (COVID-19) pandemic, reports of HCQ use being associated with prolongation of the corrected QT (QTc) interval and torsades de pointes emerged (2,3). Granted, azithromycin, which is another agent known to block potassium cardiac ion channels that prolong the QTc, was frequently co-administered in patients infected with COVID-19 who were also treated with HCQ. In addition, the patients receiving these medications were usually hospitalized and often quite ill. These patients often had some evidence of cardiac involvement secondary to COVID-19, which itself was found to be cardiotoxic (4,5). All of these factors could increase the risk of cardiac arrhythmias.

In this issue of *Arthritis Care & Research*, Gupta et al report on the reduction of atrial fibrillation (AF) incidence among patients with systemic lupus erythematosus (SLE) who were treated with HCQ (6). In this retrospective cohort analysis of adult patients with SLE from December 1, 2014 to May 30, 2017, a total of 1,647 patients were included, of which 917 were HCQ users and 730 were nonusers. Patients with previous AF were excluded from the analysis, and episodes of AF in the first year of use of HCQ were not counted as end points, allowing for a run-in period. HCQ users were older, less likely to have hypertension, and more likely to have coronary artery disease, heart failure, and diabetes mellitus, which are all risk factors for AF. All comorbidities, except for the lower incidence of hypertension, would favor a higher risk of AF in HCQ users.

A total of 23 AF events were captured, including 3 in HCQ users and 20 in nonusers. Despite the more common risk factors for AF in HCQ users, even in the unadjusted analysis, administration of HCQ was associated with decreased incidence of AF. When controlling for AF risk factors with logistic regression analysis, the odds ratio for AF was 0.12, with a 95% confidence interval of 0.034–0.39 (P = 0.0005). The propensity score matching method showed essentially the same result.

There were 4 episodes of ventricular arrhythmia, including 1 ventricular fibrillation in a nonuser and 2 ventricular tachycardias and 1 torsade de pointes in the HCQ user group, of which none were fatal. Although numerically higher in the HCQ users, this was not statistically significant. Still, based on the mechanism of action of HCQ on cardiac ion channels, an increase in risk of ventricular arrhythmias is plausible.

While this study was intended to highlight the potential benefit of HCQ on AF incidence, it supports the concept that HCQ possesses cardiac ion channel–blocking properties and should be viewed in this light. HCQ and CQ are derivatives of quinidine. Quinidine is one of the earliest-developed anti-arrhythmic drugs that is efficacious for both atrial and ventricular arrhythmias but has a proarrhythmic risk of torsades de pointes (7). Its efficacy lies primarily in blockade of the rapid cardiac delayed-rectifier potassium current (I_{Kr}), which HCQ and CQ block as well. Blockade of this channel prolongs repolarization, leading to increased refractory time, the basis of its anti-arrhythmic effect in cardiac tissues. Not only is this channel the abnormal channel in long QT syndrome type 1 (8), but it is the same channel that is blocked by anti-arrhythmic drugs developed to treat atrial and ventricular arrhythmias, such as sotalol, dofetilide, and amiodarone.

HCQ and CQ are moderate I_{Kr} blockers, similar to macrolides, antipsychotics, and antifungals in their potency to block I_{Kr}. Like anti-arrhythmic drugs, HCQ and CQ do have predictable actions on the QT interval (9). This has been observed in animal studies and now more recently in COVID-19–infected patients treated with HCQ, particularly when co-administered with azathioprine, another I_{Kr} blocker (3,5). There are also now both case reports and larger series that show HCQ does have QT prolongation effects in individuals with SLE and those with rheumatoid arthritis in the pre–COVID-19 era (9–11). Multiple randomized and observational studies have shown that anti-arrhythmic agents that block I_{Kr} have

[1]Julianna Desmarais, MD: Oregon Health & Science University, Portland; [2]Mark S. Link, MD: UT Southwestern Medical Center, Dallas, Texas.

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

Address correspondence to Mark S. Link, MD, UT Southwestern Medical Center, 2001 Inwood Road, Suite WC05.852, Dallas, TX 75390-9254. Email: Mark.Link@UTSouthwestern.edu.

Submitted for publication October 8, 2020; accepted in revised form October 20, 2020.

been associated with a reduction in AF risk (12). However, these agents raise the concern for prolongation of QTc and, accordingly, increased risk of ventricular arrhythmias, including torsades de pointes and subsequent sudden cardiac death.

The study by Gupta et al (6) does have a limitation of being a retrospective study, and while logistic regression and propensity score matching methods showed an association with reduced AF incidence, this is an association and not necessarily a causal relationship. Residual confounding and indication bias for HCQ use or nonuse could still be playing a role. In addition, the authors do not report QTc changes with HCQ, nor adherence to HCQ, which is a known issue (13). Still, the authors are to be commended for the evaluation of HCQ as a potential anti-arrhythmic agent and not just as a proarrhythmic agent. As lupus may be a risk for AF, and those with AF had higher mortality in 1 study, a medication that is associated with a lower incidence of AF would be beneficial in this high-risk population (14,15).

This current study adds to the data that there is a cardiac effect of HCQ. Further studies that more clearly delineate HCQ adherence and any association of blood levels of HCQ as it relates to its effect on the I_{Kr} channel will be helpful. In addition, larger data-based studies on HCQ exposure and its impact on the QTc are necessary before drawing conclusions regarding HCQ cardiac effects. At this point, it is unclear about what type of screening should be done as we follow these patients. Finally, we need to be mindful that HCQ is and has been used for years in patients with SLE and rheumatoid arthritis with clear benefits in mitigating disease activity and progression, and this benefit may far outweigh any particular risk.

AUTHOR CONTRIBUTIONS

Both authors drafted the article, revised it critically for important intellectual content, approved the final version to be published, and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Woosley RL, Heise CW, Gallo T, Tate J, Woosley D, Romero KA. QT drugs list. 2020. URL: www.credibleMeds.org.
2. Chorin E, Wadhwani L, Magnani S, Dai M, Shulman E, Nadeau-Routhier C, et al. QT interval prolongation and torsade de pointes in patients with covid-19 treated with hydroxychloroquine/azithromycin. Heart Rhythm 2020;17:1425–33.
3. Maraj I, Hummel JP, Taoutel R, Chamoun R, Workman V, Li C, et al. Incidence and determinants of QT interval prolongation in covid-19 patients treated with hydroxychloroquine and azithromycin. J Cardiovasc Electrophysiol 2020;31:1903–7.
4. Shi S, Qin M, Shen B, Cai Y, Liu T, Yang F, et al. Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China. JAMA Cardiol 2020;5:802–10.
5. Kochi AN, Tagliari AP, Forleo GB, Fassini GM, Tondo C. Cardiac and arrhythmic complications in patients with COVID-19. J Cardiovasc Electrophysiol 2020;31:1003–8.
6. Gupta A, Shields K, Manzi S, Wasko MC, Sharma T. Association of Hydroxychloroquine use with decreased incident atrial fibrillation in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2021;73:828–32.
7. Selzer A, Wray HW. Quinidine syncope. Paroxysmal ventricular fibrillation occurring during treatment of chronic atrial arrhythmias. Circulation 1964;30:17–26.
8. Sanguinetti MC, Jiang C, Curran ME, Keating MT. A mechanistic link between an inherited and an acquired cardiac arrhythmia: Herg encodes the ikr potassium channe/. Cell 1995;81:299–307.
9. Hooks M, Bart B, Vardeny O, Westanmo A, Adabag S. Effects of hydroxychloroquine treatment on QT interval. Heart Rhythm 2020; 17:1930–5.
10. Chen CY, Wang FL, Lin CC. Chronic hydroxychloroquine use associated with QT prolongation and refractory ventricular arrhythmia. Clin Toxicol (Phila) 2006;44:173–5.
11. Morgan ND, Patel SV, Dvorkina O. Suspected hydroxychloroquine-associated QT-interval prolongation in a patient with systemic lupus erythematosus. J Clin Rheumatol 2013;19:286–8.
12. Lafuente-Lafuente C, Longas-Tejero MA, Bergmann JF, Belmin J. Antiarrhythmics for maintaining sinus rhythm after cardioversion of atrial fibrillation. The Cochrane Database of Systematic Reviews 2012:CD005049.
13. Feldman CH, Yazdany J, Guan H, Solomon DH, Costenbader KH. Medication nonadherence is associated with increased subsequent acute care utilization among medicaid beneficiaries with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2015;67:1712–21.
14. Lim SY, Bae EH, Han KD, Jung JH, Choi HS, Kim CS, et al. Systemic lupus erythematosus is a risk factor for atrial fibrillation: a nationwide, population-based study. Clin Exp Rheumatol 2019;37:1019–25.
15. Chen SK, Barbhaiya M, Solomon DH, Guan H, Yoshida K, Feldman CH, et al. Atrial fibrillation/flutter hospitalizations among US Medicaid recipients with and without systemic lupus erythematosus. J Rheumatol 2020;47:1359–65.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 772–780
DOI 10.1002/acr.24182

American College of Rheumatology
Empowering Rheumatology Professionals

Foot and Leg Muscle Weakness in People With Midfoot Osteoarthritis

John B. Arnold,[1] Jill Halstead,[2] Andrew J. Grainger,[3] Anne-Maree Keenan,[4] Catherine L. Hill,[5] and Anthony C. Redmond[4]

Objective. To compare foot and leg muscle strength in people with symptomatic midfoot osteoarthritis (OA) with asymptomatic controls, and to determine the association between muscle strength, foot pain, and disability.

Methods. Participants with symptomatic midfoot OA and asymptomatic controls were recruited for this cross-sectional study from general practices and community health clinics. The maximum isometric muscle strength of the ankle plantarflexors, dorsiflexors, invertors and evertors, and the hallux and lesser toe plantarflexors was measured using hand-held dynamometry. Self-reported foot pain and foot-related disability were assessed with the Manchester Foot Pain and Disability Index. Differences in muscle strength were compared between groups. Multivariable regression was used to determine the association between muscle strength, foot pain, and disability after adjusting for covariates.

Results. People with midfoot OA (n = 52) exhibited strength deficits in all muscle groups, ranging from 19% (dorsiflexors) to 30% (invertors) relative to the control group (n = 36), with effect sizes of 0.6–1.1 ($P < 0.001$). In those with midfoot OA, ankle invertor muscle strength was negatively and independently associated with foot pain (β = –0.026 [95% confidence interval (95% CI) –0.051, –0.001]; $P = 0.045$). Invertor muscle strength was negatively associated with foot-related disability, although not after adjustment for depressive symptoms (β = –0.023 [95% CI –0.063, 0.017]; $P = 0.250$).

Conclusion. People with symptomatic midfoot OA demonstrate weakness in the foot and leg muscles compared to asymptomatic controls. Preliminary indications from this study suggest that strengthening of the foot and leg muscles may offer potential to reduce pain and improve function in people with midfoot OA.

INTRODUCTION

Foot osteoarthritis (OA) is a common cause of foot pain in older adults, affecting 1 in 6 adults ages >50 years in the UK (1). One of the most frequent presentations is symptomatic midfoot OA (12%), which affects the talonavicular (5.8%), navicular-first cuneiform (5.2%), or cuneometatarsal joints (3.9–6.8%) (2). Midfoot OA is associated with significant pain (2,3) and difficulty in walking (3) and climbing stairs (4). Severe midfoot OA may cause foot deformity, changes in foot posture, and difficulty with finding suitable footwear (5). Symptoms appear to change little over time, with midfoot OA causing persistent foot pain and foot-related disability over 18 months (6).

Demonstrated risk factors for midfoot OA include female sex, age, obesity, intermediate/routine occupational class, previous foot/ankle injury, and pain in other weight-bearing joints (2). Within the foot, midfoot OA is associated with bony malalignment, resulting in reduced medial longitudinal arch height (7) and a more pronated foot posture (8). This is accompanied by reduced sagittal plane range-of-motion in the medial longitudinal arch (4) and

The views expressed herein are those of the authors and do not necessarily represent those of the NHS, the NIHR, Health Education England, or the Department of Health.

Supported by the NIHR Leeds Biomedical Research Centre, the Leeds Experimental Osteoarthritis Treatment Centre funded by Arthritis Research UK (grant 20083), and the NIHR Clinical Research Network. Dr. Arnold is a National Health and Medical Research Council of Australia Early Career research fellow (ID: 1120560). Dr. Halstead's work was supported by the Health Education England/NIHR Integrated Clinical Academic Programme Bridging Scheme. Drs. Redmond and Keenan are NIHR senior investigators.

[1]John B. Arnold, PhD: University of Leeds, NIHR Leeds Biomedical Research Centre, Leeds, UK, and University of South Australia, Adelaide, South Australia; [2]Jill Halstead, PhD, FFPM (Glas), RCPS: University of Leeds, NIHR Leeds Biomedical Research Centre and Leeds Community Healthcare NHS Trust, Leeds, UK; [3]Andrew J. Grainger, BM, BS, FRCR, FRCP: University of Leeds, NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK; [4]Anne-Maree Keenan, PhD, FFPM, RCPS (Glas), Anthony C. Redmond, PhD, FFPM, RCPS (Glas), FCPM: University of Leeds, NIHR Leeds Biomedical Research Centre, Leeds, UK; [5]Catherine L. Hill, MBBS, MD, MSc, FACP: The Queen Elizabeth Hospital and University of Adelaide, Adelaide, South Australia.

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

Address correspondence to John B. Arnold, PhD, School of Health Sciences, University of South Australia, City East Campus, Adelaide, South Australia 5001, Australia. Email: john.arnold@unisa.edu.au.

Submitted for publication September 26, 2019; accepted in revised form March 3, 2020.

SIGNIFICANCE & INNOVATIONS

- To the best of our knowledge, this study is the first investigation of foot muscle strength in people with midfoot osteoarthritis and its relationship to foot pain and foot-related disability.
- Foot and leg muscle strength is reduced in all muscle groups in people with midfoot osteoarthritis compared to asymptomatic controls.
- Muscle strength was independently and inversely related to foot pain in people with midfoot osteoarthritis.

elevated forces and pressures under the midfoot during walking (7,9). Despite a growing understanding of the clinical features and functional consequences of midfoot OA, previous studies have focused on selected structural and biomechanical components, such as radiographic alignment, foot motion, and plantar pressures. Given the importance of muscle strength for joint stability and control, and the relationship between muscle weakness and OA in other joints (10,11), understanding muscle function in midfoot OA warrants further investigation. To the best of our knowledge, however, no studies have investigated muscle strength in people with midfoot OA.

Muscle weakness is a hallmark of OA at other joints such as the hip (10), knee (12), and hand (13) and is associated with pain (14), joint instability (15), and performance-based (15) and self-reported physical function (16). Deficits in muscle strength appear early in OA (17) and, in the knee, have been associated with incident radiographic disease (18) as well as symptomatic and functional decline (11). Muscle strengthening exercises are a core component of OA management and are included in international clinical guidelines (19–21). There is, however, little research on muscle strength in people with foot OA, particularly in midfoot OA. One prior study of first metatarsophalangeal joint OA investigated the relationship between symptoms and demographic and clinical characteristics, including plantarflexion strength of the hallux (22). This study showed that hallux plantarflexion strength was negatively, although weakly, associated with first metatarsophalangeal joint pain. Whether foot and leg muscle weakness is present in people with midfoot OA has not been investigated. Furthermore, whether muscle strength is associated with patient-reported outcomes in midfoot OA, such as pain and function, has not been evaluated. Greater understanding of whether foot and leg muscle weakness is a feature of midfoot OA has potential clinical implications, as muscle strength is modifiable (23) and may be a viable target for treatment. Pain and disability are the main reasons why people with OA seek treatment (24), therefore identification of the factors associated with symptoms has the potential to improve the design of treatments for this condition.

The aims of this study were to compare foot and leg muscle strength in people with symptomatic midfoot OA with asymptomatic controls and to determine whether muscle strength was associated with self-reported pain and foot-related disability. It was hypothesised a priori that people with midfoot pain and OA would present with foot and leg muscle weakness, and that muscle strength would be negatively associated with pain and foot-related disability.

PATIENTS AND METHODS

Study design and recruitment. This was a cross-sectional study involving people with midfoot pain and OA and asymptomatic controls. Participants were recruited from the community via advertisements, general practitioners, and health clinics. Ethics approval was obtained from the Leeds East Research Ethics Committee (17/YH/0261). All participants provided written informed consent prior to their involvement.

Participants. Symptomatic participants were ages >40 years, had pain in the midfoot for >3 months with an average weekly pain severity of ≥3 of 10 on an 11-point numerical rating scale (NRS) that occurred with or worsened following weight-bearing activities. The presence of midfoot pain was confirmed by participants marking the site of pain on a foot manikin (25,26), and supplemented by clinical examination to assess whether pain was reported on palpation of the talonavicular (TNJ), naviculocuneiform (NCJ), or cuneiform-metatarsal (CMJ) joints. Weight-bearing dorsoplantar and lateral radiographs were used to grade OA in either the TNJ, NCJ, or first or second CMJ by a musculoskeletal radiologist (AJG) using the La Trobe Foot Atlas (27). An established case definition was used, where a joint was considered to have OA with a score of ≥2 for osteophytes or joint space narrowing (JSN) on either the dorsoplantar or lateral views (27). To establish intrarater reliability, scoring was repeated on 20 participants, 3 months apart, without reference to the first set of scores. Exclusion criteria were >30 minutes of early morning stiffness in the feet, inflammatory arthritis, muscle or connective tissue disease, neurologic conditions, corticosteroid injection to the foot in the past 6 months, stress fracture or history of foot surgery, or contraindications to radiographs. Concurrent knee or hip pain was permitted if the pain intensity was not greater than their midfoot pain and was quiescent (average daily pain less than midfoot pain and <2 in the past week on NRS).

Control group participants were required to be age >40 years and free from foot or lower extremity joint pain. This was verified using an 11-point NRS for foot pain and a body pain manikin. Additional exclusion criteria for controls were presence of radiographic OA (osteophytes or JSN >1 on either view in any of the midfoot joints [TNJ, NCJ, first or second CMJ]), contraindications to radiograph, inflammatory arthritis, muscle or connective tissue disease, neurologic conditions, stress fracture, or lower extremity bone or joint surgery in the past 12 months. A meaningful a priori sample size calculation was not performed due to the

lack of prior research on muscle strength in people with midfoot OA. Therefore, the sample size, including unbalanced sampling of controls, was dictated by the period of recruitment for this study (12 months) and available funding.

Muscle strength testing. The maximal isometric strength of the leg and foot muscles was measured using a CITEC handheld dynamometer (CIT Technics). The device has a range of 0–500 newtons (N) and, according to manufacturer's data, was factory-calibrated to a sensitivity of 0.1%. Testing was performed by an experienced clinician (JBA) using standardized protocols, which have well-established intrarater (intraclass correlation coefficient [ICC] 0.83–0.94) and interrater reliability (ICC 0.77–0.88) (28). All testing was performed by the same researcher, with the participants in a supine position and the lower extremity stabilized proximal to the ankle joint. The muscle groups that were evaluated included ankle plantarflexors, dorsiflexors, invertors and evertors, hallux plantarflexors, and lesser toe plantarflexors.

For plantarflexion strength, the dynamometer was positioned on the plantar surface of the foot just proximal to the first metatarsal head, and for dorsiflexion it was placed on the dorsal surface of the foot just proximal to the metatarsal heads. To prevent movement during plantarflexion strength tests, the examiner anchored the dynamometer on the anterior aspect of the participants' thigh. For inversion, the dynamometer was placed on the medial border of the foot at the midpoint of the shaft of the first metatarsal, and for eversion it was placed on the lateral border of the foot over the midpoint of the fifth metatarsal. Hallux plantarflexor strength involved positioning of the dynamometer on the plantar surface of the interphalangeal joint and on the plantar surface of the toes for lesser toe strength. To standardize joint position across feet of different sizes with the same dynamometer, both the hallux and lesser toes were dorsiflexed into the participants comfortable end range of motion, as per the original protocol (28). The ankle was also placed in a plantarflexion during testing of the hallux and lesser toe muscles to prevent co-contraction of the ankle plantarflexors.

Before testing, the required movement was passively demonstrated by the examiner. This was followed by asking the participants to perform the movement against the dynamometer to ensure the correct action could be performed. The "make" technique was used requiring participants to exert a maximum voluntary contraction (MVC) against the dynamometer. Three valid MVCs of 3–5 seconds were obtained for each muscle group, with 15 seconds rest in between each trial (29). Verbal encouragement was standardized during the contractions, with the examiner telling each participant to "go ahead-push-push-push-push and relax" (30). The mean value of 3 trials was used for analysis (28). For participants with OA, the symptomatic side was the index foot; in cases of bilateral OA, only the most painful foot was tested. For controls, the right side was tested. To account for any differences in height or weight between groups, muscle strength data (in newtons) were normalized to body mass multiplied by height (% weight × height).

Foot pain and disability assessment. Pain severity in the past week, past month, and while walking was documented with an 11-point NRS for each, ranging from 0 (no pain) to 10 (worst pain imaginable). Pain and foot-related disability were assessed using the Manchester Foot Pain and Disability Index (MFPDI) (31), a 19-item questionnaire with subscales of foot pain (5 items), disability (10 items), appearance (2 items), and work or leisure (2 items). Each item is scored from 0 (none of the time) to 2 (on most days/every day). Pain subscale scores range from 0 to 10 and function scores from 0 to 20, with higher scores indicating more pain or worse foot-related disability. The MFPDI, which has been previously used in people with midfoot OA (2,3,6), displays good construct validity and internal consistency (31). Prior to analysis, raw scores were converted to Rasch-transformed interval level scores.

Other clinical characteristics. Due to the relationship and importance of depression to the development and experience of foot pain (32,33), information on depressive symptoms was obtained by participants completing the Hospital Anxiety and Depression Scale (HADS) (34), a 14-item questionnaire with 7 of these items relating to depressive symptoms (scored 0–3) with a total subscale score ranging 0–21. The psychometric properties of the HADS have been previously established (35). Questionnaires were also administered to capture general (EuroQol 5-domain) (36) and OA-specific (OA-QoL) health-related quality of life (37). Foot posture was quantified using the 6-item version of the Foot Posture Index (FPI-6), a validated and reliable clinical measure of foot posture (38,39). Each participant's foot posture (total score for the index foot) was classified according to cut points from normative data as supinated (score <0), normal (0–5), or pronated (≥6) (40).

Statistical analysis. Descriptive statistics were generated for participant characteristics, symptoms (pain NRS and MFPDI), and muscle strength scores. Normal distributions for muscle strength, pain NRS, and MFPDI scores were determined using histograms and Shapiro-Wilks tests. Independent sample *t*-tests and chi-square tests were used to compare participant characteristics and muscle strength between the midfoot pain and OA and asymptomatic control groups. Equality of variances was confirmed with Levene's test. Consistent with previous studies in foot OA (1), for the primary analysis the case definition for absence of radiographic OA in the midfoot included JSN or osteophyte grade of <2. We also conducted a further sensitivity analysis to evaluate differences in muscle strength between the midfoot OA group and asymptomatic controls using definitions of grade 0 (n = 19) and grade >0 (n = 17) for JSN or osteophytes in the midfoot joints. Differences in muscle strength were also summarized as percentage

difference (%) and with standardized effect sizes (Cohen's *d* coefficient). Intrarater reliability of radiographic scoring of foot OA was determined using percent of agreement and weighted kappa with quadratic weights.

Pearson's correlation coefficients were used to determine the strength and direction of the univariable relationship between muscle strength and MFPDI pain and function. Multivariable linear regression was used to determine the association between muscle strength and MFPDI pain (model 1) and MFPDI function (model 2), with age, sex, and body mass index (BMI) as covariates. To avoid issues with multicollinearity of predictors in the multivariable models, only the muscle strength group that displayed the strongest univariate relationship with MFPDI pain and function scores was included in the model. Due to the relationship between depression and pain, the depressive symptoms score was also entered into each model. We also adjusted for radiographic disease severity in the midfoot, represented by the total sum score of JSN and osteophytes for the TNJ, NCJ, and first and second CMJ. Results are presented as adjusted unstandardized regression coefficients (β) with 95% confidence intervals (95% CIs). The amount of variance explained by each model was determined using the adjusted r^2. All assumptions for the regression analyses were tested and met, including linearity of relationships and independence, homoscedasticity, and normality of residuals. Statistical significance was set at *P* less than 0.05. All analyses were conducted using SPSS, version 21.

RESULTS

Descriptive characteristics. Fifty-two people with midfoot OA and 36 asymptomatic controls completed all testing (Table 1). The mean age of the midfoot OA group was 62 years (73% women), with a BMI of 29.2 kg/m^2, compared to the asymptomatic controls (mean age 63 years, 66% women, BMI 27.2 kg/m^2). There were no statistically significant differences in age, sex, or BMI between groups ($P > 0.05$).

Clinical characteristics. The midfoot OA group reported moderate levels of pain over the past 24 hours (mean ± SD 3.7 ± 2.2), with slightly higher average pain over the past week (mean ± SD 4.2 ± 32.2) and while walking (mean ± SD

Table 1. Descriptive and clinical characteristics of midfoot OA and control groups*

	Midfoot OA group (n = 52)	Control group (n = 36)	*P*
Age, years	62.2 ±11.4	63.6 ± 11.7	0.586
Female sex, %	73	66	0.517
BMI (kg/m^2)	29.2 ± 5.4	27.2 ± 4.8	0.053
Joint-specific radiographic OA, no. (%)†			
Talonavicular joint	11 (21)	–	
Navicular-first cuneiform joint	21 (40)	–	
First cuneiform-metatarsal joint	18 (35)	–	
Second cuneiform-metatarsal joint	38 (73)	–	
Foot pain and functional limitation			
Dorsal midfoot pain, no. (%)	42 (81)	–	
Plantar midfoot pain, no. (%)	10 (19)	–	
Foot pain severity			
Average in past 24 hours (0–10 NRS)	3.7 ± 2.2	–	
Average in past week (0–10 NRS)	4.2 ± 2.2	–	
On walking in past week (0–10 NRS)	5.0 ± 2.6	–	
MFPDI Rasch pain‡	5.95 ± 1.6	–	
MFPDI Rasch function‡	8.62 ± 3.0	–	
Quality of life and mental health			
OA quality of life§	5.52 ± 5.9	1.08 ± 2.5	<0.001
EQ overall health (0–100)¶	68.30 ± 20.9	86.72 ± 11.9	<0.001
HADS depression#	3.87 ± 5.4	1.94 ± 3.0	0.035
HADS anxiety#	5.31 ± 6.0	3.72 ± 3.7	0.131
Foot Posture Index**			
Supinated (<0), no. (%)	4 (8)	7 (19)	
Normal (0–5), no. (%)	19 (36)	18 (50)	
Pronated (≥6), no. (%)	29 (56)	11 (31)	0.044

* Values are the mean ± SD unless indicated otherwise. BMI = body mass index; EQ = EuroQol; HADS = Hospital Anxiety and Depression Scale; MFPDI = Manchester Foot Pain and Disability Index; NRS = numerical rating scale; OA = osteoarthritis.
† Joint-specific OA does not equal 100%, as >1 midfoot joint may have OA.
‡ Higher values indicate more foot pain or foot-related disability.
§ Higher values indicate poorer OA-related quality of life.
¶ Higher values indicate better health-related quality of life.
# Higher values indicate more depression or anxiety symptoms.
** Foot Posture Index scores are for the study foot only.

Table 2. Comparison of foot and leg muscle strength between people with midfoot pain and osteoarthritis (OA) and asymptomatic controls*

Muscle strength (% weight × height)	Midfoot OA group (n = 52)	Control group (n = 36)	% difference	Effect size (Cohen's *d*)	*P*
Ankle plantarflexion	141 ± 48†	192 ± 61	26	0.9	<0.001
Ankle dorsiflexion	88 ± 33†	109 ± 30	19	0.6	<0.001
Ankle inversion	62 ± 22‡	89 ± 28	30	1.1	<0.001
Ankle eversion	67 ± 20†	90 ± 27	26	1.0	<0.001
Lesser toes plantarflexion	62 ± 20†	79 ± 23	22	0.8	<0.001
Hallux plantarflexion	62 ± 19†	85 ± 26	27	1.0	<0.001

* Values are the mean ± SD unless indicated otherwise. Percent difference is calculated relative to asymptomatic controls.
† n = 50, as 2 participants limited by pain on movement.
‡ n = 48, as 4 participants limited by pain on movement.

5.0 ± 2.6). Most participants reported dorsally located midfoot pain (81%) compared to plantar midfoot pain (19%). Thirty participants with midfoot OA had unilateral midfoot pain (57%), and 43% had bilateral midfoot pain. Radiographic OA was most commonly present in the second CMJ (73%), followed by the NCJ (40%), first CMJ (35%), and TNJ (21%). Intrarater reliability of radiographic scoring was almost perfect (percent agreement = 92%; κ_W = 0.92 [95% CI 0.90, 0.95]). People with OA reported poorer OA-specific and general health-related quality of life, as well as a higher level of depressive symptoms, compared to asymptomatic controls. A greater proportion of people with midfoot OA had a pronated foot posture (FPI ≥6) compared to controls, with fewer in the normal and supinated categories (Table 1).

Muscle strength. People with midfoot pain and OA displayed strength deficits in all muscle groups compared to asymptomatic controls (Table 2). The magnitude of weakness ranged from 19% (dorsiflexion) to 30% (inversion), equating to effect sizes of Cohen's *d* coefficient = 0.6 to 1.1 (Figure 1). Except for ankle dorsiflexion, differences existed regardless of whether people with midfoot OA were compared to controls with grade 0 for JSN or osteophytes (n = 19) or with those with grade >0 (n = 17) (*P* < 0.001–0.079) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24182/abstract).

Relationship between muscle strength, pain, and function. In bivariate analyses, muscle strength was negatively correlated with pain and foot-related disability for all muscle groups (Table 3) except for hallux plantarflexion strength and MFPDI pain. Ankle invertor muscle strength was most strongly associated with both MFPDI pain (*r* = –0.320, *P* = 0.027) (Figure 2) and MFPDI function (*r* = –0.349, *P* = 0.015).

Multivariable associations between invertor muscle strength, MFPDI pain, and MFPDI function are presented in Table 4. Multivariable regression analysis revealed that ankle invertor muscle strength was independently associated with foot pain (β = –0.026 [95% CI –0.051, –0.001]; *P* = 0.045) after adjusting for age, sex, BMI, radiograph severity, and depressive symptoms (Table 4). Depressive symptoms were positively associated with pain (β = 0.127 [95% CI 0.004, 0.251]; *P* = 0.044). Invertor muscle strength was also negatively associated with foot-related disability (β = –0.023 [95% CI –0.063, 0.017]; *P* = 0.250) (Table 4), although not after adjusting for the HADS depression domain, which was positively associated with foot-related disability (β = 0.286

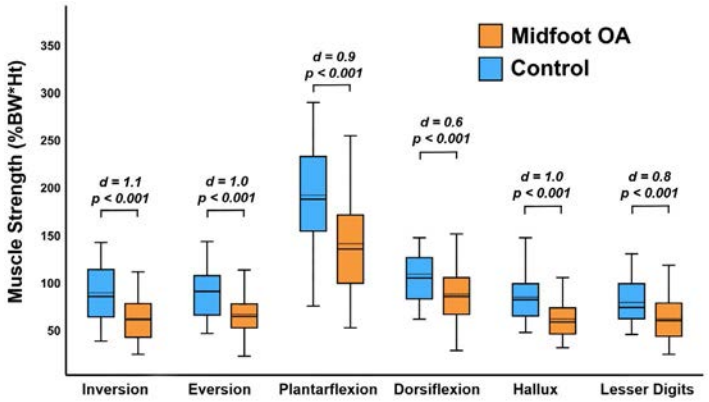

Figure 1. Box plot showing the muscle strength (% body weight [BW] × height [Ht]) for foot and leg muscle groups for midfoot osteoarthritis (OA) and control participants. Horizontal lines and error bars show the median and interquartile range. Dotted lines indicate mean value with corresponding effect size (Cohen's *d* coefficient) and *P* for differences in mean between groups for each variable.

Table 3. Univariate relationships between foot and leg isometric muscle strength variables and Manchester Foot Pain and Disability Index (MFPDI) pain and function subscales

Muscle strength (% body weight × height)	MFPDI pain		MFPDI function	
	r	*P*	*r*	*P*
Ankle plantarflexion*	−0.034	0.813	−0.221	0.122
Ankle dorsiflexion*	−0.176	0.222	−0.155	0.282
Ankle inversion†	−0.320	0.027	−0.349	0.015‡
Ankle eversion*	−0.178	0.216	−0.303	0.033‡
Lesser toes plantarflexion*	−0.279	0.053	−0.346	0.015‡
Hallux plantarflexion*	0.043	0.767	−0.125	0.387

* n = 50, as 2 participants limited by pain on movement.
† n = 48, as 4 participants limited by pain on movement.
‡ $P < 0.05$.

[95% CI 0.092, 0.480]; $P = 0.005$). As lesser toe plantarflexion and ankle eversion strength were significantly associated with foot-related disability in univariate analyses (Table 3), we substituted these variables in the multivariable analyses, and they were found not to be associated with pain or foot-related disability (data not shown). The total variance explained by the independent variables of age, sex, BMI, radiograph severity, invertor strength, and depressive symptoms was 14% for foot pain and 29% for foot-related disability.

DISCUSSION

In this study, we aimed to compare foot and leg muscle strength in people with symptomatic midfoot OA and healthy controls, and to determine whether muscle strength was associated with self-reported pain and foot-related disability. Our primary hypothesis was confirmed; we found that muscle strength was impaired in all muscle groups by 19% to 30% in people with midfoot OA. Our secondary hypothesis that muscle strength would be cross-sectionally associated with foot pain and foot-related disability was partially supported. Invertor muscle strength was independently associated with pain after adjustment for covariates. Although invertor strength was negatively associated with foot-related disability, this association was not statistically significant after adjustment for depressive symptoms.

Muscle weakness has been identified as a clinical feature of OA at other lower extremity joints, including the hip (10), knee (12), and hand (13). This is the first published study to investigate muscle strength in midfoot OA. Reductions in maximal isometric strength were observed across all foot and leg muscle groups in people with midfoot OA, with the largest differences in the ankle invertor group. This may be expected, given that radiographic midfoot OA was present in the joints along the medial arch (TNJ, NCJ, first and second CMJ) (1), where the tibialis posterior muscle (a primary hindfoot invertor) attaches to the adjacent tarsal bones and metatarsals. Deficits in intrinsic foot muscle strength were observed, including muscles that flex the lesser toes and hallux, which are responsible for stiffening the metatarsophalangeal joints to facilitate push-off during walking (41). We did not objectively quantify physical performance, but these findings suggest that muscle weakness in the foot and leg may partially explain deficits in functional ability seen in people with midfoot OA, such as difficulty walking and descending stairs (3,4).

Evidence from longitudinal studies undertaken in knee OA suggests that reduced knee extensor muscle strength is associated with incident tibiofemoral OA (18) and increased risk of symptomatic and functional decline, particularly in women (11). Although it is plausible that muscle weakness plays a role in the pathogenesis of midfoot OA, there are other factors that require consideration. For example, in hand OA, the relationship between grip strength and incident radiographic OA differed by site (metacarpal, proximal, and distal interphalangeal joints), and higher grip strength was associated with incident disease in men but not women (42). The impact of muscle weakness on structural disease in midfoot OA may also be joint specific, and the impact on prognosis may differ according to the site of foot OA, which tends to cluster in the midfoot and first metatarsophalangeal joint (3). The interaction of muscle strength and malalignment may also be important, particularly as people with midfoot OA have flatter feet than asymptomatic controls (7). As this study was cross-sectional, we were not able to determine the temporal nature of the relationship between muscle weakness and midfoot OA. Future prospective longitudinal studies in foot OA would be beneficial to clarify the nature and strength of these relationships.

The relationship between muscle weakness, foot pain, and self-reported function is complex and multifactorial. In this study, the models only explained a modest amount of variance in pain

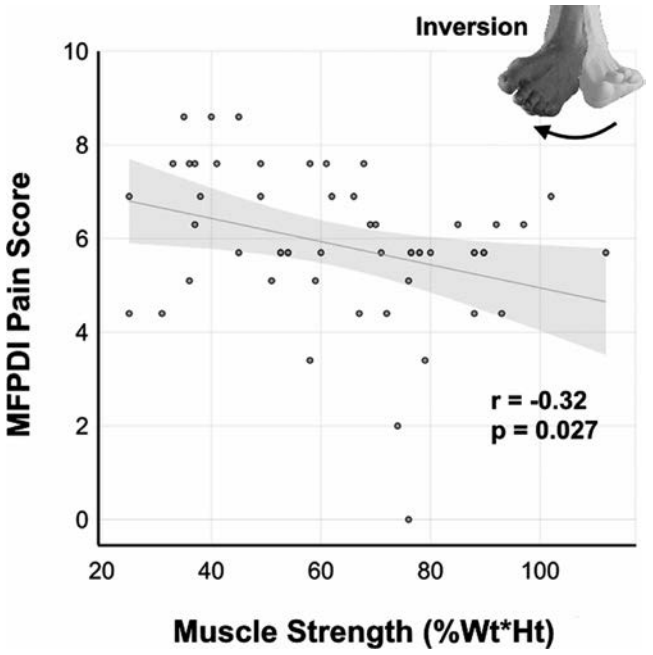

Figure 2. Relationship between invertor muscle strength (% weight [Wt] × height [Ht]) and Manchester Foot Pain and Disability Index (MFPDI) pain in people with midfoot osteoarthritis. Circles represent individual participants in the midfoot OA group.

Table 4. Relationship between invertor muscle strength and Manchester Foot Pain and Disability Index pain and function (outcomes)*

Variable	Unstandardized coefficients β	SE	95% CI	*P*
Model 1: pain				
Age	−0.012	0.027	−0.066, 0.042	0.660
Sex	0.234	0.566	−0.910, 1.378	0.682
BMI	−0.049	0.055	−0.159, 0.061	0.374
Invertor strength	−0.026	0.013	−0.051, −0.001	0.045†
Radiograph severity	0.070	0.044	−0.018, 0.159	0.117
Depressive symptoms	0.127	0.061	0.004, 0.251	0.044*
Model 2: function				
Age	0.074	0.042	−0.010, 0.159	0.081
Sex	1.030	0.886	−0.759, 2.820	0.252
BMI	0.077	0.085	−0.095, 0.250	0.370
Invertor strength	−0.023	0.020	−0.063, 0.017	0.250
Radiograph severity	−0.041	0.069	−0.180, 0.098	0.555
Depressive symptoms	0.286	0.096	0.092, 0.480	0.005*

* 95% CI = 95% confidence interval; BMI = body mass index.
† $P < 0.05$.

(14%) and foot-related disability (29%), with radiographic OA score not independently associated with either outcome. We also determined that deficits in muscle strength in people with midfoot OA existed compared to controls regardless of whether the controls demonstrated minor incidental radiographic features of midfoot OA. This suggests that pain, rather than established radiographic features, likely explained the differences in this model. These results are consistent with other sites of small-joint OA, such as the hand, where radiographic OA explains only a small amount of variance of hand pain and physical function (43,44), with pain mediating the relationship between radiographic disease and self-reported function and strength (43). Studies using magnetic resonance imaging (MRI) and ultrasound-detected OA features, which indicate OA disease activity, have revealed stronger associations between bone marrow lesions and synovitis with pain and function (45,46). Relationships between MRI-detected features of foot OA and symptoms are yet to be explored but offer an opportunity to focus on earlier disease.

Strengthening exercises are associated with moderate improvements in pain, function, physical performance, and small improvements in quality of life compared to usual care in people with hip and knee OA (47). Exercises for hand OA promoting strengthening and joint stability have shown small beneficial effects on pain, function, and joint stiffness, with few adverse events, although overall the quality of evidence is low (48). Given existing knowledge of the role of exercise for people with OA in other joints, further studies appear warranted to investigate whether muscle strengthening is a feasible and effective method to decrease pain and improve function in people with midfoot OA. Importantly, person-level psychosocial factors also influence the report of symptoms, with numerous studies identifying poorer psychological well-being to be associated with the development (49) and severity (32) of persistent foot pain. Our study in midfoot OA was able to examine the influence of depressive symptoms on pain and function and found independent associations for both outcomes, underscoring the importance of psychosocial factors in foot OA.

There are limitations to this study. Muscle strength assessment was conducted with the examiner aware of the participant's clinical status. Blinding is difficult, if not impossible, to achieve when participants have OA involving pain and deformity. To mitigate this difficulty, we used standardized, reliable protocols to obtain maximum force output from participants during testing. To be included in the control group, participants had to show grade 0 or 1 changes for JSN or osteophytes in all the midfoot joints. While this is the usual accepted criteria for absence of OA, 17 participants had grade 1 for either JSN or osteophytes in the midfoot. These participants were confirmed, however, to have no foot symptoms or history of foot injury likely to predispose them to OA. The sample size for the control group was chosen to be as large as could be practically achieved within the time and resource constraints, and consequently the control group included fewer participants than the OA group. Although we adjusted for important confounders of foot pain in the multivariable analyses, the number of participants with midfoot OA also limited the number of independent variables included in the models. Given that the amount of explained variance in pain and function in the multivariable analysis was low, other factors not investigated are likely to be associated with foot pain and foot-related disability in midfoot OA.

In conclusion, people with symptomatic midfoot OA demonstrate weakness in the foot and leg muscles compared to asymptomatic controls. In those with midfoot pain and OA, ankle invertor muscle strength was independently and negatively associated with pain after adjusting for covariates. Ankle invertor strength was also associated with foot-related disability, however, not after adjusting for depressive symptoms. Longitudinal studies are required to establish whether foot and

leg muscle weakness has implications for structural and symptomatic decline. Strengthening of the foot and leg muscles may offer potential to reduce pain and improve function in people with midfoot OA.

ACKNOWLEDGMENTS

The authors thank the radiographers in the Department of Radiology, Chapel Allerton Hospital for their assistance with taking radiographs and all volunteers who participated.

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. Arnold 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. Arnold, Halstead, Hill, Redmond.
Acquisition of data. Arnold, Halstead.
Analysis and interpretation of data. Arnold, Halstead, Grainger, Keenan, Hill, Redmond.

REFERENCES

1. 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: cross-sectional findings from the Clinical Assessment Study of the Foot. Ann Rheum Dis 2015;74:156–63.
2. Thomas MJ, Peat G, Rathod T, Marshall M, Moore A, Menz HB, et al. The epidemiology of symptomatic midfoot osteoarthritis in community-dwelling older adults: cross-sectional findings from the Clinical Assessment Study of the Foot. Arthritis Res Ther 2015;17.
3. Rathod T, Marshall M, Thomas MJ, Menz HB, Myers HL, Thomas E, et al. Investigations of potential phenotypes of foot osteoarthritis: cross-sectional analysis from the Clinical Assessment Study of the Foot. Arthritis Care Res 2016;68:217–27.
4. Rao S, Baumhauer JF, Tome J, Nawoczenski DA. Comparison of in vivo segmental foot motion during walking and step descent in patients with midfoot arthritis and matched asymptomatic control subjects. J Biomech 2009;42:1054–60.
5. Mann RA, Prieskorn D, Sobel M. Mid-tarsal and tarsometatarsal arthrodesis for primary degenerative osteoarthrosis or osteoarthrosis after trauma. J Bone Joint Surg Am 1996;78:1376–85.
6. Downes TJ, Chesterton L, Whittle R, Roddy E, Menz HB, Marshall M, et al. Symptomatic course of foot osteoarthritis phenotypes: an 18-month prospective analysis of community-dwelling older adults. Arthritis Care Res (Hoboken) 2018;70:1107–12.
7. Menz HB, Munteanu SE, Zammit GV, Landorf KB. Foot structure and function in older people with radiographic osteoarthritis of the medial midfoot. Osteoarthritis Cartilage 2010;18:317–22.
8. Arnold JB, Marshall M, Thomas MJ, Redmond AC, Menz HB, Roddy E. Midfoot osteoarthritis: potential phenotypes and their associations with demographic, symptomatic and clinical characteristics. Osteoarthritis Cartilage 2019;27:659–66.
9. Rao S, Baumhauer J, Nawoczenski D. Is barefoot regional plantar loading related to self-reported foot pain in patients with midfoot osteoarthritis. Osteoarthritis Cartilage 2011;19:1019–25.
10. Loureiro A, Mills PM, Barrett RS. Muscle weakness in hip osteoarthritis: a systematic review. Arthritis Care Res 2013;65:340–52.
11. Culvenor AG, Ruhdorfer A, Juhl C, Eckstein F, Øiestad BE. Knee extensor strength and risk of structural, symptomatic, and functional decline in knee osteoarthritis: a systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2017;69:649–58.
12. O'Reilly SC, Jones A, Muir KR, Doherty M. Quadriceps weakness in knee osteoarthritis: the effect on pain and disability. Ann Rheum Dis 1998;57:588–94.
13. Dominick KL, Jordan JM, Renner JB, Kraus VB. Relationship of radiographic and clinical variables to pinch and grip strength among individuals with osteoarthritis. Arthritis Rheum 2005;52:1424–30.
14. Muraki S, Akune T, Teraguchi M, Kagotani R, Asai Y, Yoshida M, et al. Quadriceps muscle strength, radiographic knee osteoarthritis and knee pain: the ROAD study. BMC Musculoskelet Disord 2015;16.
15. Hurley MV, Scott DL, Rees J, Newham DJ. Sensorimotor changes and functional performance in patients with knee osteoarthritis. Ann Rheum Dis 1997;56:641–8.
16. McAlindon T, Cooper C, Kirwan J, Dieppe P. Determinants of disability in osteoarthritis of the knee. Ann Rheum Dis 1993;5258–62.
17. Palmieri-Smith RM, Thomas AC, Karvonen-Gutierrez C, Sowers MF. Isometric quadriceps strength in women with mild, moderate, and severe knee osteoarthritis. Am J Phys Med Rehabil 2010;89:541–8.
18. Øiestad B, Juhl C, Eitzen I, Thorlund J. Knee extensor muscle weakness is a risk factor for development of knee osteoarthritis. A systematic review and meta-analysis. Osteoarthritis Cartilage 2015;23:171–7.
19. Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis 2013;72:1125–35.
20. Kloppenburg M, Kroon FP, Blanco FJ, Doherty M, Dziedzic KS, Greibrokk E, et al. 2018 update of the EULAR recommendations for the management of hand osteoarthritis. Ann Rheum Dis 2018;78:16–24.
21. 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.
22. Munteanu SE, Zammit GV, Menz HB. Factors associated with foot pain severity and foot-related disability in individuals with first metatarsophalangeal joint OA. Rheumatology (Oxford) 2011;51:176–183.
23. Bartholdy C, Juhl C, Christensen R, Lund H, Zhang W, Henriksen M. The role of muscle strengthening in exercise therapy for knee osteoarthritis: a systematic review and meta-regression analysis of randomized trials. Semin Arthritis Rheum 2017;47:9–21.
24. Paskins Z, Sanders T, Hassell AB. What influences patients with osteoarthritis to consult their GP about their symptoms? A narrative review. BMC Fam Pract 2013;14:195.
25. Chatterton BD, Muller S, Thomas MJ, Menz HB, Rome K, Roddy E. Inter and intra-rater repeatability of the scoring of foot pain drawings. J Foot Ankle Res 2013;6:44.
26. Garrow AP, Silman AJ, Macfarlane GJ. The Cheshire Foot Pain and Disability Survey: a population survey assessing prevalence and associations. Pain 2004;110:378–84.
27. 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.
28. Spink MJ, Fotoohabadi MR, Menz HB. Foot and ankle strength assessment using hand-held dynamometry: reliability and age-related differences. Gerontology 2010;56:525–32.
29. Wang CY, Olson SL, Protas EJ. Test-retest strength reliability: hand-held dynamometry in community-dwelling elderly fallers. Arch Phys Med Rehabil 2002;83:811–5.
30. Thorborg K, Bandholm T, Hölmich P. Hip-and knee-strength assessments using a hand-held dynamometer with external belt-fixation

are inter-tester reliable. Knee Surg Sports Traumatol Arthrosc 2013; 21:550–5.

31. Garrow AP, Papageorgiou AC, Silman AJ, Thomas E, Jayson MI, Macfarlane GJ. Development and validation of a questionnaire to assess disabling foot pain. Pain 2000;85:107–13.
32. Awale A, Dufour AB, Katz P, Menz HB, Hannan MT. Link between foot pain severity and prevalence of depressive symptoms. Arthritis Care Res (Hoboken) 2016;68:871–6.
33. Butterworth PA, Urquhart DM, Cicuttini FM, Menz HB, Strauss BJ, Proietto J, et al. Relationship between mental health and foot pain. Arthritis Care Res (Hoboken) 2014;66:1241–5.
34. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–70.
35. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. J Psychosom Res 2002;52:69–77.
36. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20:1727–36.
37. Keenan AM, Mckenna SP, Doward LC, Conaghan PG, Emery P, Tennant A. Development and validation of a needs-based quality of life instrument for osteoarthritis. Arthritis Rheum 2008;59:841–8.
38. Redmond AC, Crosbie J, Ouvrier RA. Development and validation of a novel rating system for scoring standing foot posture: the Foot Posture Index. Clin Biomech (Bristol, Avon) 2006;21: 89–98.
39. Cornwall MW, McPoil TG, Lebec M, Vicenzino B, Wilson J. Reliability of the modified foot posture index. J Am Podiatr Med Assoc 2008;98:7–13.
40. Redmond AC, Crane YZ, Menz HB. Normative values for the foot posture index. J Foot Ankle Res 2008;1:6.
41. Farris DJ, Kelly LA, Cresswell AG, Lichtwark GA. The functional importance of human foot muscles for bipedal locomotion. Proc Natl Acad Sci 2019;116:1645–50.
42. Chaisson CE, Zhang Y, Sharma L, Kannel W, Felson DT. Grip strength and the risk of developing radiographic hand osteoarthritis: results from the Framingham Study. Arthritis Rheum 1999;42:33–8.
43. Jones G, Cooley H, Bellamy N. A cross-sectional study of the association between Heberden's nodes, radiographic osteoarthritis of the hands, grip strength, disability and pain. Osteoarthritis Cartilage 2001;9:606–11.
44. Haugen IK, Slatkowsky-Christensen B, Bøyesen P, van der Heijde D, Kvien TK. Cross-sectional and longitudinal associations between radiographic features and measures of pain and physical function in hand osteoarthritis. Osteoarthritis Cartilage 2013;21:1191–8.
45. Haugen IK, Bøyesen P, Slatkowsky-Christensen B, Sesseng S, van der Heijde D, Kvien TK. Associations between MRI-defined synovitis, bone marrow lesions and structural features and measures of pain and physical function in hand osteoarthritis. Ann Rheum Dis 2012;71:899–904.
46. Haugen IK, Slatkowsky-Christensen B, Bøyesen P, Sesseng S, van der Heijde D, Kvien TK. MRI findings predict radiographic progression and development of erosions in hand osteoarthritis. Ann Rheum Dis 2016;75:117–23.
47. Goh SL, Persson MS, Stocks J, Hou Y, Welton NJ, Lin J, et al. Relative efficacy of different exercises for pain, function, performance and quality of life in knee and hip osteoarthritis: systematic review and network meta-analysis. Sports Medicine 2019;49:743–61.
48. Østerås N, Kjeken I, Smedslund G, Moe RH, Slatkowsky-Christensen B, Uhlig T, et al. Exercise for hand osteoarthritis: a Cochrane systematic review. J Rheumatol 2017;44:1850–8.
49. Gill TK, Menz HB, Landorf KB, Arnold JB, Taylor AW, Hill CL. Predictors of foot pain in the community: the North West Adelaide health study. J Foot Ankle Res 2016;9:23.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 781–787
DOI 10.1002/acr.24371

American College of Rheumatology
Empowering Rheumatology Professionals

Effects of Vitamin D Supplementation on Disabling Foot Pain in Patients With Symptomatic Knee Osteoarthritis

Liudan Tu,[1] Shuang Zheng,[2] Flavia Cicuttini,[3] Xingzhong Jin,[4] Weiyu Han,[5] Zhaohua Zhu,[5] Benny Antony,[2] Tania Winzenberg,[2] Graeme Jones,[2] Jieruo Gu,[6] Anita E. Wluka,[3] and Changhai Ding[7]

Objective. The present study was undertaken to determine whether vitamin D supplementation or maintaining sufficient vitamin D level reduces foot pain over 2 years in patients with symptomatic knee osteoarthritis (OA).

Methods. A post hoc study was conducted from a randomized, double-blind, placebo-controlled trial named the Vitamin D Effect on Osteoarthritis (VIDEO) study. Symptomatic knee OA patients with serum 25-hydroxyvitamin D levels between 12.5 nmoles/liter and 60 nmoles/liter were included and randomly allocated to either monthly vitamin D_3 or placebo treatment (1:1) for 2 years. Manchester Foot Pain and Disability Index (MFPDI) was used to evaluate foot pain and disabling foot pain was defined as at least 1 of the 10 functional limitation items (items 1–9 and 11) being documented as on "most/every day(s)" in the last month. A repeated-measures, mixed-effects model was used to analyze the change of MFPDI scores between groups adjusting for potential confounders.

Results. A total of 413 patients with a mean age of 63.2 years (49.7% males) were enrolled and 340 completed the study. The mean MFPDI score was 22.8 ± 7.3, with 23.7% of participants having disabling foot pain at baseline. There were significant differences in MFPDI scores change between groups over 2 years, with more improvements in the vitamin D group than in the placebo group (–0.03 versus 1.30; $P = 0.013$) and more improvement in those maintaining sufficient vitamin D levels (n = 226) than those who did not (n = 114) (–0.09 versus 2.19; $P = 0.001$).

Conclusion. Vitamin D supplementation and maintenance of sufficient vitamin D levels may improve foot pain in those with knee OA.

INTRODUCTION

Osteoarthritis (OA) is a chronic disease worldwide and characterized by joint pain and deformity. In those >60 years of age, the global prevalence of OA is ~10% in men and 20% in women; the financial burden is estimated as high as 1.0–2.5% of the gross domestic product in Western countries (1). Foot pain, a common musculoskeletal pain, often defined as pain in the foot and/or ankle (2), affects nearly 1 in 5 older individuals in the community (3–7) and has a detrimental impact on health-related quality of life. Foot pain often coexists with knee pain, and concurrent foot pain leads to impaired physical activity, lower quality of life, and increased levels of depression in patients with knee OA when compared with the general population (4,5). In addition, in a survey of 8,990 older individuals, most patients with knee pain also had pain in multiple joints, and the severity of knee pain and related disability was worse in the presence of pain elsewhere (6). Given that patients with knee OA are more likely to have foot pain and increased severity of foot pain, management of foot pain in patients with OA is of priority.

Vitamin D deficiency and insufficiency, which are generally defined as a serum 25-hydroxyvitamin D concentration of <50

ClinicalTrials.gov identifier: NCT01176344.

Supported by the Australian National Health and Medical Research Council (project code 605501). Drs. Winzenberg and Wluka are recipients of a National Health and Medical Research Council Career Development Fellowship. Dr. Jones is recipient of a National Health and Medical Research Council Practitioner Fellowship. Dr. Ding is recipient of an Australian Research Council Future Fellowship.

[1]Liudan Tu, MD: The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, and University of Tasmania, Hobart, Tasmania, Australia; [2]Shuang Zheng, PhD, Benny Antony, PhD, Tania Winzenberg, PhD, Graeme Jones, MD: University of Tasmania, Hobart, Tasmania, Australia; [3]Flavia Cicuttini, PhD, Anita E. Wluka, PhD: Monash University, Melbourne, Victoria, Australia; [4]Xingzhong Jin, PhD: University of Tasmania, Hobart, Tasmania, and the University of Sydney, New South Wales, Australia; [5]Weiyu Han, PhD, Zhaohua Zhu, PhD: University of Tasmania, Hobart, Tasmania, Australia, and Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, China; [6]Jieruo Gu, MD: The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China; [7]Changhai Ding, MD: The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, University of Tasmania, Hobart, Tasmania, Australia, Monash University, Melbourne, Victoria, Australia, and Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, China.

Drs. Tu and Zheng contributed equally to this work.

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

Address correspondence to Changhai Ding, MD, Clinical Research Centre, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, China. Email: Changhai.Ding@utas.edu.au.

Submitted for publication November 25, 2019; accepted in revised form June 30, 2020.

SIGNIFICANCE & INNOVATIONS

- Foot pain is common in patients with symptomatic osteoarthritis (OA), and vitamin D deficiency is related to chronic pain.
- This study suggests that vitamin D supplementation may be beneficial for foot pain in patients with OA.

nmoles/liter and between 50 and 75 nmoles/liter, respectively (8), are common all around the world, presenting in ~45% of adults in Australia (9). Previous epidemiologic studies have reported that vitamin D deficiency is associated with chronic musculoskeletal pain and depression, but underlying mechanisms are complex and unclear (10–12). Studies exploring the effect of vitamin D supplementation on nonspecific chronic pain in the adult population and in those with rheumatoid arthritis and osteoporosis have had inconsistent findings (12,13). In addition, the efficacy of vitamin D supplementation on knee pain in patients with knee OA is inconsistent between the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain score and the visual analog scale (VAS) knee pain score (14,15). To date, there have been no studies exploring the effect of vitamin D supplementation on foot pain in patients with knee OA. Therefore, the aim of our study was to explore whether vitamin D supplementation or maintaining sufficient vitamin D level reduces foot pain in patients with symptomatic knee OA initially deficient in vitamin D.

PATIENTS AND METHODS

Patients and trial design. A post hoc study was conducted based on a randomized, double-blind, placebo-controlled trial named the Vitamin D Effect on Osteoarthritis (VIDEO) Study (ClinicalTrials.gov identifier: NCT01176344), in which the primary outcomes were tibial cartilage volume and knee pain among patients with symptomatic knee OA (16). Patients who suffered from symptomatic knee OA for at least 6 months, who had knee pain of 20–80 mm on a 100-mm VAS and serum 25-hydroxyvitamin D levels between 12.5 nmoles/liter and 60 nmoles/liter, and who were ages 50–79 years were included. Patients with grade 3 radiographic changes on the Altman and Gold atlas, severe knee pain on standing (>80 mm on a 100-mm VAS), other rheumatic diseases, such as rheumatoid arthritis, psoriatic arthritis, and lupus, contraindication to magnetic resonance imaging (MRI), cancer, severe cardiac or renal impairment, hypersensitivity to vitamin D, anticipated knee or hip surgery within the next 2 years, and a history of taking vitamin D within the previous 1 month were excluded from this study. After signing the written consent form, patients were randomly allocated to 24 months of vitamin D or placebo treatment at a ratio of 1:1. A monthly capsule containing 50,000 international units (IU) (1.25 mg) of vitamin D_3 (cholecalciferol) or placebo was given to patients, and assessments were conducted at baseline and at months 3, 6, 12, and 24. This study was approved by the Ethics Committees in Tasmania and Melbourne (reference numbers H1040 and CF10/1182-2010000616, respectively).

Measurements. *Manchester Foot Pain and Disability Index (MFPDI) questionnaires.* The MFPDI questionnaire was used to measure foot pain of patients at months 0, 3, 6, 12, and 24. The MFPDI was developed to measure foot pain and disability in the elderly (17). It is a proven, useful, and valid instrument for assessing foot pain in the older population and has been used in both observational studies and randomized controlled trials (18). Each item was scored either 1 (none of the time), 2 (on some days), or 3 (on most days/every day). The total score was calculated by summing the scores of 17 items, with a possible score range of 17 to 54. A higher score indicates greater disability. Disabling foot pain was defined when at least 1 of the 10 functional limitation items (items 1–9 and 11) was documented as "most/every day(s)" in the last month (18).

Knee structure measurements. Radiographic OA was assessed at baseline by an anteroposterior radiograph of the patient in a standing, semiflexed position, as per the Altman atlas (19), according to the protocol of the Osteoarthritis Research Society International atlas for scoring osteophytes and joint space narrowing. MRI scans with a commercial transmit–receive extremity coil at baseline and 2 years of the study knee were obtained according to standardized protocol. T2-weighted/proton density–weighted, fast spin-echo sequences were used to assess cartilage defects and bone marrow lesions; details on the protocol have been described previously (20).

WOMAC. The WOMAC scale (21) was used to detect knee symptoms at months 0, 3, 6, 12, and 24. The sum of pain (range 0–500), stiffness (range 0–200), and physical function (range 0–1,700) subscales was calculated as the total WOMAC score (range 0–2,400).

Serum 25-hydroxyvitamin D measurement and definition of vitamin D status. Serum 25-hydroxyvitamin D levels were assayed using direct competitive chemiluminescent immunoassays at screening, month 3, and month 24. Patients whose serum 25-hydroxyvitmin D level was >50 nmoles/liter at both month 3 and month 24 were classified into the group that maintained sufficient vitamin D, and those whose serum 25-hydroxyvitmin D level was <50 nmoles/liter at either month 3 or month 24 were classified into the group that did not maintain sufficient vitamin D.

Physical activity. Physical activity was assessed using the International Physical Activity Questionnaire short version, which has been proven to be valid and reliable in monitoring population levels of physical activity among older adults in diverse settings (22). Based on the scoring protocol, we classified physical activity status as insufficiently active, sufficiently active, and highly active.

Other measurements. Body height and weight were measured at baseline, and body mass index (BMI; kg/m^2) was calculated. Height was measured to the nearest 0.1 cm (with shoes removed) using a Leicester Height Measure stadiometer (Invicta

Plastics), and weight was measured to the nearest 0.1 kg (with shoes and bulky clothing removed) using Heine S-7307 electronic scales (Heine).

Statistical analysis. Baseline characteristics between patients with and without disabling foot pain were compared using independent *t*-tests or chi-square tests. A repeated-measures mixed-effects model with terms for treatment, time, and treatment by time and adjustment for age, sex, and BMI was used to analyze the change in MFPDI scores over 24 months for the group including vitamin D versus the placebo group and the group that maintained sufficient vitamin D versus the group that did not maintain sufficient vitamin D. Multilevel mixed-effects models were used to deal with missing data caused by loss to follow-up and nonresponses. A subgroup analysis exploring the effects of vitamin D supplementation and maintaining sufficient vitamin D levels on the relief of foot pain in patients with disabling foot pain was performed at baseline. All tests were 2-sided, and *P* values less than 0.05 were considered significant. Stata, version 12.0, was used to perform statistical analyses.

RESULTS

Baseline characteristics. A total of 413 patients were included and randomized to receive either vitamin D (n = 209) or placebo (n = 204) treatment. After 24 months, 340 patients completed the trial; there were no significant differences in baseline characteristics between patients who completed the study and those who did not. The average age of patients was 63.2 years, with a mean BMI of 29.6 kg/m^2; 49.7% of them were female. The mean ± SD MFPDI score was 22.8 ± 7.3. For those reporting foot pain at baseline (n = 214), 74 patients (34.6%) reported pain in the toes, 49 (22.9%) reported pain in the ball of foot, 48 (22.4%) reported pain in the arch, 43 (20%) reported pain in the whole foot, and 37 (17.3%) reported pain in the heel. Disabling foot pain was present in 23.7% (n = 98) of patients according to MFPDI case definition. There were also no significant differences in baseline characteristics, MFPDI scores, prevalence of disabling foot pain, knee symptoms, and knee structure measurements between the vitamin D group and the placebo group (Table 1). At the same time, the baseline serum vitamin D level was higher in the group that maintained sufficient vitamin D when compared with the group that did not maintain sufficient vitamin D (45.2 versus 41.5; *P* = 0.01), while other characteristics were similar for the 2 groups.

Vitamin D supplementation and change in MFPDI scores. Over 24 months, MFPDI scores remained largely unchanged in the vitamin D group (–0.03 [95% confidence interval (95% CI) –0.80, 0.74]) while worsening in the placebo group (1.30 [95% CI 0.51, 2.09]) (Figure 1). There were significant differences in change of MFPDI scores between groups in the mixed-effects model after including all time points adjusted for age, sex, and BMI (between-group difference –1.32 [95% CI –2.43, –0.22], *P* = 0.013) (Table 2 and Figure 1).

In subgroup analyses, for patients with disabling foot pain at baseline, although those who received vitamin D treatment had lower changes in MFPDI scores compared with the placebo group after 24 months, the difference was not statistically

Table 1. Baseline characteristics of participants between the group that maintained sufficient vitamin D and the group that had insufficient vitamin D*

	Vitamin D group (n = 209)	Placebo group (n = 204)	*P*
Age, years	63.55 ± 6.88	62.85 ± 7.22	0.32
Female, no. (%)	106 (50.72)	102 (50)	0.92
Body mass index	29.57 ± 5.39	29.64 ± 4.62	0.88
Baseline 25(OH)D level, nmoles/liter	43.74 ± 11.79	43.81 ± 12.66	0.95
MFPDI score (range 0–34)	21.85 ± 6.83	22.66 ± 7.49	0.27
Disabling foot pain, no. (%)	101 (47.42)	112 (52.58)	0.30
Physical activity, no. (%)			0.71
Insufficiently active	38 (19.0)	40 (20.83)	
Sufficiently active	82 (41.0)	71 (36.98)	
Highly active	80 (40)	81 (42.19)	
WOMAC score			
Pain (range 0–500)	137.88 ± 88.82	134.74 ± 83.42	0.71
Function (range 0–1700)	487.94 ± 318.14	467.59 ± 292.79	0.50
Stiffness (range 0–200)	61.48 ± 41.53	61.74 ± 40.08	0.95
Knee structure measurement			
Total radiographic OA (range 0–18)	8.26 ± 5.56	8.31 ± 4.91	0.93
Total cartilage defects (range 0–24)	14.84 ± 4.08	14.42 ± 3.94	0.29
Total bone marrow lesions (range 0–45)	3.15 ± 3.22	3.59 ± 3.23	0.17

* Values are the mean ± SD unless indicated otherwise. Two-tailed Student's *t*-tests were used for differences between means. Chi-square tests were used for proportions (percentages), and Wilcoxon's rank sum tests were used for differences between medians. 25(OH)D = 25-hydroxyvitamin D; MFPDI = Manchester Foot Pain and Disability Index; OA = osteoarthritis; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

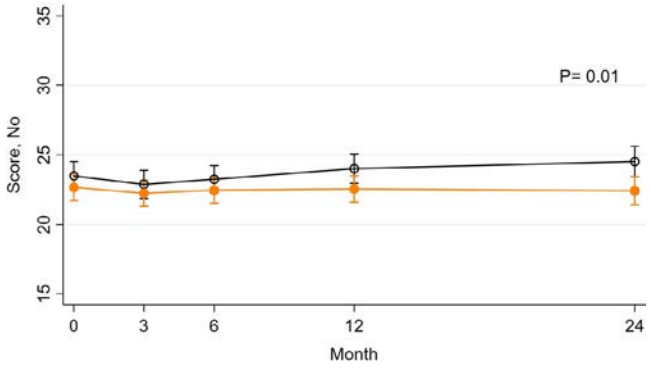

Figure 1. Change in total Manchester Foot Pain and Disability Index scores (range 0–34) in the vitamin D supplementation group (orange) and the placebo group (black).

significant (a change of –4.86 [95% CI –6.79, –2.93] in the vitamin D group versus –2.30 [95% CI –5.33, –0.31] in the placebo group; between-group difference –2.56 [95% CI –5.33, 0.21], *P* = 0.07). In patients without disabling foot pain at baseline, the between-group difference was smaller and also not statistically significant (a change of 1.36 [95% CI 0.59, 2.14] in the vitamin D group versus 2.46 [95% CI 1.66, 3.25] in the placebo group; between-group difference –1.09 [95% CI –2.20, 0.02], *P* = 0.05) (Table 2). For female patients, there was no significant difference in change of MFPDI scores between the vitamin D group and the placebo group. In male patients, vitamin D supplementation significantly improved MFPDI scores when compared with the placebo group (Table 2). However, there was no significant interaction between sex and vitamin D supplementation on change in MFPDI scores.

Maintaining sufficient vitamin D levels and change in MFPDI scores. In post hoc analyses comparing patients who maintained vitamin D sufficiency to those who did not, MFPDI scores, over 2 years, decreased in those maintaining vitamin D sufficiency (2.19 [95% CI 1.21, 3.18]) but increased in those who did not (–0.09 [95% CI –0.79, 0.61]); the between-group difference was –2.29 (95% CI –3.49, –1.08; *P* = 0.001) after adjusting for age, sex, BMI, serum 25(OH)D level, and baseline MFPDI score (Table 3 and Figure 2).

In subgroup analyses, for those with disabling foot pain at baseline, there was a greater decrease in MFPDI scores in those who maintained vitamin D sufficiency (–0.14 [95% CI –2.75, 2.48]) compared to those who did not (–4.63 [95% CI –6.35, –2.92]); the between-group difference was –4.49 (95% CI –7.62, –1.37; *P* = 0.005) (Table 3). There were similar findings in patients without disabling foot pain at baseline (Table 3). For female patients, there was no significant difference between maintaining sufficient vitamin D and not maintaining sufficient vitamin D in change of MFPDI scores (Table 3). In male patients, significant improvement of MFPDI scores was found in the group that maintained sufficient vitamin D compared to the group that did not maintain sufficient vitamin D; however, there was no significant interaction between sex and maintaining sufficient vitamin D in change in MFPDI scores.

DISCUSSION

To the best of our knowledge, this study is the first to investigate the effects of supplementing vitamin D and maintaining sufficient vitamin D level on foot pain in patients with symptomatic knee OA. In this sample, 51.8% of participants with knee OA

Table 2. Effect of vitamin D supplementation on change in Manchester Foot Pain and Disability Index (MFPDI) scores over 2 years*

	Mean change (95% CI)	Between-group difference change, mean (95% CI)	*P*
Whole sample		–1.32 (–2.43, –0.22)†	0.013†
Placebo group (n = 204)	1.30 (0.51, 2.09)†		
Vitamin D group (n = 209)	–0.03 (–0.80, 0.74)		
Those without disabling foot pain at baseline		–1.09 (–2.20, 0.02)	0.05
Placebo group (n = 153)	2.46 (1.66, 3.25)†		
Vitamin D group (n = 162)	1.36 (0.59, 2.14)†		
Those with disabling foot pain at baseline		–2.56 (–5.33, 0.21)	0.07
Placebo group (n = 51)	–2.30 (–5.33, –0.31)†		
Vitamin D group (n = 47)	–4.86 (–6.79, –2.93)†		
Female patients		–0.97 (–2.65, 0.70)	0.26
Placebo group (n = 102)	1.19 (–0.06, 2.43)		
Vitamin D group (n = 106)	0.22 (–0.91, 1.34)		
Male patients		–1.59 (–3.05, –0.13)†	0.03†
Placebo group (n = 102)	1.32 (0.31, 2.33)†		
Vitamin D group (n = 103)	–0.27 (–1.33, 0.78)		

* Changes in outcomes are generated from mixed-effects models adjusted for age, sex, and body mass index. Between-group differences were calculated with values from the vitamin D group minus values from the placebo group, or values from the group that maintained sufficient vitamin D minus values from the group that had insufficient vitamin D.
† Significant.

Table 3. Effect of vitamin D status on change in Manchester Foot Pain and Disability Index (MFPDI) scores over 2 years*

	Mean change (95% CI)	Between-group difference change, mean (95% CI)	*P*
Whole sample		−2.29 (−3.49, −1.08)†	0.001†
Not maintaining sufficient vitamin D (n = 114)	2.19 (1.21, 3.18)†		
Maintaining sufficient vitamin D (n = 226)	−0.09 (−0.79, 0.61)		
Those without disabling foot pain at baseline		−1.55 (−2.76, −0.33)†	0.01†
Not maintaining sufficient vitamin D (n = 91)	2.76 (1.77, 3.74)†		
Maintaining sufficient vitamin D (n = 174)	1.21 (0.50, 1.93)†		
Those with disabling foot pain at baseline		−4.49 (−7.62, −1.37)†	0.005†
Not maintaining sufficient vitamin D (n = 23)	−0.14 (−2.75, 2.48)		
Maintaining sufficient vitamin D (n = 52)	−4.63 (−6.35, −2.92)†		
Female patients		−1.21 (−2.98, 0.56)	0.18
Not maintaining sufficient vitamin D (n = 62)	1.52 (0.13, 2.91)†		
Maintaining sufficient vitamin D (n = 98)	0.31 (−0.79, 1.40)		
Male patients		−2.92 (−4.54, −1.31)†	0.001†
Not maintaining sufficient vitamin D (n = 52)	2.81 (1.45, 4.17)†		
Maintaining sufficient vitamin D (n = 128)	−0.11 (−0.99, 0.77)		

* Changes in outcomes are generated from mixed-effects models adjusted for age, sex, body mass index, 25-hydroxyvitamin D level, and baseline MFPDI score. Between-group differences were calculated with values from the vitamin D group minus values from the placebo group, or values from the group that maintained sufficient vitamin D minus values from the group that had insufficient vitamin D.
† Significant.

and vitamin D deficiency reported foot pain. Foot pain and disability scores (using the MFPDI) decreased more over 24 months in the treatment group that maintained sufficient vitamin D than in the placebo group and the group that did not maintain sufficient vitamin D. Our results suggest that foot pain is common and that maintaining sufficient vitamin D levels over 24 months may have beneficial effects on foot pain in patients with knee OA.

Foot pain is a common condition in patients with OA. A recent cross-sectional study using data from the Osteoarthritis Initiative reported that one-fourth of individuals with knee OA experienced concurrent foot pain, with the majority (55%) reporting pain in both feet. Furthermore, patients with knee OA with foot/ankle symptoms reported worse scores on all WOMAC subscales, including

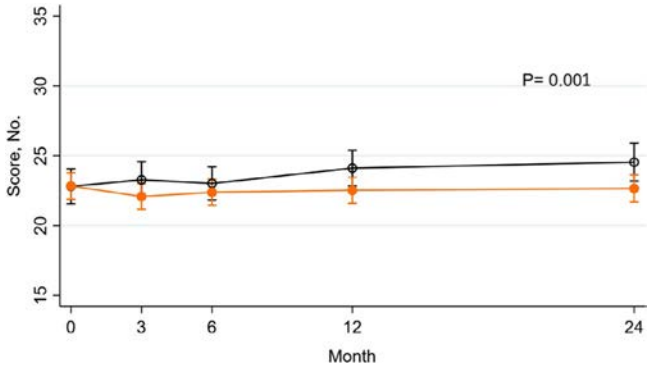

Figure 2. Change in total Manchester Foot Pain and Disability Index scores (range 0–34) in the groups that did maintain (orange) and did not maintain (black) vitamin D sufficiency. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24371/abstract.

the total score, worse health outcomes, and poorer physical function compared with those without foot/ankle symptoms (5). In our study, more than one-half of the patients (51.8%) reported foot pain at baseline, and patients with foot pain had lower quality of life and higher rates of depression, which is similar to results of previous studies. In another study, foot/ankle symptoms in either or both feet significantly increased the odds of developing knee symptoms and symptomatic radiographic knee OA in individuals at risk of the disease (23). Additionally, in patients with symptomatic radiographic knee OA, the presence of foot/ankle symptoms was associated with increased risk of knee pain over 4 years (24). Owing to the coexistent relationships between foot/ankle symptoms and knee OA, more attention should be paid to management of foot pain in patients with OA and in those at risk of knee OA.

Although there is a growing body of evidence suggesting that a low level of vitamin D is associated with chronic pain, no clinical study has been conducted to explore the effect of vitamin D supplementation on foot pain. Furthermore, studies examining whether vitamin D supplementation is beneficial on other musculoskeletal pains are limited and have found conflicting results (10,12,13,25,26), mainly due to variations in participants, outcome measures, sample size, vitamin D dosage, and follow-up time. A recent secondary analysis of a randomized controlled trial with a large sample suggested that long-term, monthly supplementation of 100,000 IU of vitamin D did not improve pain scores or reduce analgesic dispensing in the general population (26). Similarly, a Cochrane review also concluded that a significant beneficial effect of vitamin D on chronic painful conditions across different sites was unlikely (10). However, in the current study, vitamin D supplementation and sufficient vitamin D reduced MFPDI scores

in patients with symptomatic knee OA after 24 months compared with the placebo group, particularly in patients with disabling foot pain at baseline. There are a number of reasons as to why our results vary from other studies: our participants were selected based on low baseline vitamin D levels; we specifically examined foot pain using a validated measure; our vitamin D dosage was 50,000 IU per month, with the potential to improve compliance; and our duration of treatment was 2 years. As reported before in our randomized controlled trial, 62% of patients in the placebo group gained a sufficient 25-hydroxyvitamin D level after 24 months. This may dilute the effect of vitamin D supplementation on MFPDI scores. On the other hand, consistent results were found when patients were classified into the group that maintained sufficient vitamin D and the group that did not maintain sufficient vitamin D. However, further clinical trials will be needed to determine whether vitamin D supplementation is beneficial for foot pain in patients with knee OA.

Previous studies have found that low levels of vitamin D are associated with chronic pain (12,27,28); but there is no previous work linking vitamin D deficiency to foot pain. In one population-based, cross-sectional study of 958 older adults, a lower level of vitamin D was not related to foot pain, but it was related to back pain (29). In contrast, in our longitudinal study, our population with low vitamin D levels that maintained vitamin D sufficiency with supplementation had significantly decreased MFPDI scores compared with those who did not maintain vitamin D sufficiency between months 3 and 24, suggesting that correction of vitamin D deficiency might reduce foot pain over time.

Several potential mechanisms, such as bone demineralization, muscle weakness, and pain dysregulation, may link vitamin D deficiency to musculoskeletal pain. Vitamin D can modulate a number of inflammatory pathways (30), which are associated with pain sensitization. Low vitamin D level can activate proinflammatory cytokine proliferation, thus altering sensitization of peripheral and central pathways through nociceptive inflammation processing (31,32), which may be an important contributor to clinical symptoms of knee OA (33). Even so, the underlying mechanisms between vitamin D deficiency and foot pain are still unclear, and further investigations are needed.

In patients with OA, differences in sex existed in the experience of pain, and psychological factors (34) and other factors, such as foot and ankle shape, footwear habit, obesity, decline in muscle strength with aging, and ligamentous laxity, may underline the differences in sex in regard to pain. In our study, female patients reported higher MFPDI scores and more disabling foot pain at baseline. Although male patients experienced significant improvement in MFPDI scores when treated with vitamin D supplementation and maintaining sufficient vitamin D level, there were no significant interactions between sex and vitamin D supplementation or maintaining sufficient vitamin D on change in MFPDI scores. This suggests that there is no difference in sex in the effects of vitamin D supplementation or maintaining sufficient vitamin D on change in MFPDI scores.

There are several potential limitations to our study. First, this is a post hoc analysis in which foot pain was the secondary outcome in the original protocol. Second, nearly one-half of patients reported some foot pain at baseline, and only 23.7% of patients had disabling foot pain according to the definition we used in this study. Even though the MFPDI score increased less after vitamin D supplementation in those without disabling foot pain, the clinically significant difference of MFPDI score is unknown. Third, 62% of patients in the placebo group reached sufficient vitamin D levels after 24 months of follow-up, which might be due to seasonal change, outdoor physical activity, or other reasons that may lead to an underestimation of any benefit of vitamin D. In support of this, beneficial effects of vitamin D were also found in MFDPI scores after patients were divided into consistently and not consistently sufficient vitamin D groups.

In conclusion, this is the first study to show that vitamin D supplementation and maintaining sufficient vitamin D levels reduces foot pain over 2 years in patients with symptomatic knee OA. Vitamin D supplementation and maintaining a sufficient vitamin D level may improve foot pain in patients with knee OA.

ACKNOWLEDGMENTS

We thank the participants who made this study possible, and we gratefully acknowledge the role of staff and volunteers in the Vitamin D Effect on Osteoarthritis Study in collecting the data. We thank research assistants Jodi Barling, Kay Ngno, Judy Hankin, and Alice Noon, who were involved in the coordination of this study.

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. Ding 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. Cicuttini, Winzenberg, Jones, Wluka, Ding.

Acquisition of data. Han, Zhu, Antony, Ding.

Analysis and interpretation of data. Tu, Zheng, Jin, Gu, Ding.

REFERENCES

1. Hiligsmann M, Cooper C, Arden N, Boers M, Branco JC, Luisa Brandi M, et al. Health economics in the field of osteoarthritis: an expert's consensus paper from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). Semin Arthritis Rheum 2013;43:303–13.
2. Thomas MJ, Roddy E, Zhang W, Menz HB, Hannan MT, Peat GM. The population prevalence of foot and ankle pain in middle and old age: a systematic review. Pain 2011;152:2870–80.
3. Menz HB, Tiedemann A, Kwan MM, Plumb K, Lord SR. Foot pain in community-dwelling older people: an evaluation of the Manchester Foot Pain and Disability Index. Rheumatology (Oxford) 2006;45:863–7.
4. Suri P, Morgenroth DC, Kwoh CK, Bean JF, Kalichman L, Hunter DJ. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2010;62:1715–23.

5. Paterson KL, Hinman RS, Hunter DJ, Wrigley TV, Bennell KL. Impact of concurrent foot pain on health and functional status in people with knee osteoarthritis: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2015;67:989–95.
6. Croft P, Jordan K, Jinks C. "Pain elsewhere" and the impact of knee pain in older people. Arthritis Rheum 2005;52:2350–4.
7. Hill CL, Gill TK, Menz HB, Taylor AW. Prevalence and correlates of foot pain in a population-based study: the North West Adelaide health study. J Foot Ankle Res 2008;1:2.
8. Bouillon R, Carmeliet G. Vitamin D insufficiency: definition, diagnosis and management. Best Pract Res Clin Endocrinol Metab 2018;32:669–84.
9. Malacova E, Cheang PR, Dunlop E, Sherriff J, Lucas RM, Daly RM, et al. Prevalence and predictors of vitamin D deficiency in a nationally-representative sample of adults participating in the 2011–2013 Australian Health Survey. Br J Nutr 2019:1–24.
10. Straube S, Andrew Moore R, Derry S, McQuay HJ. Vitamin D and chronic pain. Pain 2009;141:10–3.
11. Anglin RE, Samaan Z, Walter SD, McDonald SD. Vitamin D deficiency and depression in adults: systematic review and meta-analysis. Br J Psychiatry 2013;202:100–7.
12. Shipton EE, Shipton EA. Vitamin D deficiency and pain: clinical evidence of low levels of vitamin D and supplementation in chronic pain states. Pain Ther 2015;4:67–87.
13. Gaikwad M, Vanlint S, Mittinity M, Moseley GL, Stocks N. Does vitamin D supplementation alleviate chronic nonspecific musculoskeletal pain? A systematic review and meta-analysis. Clin Rheumatol 2017;36:1201–8.
14. Gao XR, Chen YS, Deng W. The effect of vitamin D supplementation on knee osteoarthritis: a meta-analysis of randomized controlled trials. Int J Surg 2017;46:14–20.
15. Hussain S, Singh A, Akhtar M, Najmi AK. Vitamin D supplementation for the management of knee osteoarthritis: a systematic review of randomized controlled trials. Rheumatol Int 2017;37:1489–98.
16. Jin X, Jones G, Cicuttini F, Wluka A, Zhu Z, Han W, et al. Effect of vitamin D supplementation on tibial cartilage volume and knee pain among patients with symptomatic knee osteoarthritis: a randomized clinical trial. JAMA 2016;315:1005–13.
17. Garrow AP, Papageorgiou AC, Silman AJ, Thomas E, Jayson MI, Macfarlane GJ. Development and validation of a questionnaire to assess disabling foot pain. Pain 2000;85:107–13.
18. Roddy E, Muller S, Thomas E. Defining disabling foot pain in older adults: further examination of the Manchester Foot Pain and Disability Index. Rheumatology (Oxford) 2009;48:992–6.
19. Altman RD, Hochberg M, Murphy WA Jr, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage 1995;3 Suppl A:3–70.
20. Zhu Z, Otahal P, Wang B, Jin X, Laslett LL, Wluka AE, et al. Cross-sectional and longitudinal associations between serum inflammatory cytokines and knee bone marrow lesions in patients with knee osteoarthritis. Osteoarthritis Cartilage 2017;25:499–505.
21. 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.
22. Craig CL, Marshall AL, Sjostrom 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.
23. Paterson KL, Kasza J, Hunter DJ, Hinman RS, Menz HB, Peat G, et al. Longitudinal association between foot and ankle symptoms and worsening of symptomatic radiographic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 2017;25:1407–13.
24. Paterson KL, Kasza J, Hunter DJ, Hinman RS, Menz HB, Peat G, et al. The relationship between foot and ankle symptoms and risk of developing knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 2017;25:639–46.
25. Wu Z, Malihi Z, Stewart AW, Lawes CM, Scragg R. Effect of vitamin D supplementation on pain: a systematic review and meta-analysis. Pain Physician 2016;19:415–27.
26. Wu Z, Camargo CA Jr, Malihi Z, Bartley J, Waayer D, Lawes CM, et al. Monthly vitamin D supplementation, pain, and pattern of analgesic prescription: secondary analysis from the randomized, double-blind, placebo-controlled Vitamin D Assessment Study. Pain 2018;159:1074–82.
27. Hsiao MY, Hung CY, Chang KV, Han DS, Wang TG. Is serum hypovitaminosis D associated with chronic widespread pain including fibromyalgia? A meta-analysis of observational studies. Pain Physician 2015;18:E877–87.
28. Plotnikoff GA, Quigley JM. Prevalence of severe hypovitaminosis D in patients with persistent, nonspecific musculoskeletal pain. Mayo Clin Proc 2003;78:1463–70.
29. Hicks GE, Shardell M, Miller RR, Bandinelli S, Guralnik J, Cherubini A, et al. Associations between vitamin D status and pain in older adults: the Invecchiare in Chianti study. J Am Geriatr Soc 2008;56:785–91.
30. Krishnan AV, Feldman D. Mechanisms of the anti-cancer and anti-inflammatory actions of vitamin D. Annu Rev Pharmacol Toxicol 2011;51:311–36.
31. Kawasaki Y, Zhang L, Cheng JK, Ji RR. Cytokine mechanisms of central sensitization: distinct and overlapping role of interleukin-1β, interleukin-6, and tumor necrosis factor-α in regulating synaptic and neuronal activity in the superficial spinal cord. J Neurosci 2008;28:5189–94.
32. Glover TL, Horgas AL, Fillingim RB, Goodin BR. Vitamin D status and pain sensitization in knee osteoarthritis: a critical review of the literature. Pain Manag 2015;5:447–53.
33. Cruz-Almeida Y, King CD, Goodin BR, Sibille KT, Glover TL, Riley JL, et al. Psychological profiles and pain characteristics of older adults with knee osteoarthritis. Arthritis Care Res (Hoboken) 2013;65:1786–94.
34. Speed TJ, Richards JM, Finan PH, Smith MT. Sex moderates the effects of positive and negative affect on clinical pain in patients with knee osteoarthritis. Scand J Pain 2017;16:66–73.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 788–793
DOI 10.1002/acr.24194

BRIEF REPORT

Association of Intermittent and Constant Knee Pain Patterns With Knee Pain Severity and With Radiographic Knee Osteoarthritis Duration and Severity

Lisa C. Carlesso,[1] Gillian A. Hawker,[2] James Torner,[3] Cora E. Lewis,[4] Michael Nevitt,[5] and Tuhina Neogi,[6] for the Multicenter Osteoarthritis Study Group

Objective. To examine the relation of knee pain patterns to pain severity and to radiographic osteoarthritis (OA) severity and duration.

Methods. The Multicenter Osteoarthritis Study is a longitudinal cohort of older adults with or at risk of knee OA. Participants' Intermittent and Constant Osteoarthritis Pain (ICOAP) scores were characterized as 1) no intermittent or constant pain, 2) intermittent pain only, 3) constant pain only, and 4) a combination of constant and intermittent pain. Knee pain severity was assessed using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale and a visual analog scale (VAS). Radiographic knee OA (ROA) severity was defined as Kellgren/Lawrence grade ≥2, and ROA duration was defined according to the clinic visit at which ROA was first noted. We assessed the relation of ICOAP pain patterns to knee pain severity, ROA severity, and ROA duration using regression models with generalized estimating equations.

Results. There were 2,322 participants (mean age 68.8 years, body mass index 31.0 kg/m^2, 60% female). Higher ICOAP pain patterns, i.e., a mix of constant and intermittent pain, were associated with greater WOMAC pain severity compared with those patients without either pain pattern (odds ratio [OR] 43.2 [95% confidence interval (95% CI) 26.4–61.3]). Results were similar for the VAS (OR 71.2 [95% CI 45.7–110.9]). Those patients with more severe and longer duration of ROA were more likely to have a mix of constant and intermittent pain compared with those without either pain (OR 3.7 [95% CI 3.1–4.6] and OR 2.9 [95% CI 2.5–3.5], respectively).

Conclusion. Knee pain patterns are associated with radiographic disease stage and duration, as well as pain severity, highlighting the fact that pain patterns are important for understanding symptomatic disease progression.

INTRODUCTION

Knee osteoarthritis (OA) is a progressive disease whereby the frequency and severity of pain typically increases with worsening disease (1). Qualitative research has identified 3 specific patterns of pain in knee OA, which vary depending on the stage of the disease (2). These 3 patterns are reflected by the frequency of pain as being intermittent, constant, or a mix of constant with intermittent, whereby people experience intermittent activity-related pain, then constant pain as the disease progresses, and finally the late stage of disease is demarcated by constant pain overlaid by more severe, often unpredictable, intermittent pain. Based on this understanding, a new measure, the Intermittent and Constant Osteoarthritis Pain (ICOAP) scale was developed to capture these pain patterns, thus allowing for improved understanding of pain in the different phases of disease (3).

The psychometric properties of the ICOAP measure were initially assessed (items, subscale, and total scores) showing

The Multicenter Osteoarthritis Study is supported by the NIH (grants U01-AG-18820, U01-AG-18832, U01-AG-18947, and U01-AG-19079). Dr. Neogi's work was supported by the NIH (grants P60-AR-047785, K24-AR-070892, and R01-AR-062506).

[1]Lisa C. Carlesso, PT, PhD: McMaster University, Hamilton, Ontario, Canada; [2]Gillian A. Hawker, MD, MSc: University of Toronto and Women's College Research Institute, Toronto, Ontario, Canada; [3]James Torner, PhD: University of Iowa, Iowa City; [4]Cora E. Lewis, MD, MSPH: University of Alabama, Birmingham; [5]Michael Nevitt, PhD, MPH: University of California, San Francisco; [6]Tuhina Neogi, MD, PhD: Boston University School of Medicine, Boston, Massachusetts.

Dr. Neogi has received consulting fees from Pfizer/Lilly, Novartis, and EMD Merck-Serono (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Lisa C. Carlesso, PT, PhD, School of Rehabilitation Science, Faculty of Health Sciences, McMaster University, IAHS 415, 1400 Main Street West, Hamilton, Ontario, Canada, L8S 1C7. Email: carlesl@mcmaster.ca.

Submitted for publication June 10, 2019; accepted in revised form March 17, 2020.

SIGNIFICANCE & INNOVATIONS

- In a cross-sectional analysis of a prospective cohort of people with or at risk of knee osteoarthritis, knee pain patterns (intermittent pain, constant pain, or constant and intermittent pain) were associated with radiographic disease stage and duration, as well as with pain severity.
- These findings highlight the importance of pain patterns for understanding symptomatic disease progression.

significant correlations with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale and the Knee Injury and OA Outcomes symptom scale in a relatively small sample (3). The scale has subsequently been validated with measures of self-reported function, physical performance, and physical activity using accelerometry (4,5). However, these studies have not evaluated the relation of the 3 different pain patterns (i.e., intermittent, constant, or constant and intermittent) that appear to reflect different disease stages, regardless of pain severity, to established symptom severity assessment instruments. In addition, there has not been a specific evaluation to date regarding the original premise of ICOAP pain patterns reflecting the stage of disease in OA, typically defined radiographically. In other words, we do not know whether mild radiographic findings or early-stage disease are associated with intermittent pain, intermediate radiographic stages with constant pain, and late disease with both. Therefore, the objectives of this study were to examine the relation of ICOAP-defined pain patterns with knee pain severity, radiographic disease severity, and duration.

SUBJECTS AND METHODS

The Multicenter Osteoarthritis (MOST) Study is an NIH-funded longitudinal cohort of community-dwelling adults between ages 50 and 79 years who have or are at risk of developing knee OA at baseline. Subjects were recruited from Birmingham, Alabama, and Iowa City, Iowa. Details of the cohort have been published elsewhere (6). The study was approved by the institutional review boards at the University of Iowa; University of Alabama at Birmingham; University of California, San Francisco; and Boston University Medical Center that were in compliance with the Helsinki Declaration. The current sample comprised participants who attended the 60-month visit (baseline for this study), since it was the first time the ICOAP measure was obtained.

Pain measures. The ICOAP is an 11-item measure consisting of items for each of 2 subscales, Intermittent and Constant. Items include pain intensity, pain frequency (for the Intermittent subscale), effect on sleep and quality of life, and the extent to which the pain "upsets or worries" and "frustrates or annoys." Initial psychometric testing of the scale demonstrated good validity and reliability (3). At the 60-month visit, participants who reported at least some knee pain in the prior 30 days were asked to complete ICOAP. ICOAP data was obtained in a knee-specific manner, inquiring about symptoms over the prior 7 days. ICOAP pain patterns were defined according to responses to each respective subscale item on severity, ranging from none to extremely on a 5-point scale as follows: 1) no intermittent or constant pain, 2) intermittent pain only (of at least "mild" severity and with a frequency of at least "sometimes"), 3) constant pain only (of at least "mild" severity), and 4) a combination of constant and intermittent pain, as defined above. Pain severity was measured using a knee-specific WOMAC pain subscale (Likert version, range 0–20) inquiring about pain during the past 30 days (6). Scores were categorized as none, mild/moderate, or severe/extreme (7,8). Higher scores on the WOMAC indicate greater pain. A knee-specific visual analog scale (VAS) measured average pain severity (0–10) in the past 30 days, and was categorized as 0, 1–4, and >4 (of 10) (9).

Radiographic analysis: duration and severity. Bilateral weight-bearing fixed-flexion posteroanterior radiographs of the knee were obtained at each study visit (0, 30, and 60 months). Radiographic severity in the tibiofemoral joint was graded by 2 experienced readers blinded to clinical data according to Kellgren/Lawrence (K/L) criteria (0–4) (10). Any disagreements between readers were adjudicated by a third reader along with the first 2 readers to reach consensus. The interrater reliability weighted kappa for the K/L grade was $\kappa = 0.80$. Radiographic knee OA (ROA) was defined as K/L grade ≥2. OA duration was defined according to the clinic visit at which ROA was first noted (i.e., the longest duration was for those who had ROA at baseline; the shortest duration was for those whose ROA was identified at the 60-month visit).

Potential confounders and relevant covariates. Variables included age, sex, body mass index (BMI), widespread pain, depressive symptoms, pain catastrophizing, study site, race, and K/L grade (for the pain severity analyses) at the 60-month visit. As per previous studies, widespread pain was operationalized using a validated standard homunculus (11). A Center for Epidemiologic Studies Depression Scale score of ≥16 was used to define depressive symptoms (12). Pain catastrophizing was measured using 1 item from the Coping Strategies Questionnaire, which has been shown to be valid and reliable (13). Race was categorized as White versus other.

Statistical analysis. We first evaluated the mean WOMAC pain and mean VAS pain (outcomes) for each ICOAP pain category (exposure: none, intermittent only, constant pain only, both constant and intermittent pain), and the relation of ICOAP pain patterns (exposure) to the likelihood of having greater pain severity. We used generalized estimating equations (GEE) to account for 2 knees within an individual. We also hypothesized that those

patients with greater pain severity would be more likely to have constant pain. We therefore evaluated the relation of knee pain severity categories (WOMAC and VAS, separately) (exposures) to ICOAP pain patterns (outcomes) using proportional odds logistic regression with GEE. Similarly, we hypothesized that those with greater ROA severity and duration would likely have a more advanced ICOAP pain pattern. To evaluate this possibility, we examined the relation of ROA severity and duration (exposures) with ICOAP pain patterns (outcomes) using proportional odds logistic regression with GEE. All analyses were adjusted for age, sex, BMI, widespread pain, depressive symptoms, pain catastrophizing, and clinic site. K/L grade and race were additionally adjusted for in the knee pain severity analyses. All analyses were performed using SAS software, version 9.3.

RESULTS

At the 60-month visit, there were 2,322 participants (4,632 knees) with ICOAP data (mean ± SD age 68.8 ± 8 years, mean ± SD BMI 31.0 ± 6, 60% female). The majority of knees (62%) had neither intermittent nor constant pain, 30% had intermittent pain only, 4% had constant pain only, and 4% had both. In all, 60% of knees had no ROA, while 5.5% had incident ROA at the 60-month visit (shortest duration of OA), 5.5% had incident ROA at the 30-month visit, and 29.5% had ROA at the baseline visit (longest duration of OA) (Table 1).

Knee pain severity. By both the WOMAC and VAS scores, approximately 55% of knees had mild to moderate pain, and approximately 11% had severe/extreme pain. Mean pain scores increased with each successive ICOAP pain category, i.e., no intermittent or constant pain, intermittent pain only, constant pain only, and constant plus intermittent pain: mean WOMAC pain (1.2, 4.9, 8.2, and 9.0, respectively) and mean VAS pain (6.0, 27.5, 43.5, and 53.2, respectively). ICOAP pain patterns (as per our definitions), were associated with a greater likelihood of being in a higher pain severity category by WOMAC and VAS. Specifically, those patients with a mix of constant and intermittent pain had 43 and 71 times higher odds of having greater pain severity than those without either type of pain (WOMAC odds ratio [OR] 43.2 [95% confidence interval (95% CI) 26.4–61.3]; VAS OR 71.2 [95% CI 45.7–110.9]) (Figure 1).

Additionally, greater WOMAC and VAS pain severity categories had higher odds of being associated with constant rather than intermittent pain only. For example, those patients with severe/extreme pain and those with mild/moderate pain by WOMAC had 3.8 times (95% CI 1.5–9.4) and 1.4 times (95% CI 0.6–3.3) higher odds, respectively, of having constant versus intermittent pain only compared with those with a WOMAC pain score of 0. For VAS, those with scores >4 (OR 4.7 [95% CI 1.7–12.6]) and those with pain scores between 1 and 4 (OR 1.2 [95% CI 0.5–3.3]) similarly had higher odds of having constant versus intermittent pain only compared with those whose VAS score was 0 (Table 2).

ROA severity and duration. As shown in Table 3, those patients with greater ROA severity and a longer duration of knee ROA had higher odds of having pain patterns hypothesized to be associated with such a disease status (i.e., both intermittent and constant pain at the "highest" end of the ICOAP pain pattern). For ROA severity, there was a significant trend of having increasing odds of a higher K/L grade (1, 2, 3, or 4) with combined constant and intermittent pain (OR 1.3, 2.0, and 3.7, respectively), compared with neither pain being present. Similarly, a significant trend was observed for ROA duration. When we limited our analysis to

Table 1. Participant characteristics (n = 2,322; 4,632 knees)*

Characteristic	Value
Age, mean ± SD years	68.8 ± 8.0
Female, %	60
Body mass index, mean ± SD kg/m^2	31.0 ± 6.0
ICOAP pain patterns (knees)	
No intermittent or constant pain	2,873 (62)
Intermittent pain only	1,389 (30)
Constant pain only	185 (4)
Both constant and intermittent pain	185 (4)
Kellgren/Lawrence grade (knees)	
0	2,130 (46)
1	649 (14)
2	834 (18)
3 or 4	1,019 (22)
Radiographic knee OA status (knees)	
No OA	2,779 (60)
Incident at 60 months (shortest duration)	255 (5.5)
Incident at 30 months	255 (5.5)
Prevalent at baseline (longest duration)	1,343 (29)

* Values are the number (%), calculated on the number of knees, unless indicated otherwise. ICOAP = Intermittent and Constant Osteoarthritis Pain; OA = osteoarthritis.

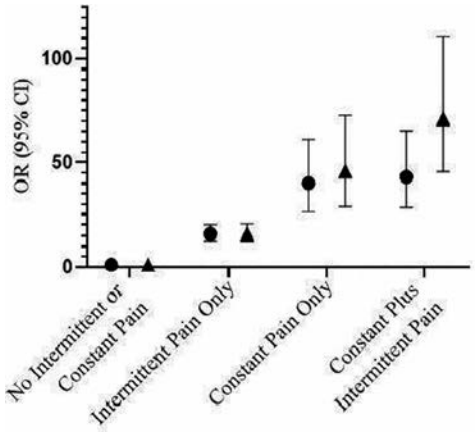

Figure 1. Maximal Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC; circles) and visual analog scale (VAS; triangles) pain category by pain patterns. Adjusted for age, sex, body mass index, catastrophizing, depressive symptoms, widespread pain, race, Kellgren/Lawrence grade, and clinic site. *P* for linear trend <0.0001 for both WOMAC and VAS. 95% CI = 95% confidence interval; OR = odds ratio.

Table 2. Association of WOMAC and VAS categories with ICOAP constant versus intermittent pain*

	Values
Maximal WOMAC knee pain†	
None (n = 1,638)	1.0 (ref.)
Mild/moderate pain (n = 2,471)	1.4 (0.6–3.3)
Severe/extreme pain (n = 514)	3.8 (1.5–9.4)
Maximal VAS knee pain†	
0 (n = 1,560)	1.0 (ref.)
1–4 (n = 2,529)	1.2 (0.5–3.3)
>4 (n = 534)	4.7 (1.7–12.6)

* Values are the odds ratio (95% confidence interval) unless indicated otherwise, adjusted for age, sex, body mass index, catastrophizing, depressive symptoms, widespread pain, race, Kellgren/Lawrence grade, and clinic site. ICOAP = Intermittent and Constant Osteoarthritis Pain; ref. = reference; VAS = visual analog scale; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† *P* for linear trend < 0.0001.

those patients with some pain (i.e., excluded those who reported no pain on ICOAP; final column of Table 3), those with the shortest duration of OA had a similar likelihood of having constant versus intermittent pain as those with no ROA (OR 0.7 [95% CI 0.4–1.3]). Those with longer durations of OA, i.e., present for at least 60 months, had 1.4 times higher odds of having constant versus intermittent pain (OR 1.4 [95% CI 1.0–2.0]).

DISCUSSION

Previously published qualitative data have suggested that with the progression of knee OA over time, the pain associated with knee OA transitions from intermittent to constant pain, punctuated by intermittent unpredictable pain (2). This transition served as the foundation for the development of the ICOAP measure. In light of this information, we sought to evaluate whether these identified conceptual clinical pain patterns, regardless of pain severity and as assessed by the ICOAP, were associated with expected increments in radiographic disease severity and longer duration of radiographic disease, as well as with greater pain severity. We found that knee pain patterns defined by the ICOAP instrument were associated with greater ROA severity and duration. We note that the ICOAP-defined pain patterns indicative of later-stage disease, regardless of symptom severity on the ICOAP instrument, were associated with greater pain severity by both WOMAC and VAS.

These ICOAP-defined pain patterns, to our knowledge, have not been analyzed in this manner in prior published work. Therefore, similar comparisons are not possible, because typically only the summed subscale scores and/or their correlations with, for example, the WOMAC pain subscale, have been published (5,14). An association of the summed scores with pain severity by WOMAC is not surprising, since the ICOAP summed score includes an element of pain severity in some of the questions. Our evaluation of the knee pain patterns defined those patterns without regard to pain severity (i.e., intermittent or constant pain was defined by their reported presence). Additionally, the ICOAP assesses symptoms during the past week, while WOMAC and VAS pain scales assessed pain in the past 30 days, allowing us to examine the implications of the knee pain patterns outside of the 1-week time frame. Notably, while the ORs were large, the 95% CIs were wide, demonstrating imprecision of the estimates, reflecting the lower prevalence of constant pain and of constant plus intermittent pain in this community-based sample. Perhaps the closest approximation of our results come from a previous study of MOST data that focused on consistency of knee pain symptoms over 1 month. Although the ICOAP was not used, knee pain severity was reported to be higher in those patients with consistent pain (present on most days over a 2-month period) compared to inconsistent pain (only present on most days over 1 month), and those with ROA were less likely to have inconsistent pain (15).

We found that greater severity and longer duration of ROA were associated with a greater likelihood of constant plus intermittent pain compared with neither pain type being present. In addition, longer ROA duration was also associated with constant pain only versus intermittent pain only, and there was a dose-response relationship. These results suggest the possibility that differing pain mechanisms underlie intermittent and constant pain. For example, constant pain found in later stages of disease severity may be representative of central pain sensitization, whereas earlier intermittent pain may be largely peripherally driven nociceptive input (16). These findings speak more broadly to the need to understand other pain dimensions, such as these pain patterns,

Table 3. Association of radiographic severity and duration of OA with ICOAP*

Knee OA severity	ICOAP 4-level outcome	Knee OA duration	ICOAP 4-level outcome	ICOAP constant vs. intermittent pain only†
K/L 0 (n = 2,130, 46%)	1.0 (ref)	No OA (n = 2,779)	1.0 (ref)	1.0 (ref)
K/L 1 (n = 649, 14%)	1.3 (1.1–1.7)	Incident OA at 60 months (shortest duration; n = 255)	1.8 (1.3–2.4)	0.7 (0.4–1.3)
K/L 2 (n = 834, 18%)	2.0 (1.6–2.5)	Incident OA at 30 months (n = 255)	2.3 (1.7–3.1)	1.5 (0.9–2.5)
K/L 3 or 4 (n = 1,019, 22%)	3.7 (3.1–4.6)	Prevalent at first study visit (longest duration; n = 1,343)	2.9 (2.5–3.5)	1.4 (1.0–2.0)
P for linear trend	<0.0001	–	<0.0001	0.03

* Values are the odds ratio (95% confidence interval) unless indicated otherwise, adjusted for age, sex, body mass index, catastrophizing, depressive symptoms, widespread pain, and clinic site. Intermittent and Constant Osteoarthritis Pain (ICOAP) pain patterns modeled as 4-level ordered outcome as defined in Subjects and Methods. Percentages are calculated based on 4,632 knees. K/L = Kellgren/Lawrence; OA = osteoarthritis; ref. = reference.
† ICOAP pain modeled as any constant pain versus intermittent pain only.

beyond pain severity alone to truly understand symptomatic disease progression.

The relation of ICOAP-defined pain patterns to ROA duration and severity is novel and lends new support to previous longitudinal studies that have demonstrated the presence of different pain patterns and their variability (17). These results provide, for the first time, proof-of-concept evidence that these pain patterns do indeed track with OA structural disease. This work supports a relationship between the pain experience and its association with ROA severity and duration, which have had conflicting correlations with pain severity (17). These results point to the likelihood that pain severity itself is not an adequate metric to understand the stage of symptomatic OA disease. Our data suggest that clinicians may be able to use the ICOAP as a tool to effectively track knee OA progression, and this approach may potentially help mitigate the so-called "structure-symptom" discordance. Studies are needed to understand what may trigger the transition to more advanced pain patterns. However, discerning at what stage of OA pain patterns change from being intermittent in nature to constant, and then to constant with unpredictable intermittent pain, will require longitudinal data in which greater variety in duration (and severity of OA) is captured along with the unpredictability of the intermittent pain.

Our main study limitation is that we were unable to ascertain the influence of the onset or frequency of unpredictable intermittent pain that occurs after a specific trigger. This information may provide further discrimination of the stage of disease. Strengths of our study include the examination and validation of ICOAP-defined pain patterns with important indicators of the stage of disease, with adjustment for known confounders and relevant covariates, in addition to our use of standardized and validated questionnaires. Further, this is the largest study to date to validate ICOAP with high-quality standardized radiographs.

In conclusion, ICOAP-derived knee pain patterns (intermittent, constant, constant and intermittent) are associated with overall pain severity symptoms, disease duration, and severity of ROA. This finding supports previous qualitative work that described a progression from intermittent to constant pain, culminating in a combination of the 2 as the OA disease process progresses. These findings highlight the need for a broader approach to understanding pain and its mechanisms that likely differ by stage of disease. Importantly, while pain severity alone is insufficient to understand disease stage and progression, these knee pain patterns appear likely to be more useful for understanding symptomatic disease progression.

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. Carlesso 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. Hawker, Neogi.
Acquisition of data. Torner, Lewis, Neogi.
Analysis and interpretation of data. Carlesso, Torner, Nevitt, Neogi.

REFERENCES

1. Pan F, Tian J, Aitken D, Cicuttini F, Jones G, de Rooij M, et al. Predictors of pain severity trajectory in older adults: a 10.7-year follow-up study. Osteoarthritis Cartilage 2018;26:1619–26.
2. Hawker GA, Stewart L, French MR, Cibere J, Jordan JM, March L, et al. Understanding the pain experience in hip and knee osteoarthritis: an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16: 415–22.
3. Hawker GA, Davis AM, French MR, Cibere J, Jordan JM, March L, et al. Development and preliminary psychometric testing of a new OA pain measure: an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16:409–14.
4. Song J, Chang AH, Chang RW, Lee J, Pinto D, Hawker G, et al. Relationship of knee pain to time in moderate and light physical activities: data from Osteoarthritis Initiative. Semin Arthritis Rheum 2018;47:683–8.
5. Davison MJ, Ioannidis G, Maly MR, Adachi JD, Beattie KA. Intermittent and constant pain and physical function or performance in men and women with knee osteoarthritis: data from the Osteoarthritis Initiative. Clin Rheumatol 2016;35:371–9.
6. 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.
7. Wang K, Kim HA, Felson DT, Xu L, Kim DH, Nevitt MC, et al. Radiographic knee osteoarthritis and knee pain: cross-sectional study from five different racial/ethnic populations. Sci Rep 2018; 8:1364.
8. Neogi T, Felson D, Niu J, Nevitt M, Lewis CE, Aliabadi P, et al. Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. BMJ 2009;339:b2844.
9. McAlindon TE, Driban JB, Henrotin Y, Hunter DJ, Jiang GL, Skou ST, et al. OARSI clinical trials recommendations: design, conduct, and reporting of clinical trials for knee osteoarthritis. Osteoarthritis Cartilage 2015;23:747–60.
10. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthrosis. Ann Rheum Dis 1957;16:494–502.
11. Leveille SG, Zhang Y, McMullen W, Kelly-Hayes M, Felson DT. Sex differences in musculoskeletal pain in older adults. Pain 2005;116:332–8.
12. Radloff L. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1:385–401.
13. Jensen MP, Keefe FJ, Lefebvre JC, Romano JM, Turner JA. One- and two-item measures of pain beliefs and coping strategies. Pain 2003;104:453–69.
14. Zhang C, Liu DH, Qu YL, Jia ZY, Wang W, Li J, et al. Transcultural adaptation and validation of the Chinese version of the intermittent and constant osteoarthritis pain (ICOAP) measure in patients with knee osteoarthritis. Osteoarthritis Cartilage 2017;25:506–12.
15. Neogi T, Nevitt M, Yang M, Curtis J, Torner J, Felson D. Consistency of knee pain: correlates and association with function. Osteoarthritis Cartilage 2010;18:1250–5.
16. Carlesso LC, Frey Law L, Wang N, Nevitt M, Lewis CE, Neogi T. The relation of pain sensitization and conditioned pain

modulation to pain patterns in knee osteoarthritis: the Multicenter Osteoarthritis Study. American College of Rheumatology Annual Conference; 2019 October; Atlanta, GA. Arthritis Rheumatol 2019;71:3351–2.

17. Nicholls E, Thomas E, van der Windt DA, Croft PR, Peat G. Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. Osteoarthritis Cartilage 2014;22:2041–50.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 794–800
DOI 10.1002/acr.24385

American College of Rheumatology
Empowering Rheumatology Professionals

Get a Grip on Factors Related to Grip Strength in Persons With Hand Osteoarthritis: Results From an Observational Cohort Study

Ida K. Haugen,[1] Janni Aaserud,[2] and Tore K. Kvien[3]

Objective. To compare levels of grip strength in individuals with hand osteoarthritis (OA) with normative values, and to examine how hand OA severity and other biopsychosocial factors are associated with grip strength.

Methods. Levels of grip strength across age groups were compared with normative values from the general population in sex-stratified analyses using 2-sample *t*-tests. Associations between radiographic hand OA severity (Kellgren/Lawrence sum score) in different joint groups and grip strength of the same hand were examined in 300 individuals from the Nor-Hand study using linear regression. Analyses were repeated using markers of pain, demographic factors, comorbidities, and psychological and social factors as independent variables. We adjusted for age, sex, and body mass index.

Results. Individuals with hand OA had lower grip strength than the general population, especially in individuals age <60 years. In thumb base joints, increasing radiographic severity (range 0–8) and the presence of pain were associated with lower grip strength (β = –0.83 [95% confidence interval (95% CI) –1.12, –0.53] and β = –2.15 [95% CI –3.15, –1.16], respectively). Negative associations with grip strength were also found for women, low education, higher comorbidity index, and higher resting heart rate.

Conclusion. Individuals with hand OA have lower grip strength than the general population. Our results support the idea that studies on thumb base OA should include grip strength as an outcome measure. However, other biopsychosocial factors should also be considered when the grip strength is being interpreted, because other factors such as sex, socioeconomic factors, physical fitness, and comorbidities are negatively associated with grip strength.

INTRODUCTION

Hand osteoarthritis (OA) is a highly prevalent disease, increasing in industrialized countries due to an aging population. It represents a considerable burden of disease of which awareness in research has increased over the last decade (1). However, there is still limited knowledge about the pathogenesis of hand OA and a lack of effective treatment options for the affected patients.

The main symptoms of hand OA are pain and aching in the affected joints. Other important symptoms include reduced grip strength, stiffness, loss of mobility, aesthetic damage, and disability (1). Studies of patient perspectives have shown that problems with gripping and reduced strength are considered important factors for hand OA patients, and grip strength is essential for the ability to carry out activities of daily life (2). Hence, activities requiring grip strength are included in patient-reported outcome measures assessing physical function in hand OA patients (3,4). Measurements of grip strength are important because they give an indication of the functional integrity of the hand (5). The international organization Outcome Measures in Rheumatology (OMERACT) recommends hand strength as part of the core domains for all hand OA studies, and assessment of grip strength, or alternatively pinch strength, is the current recommended instrument to assess hand strength (1). Previous studies have shown that increasing radiographic severity of hand OA is associated with reduced grip strength. In particular, a higher risk of reduced grip and pinch strength is observed in patients with radiographic OA in the thumb, including the first carpometacarpal (CMC1) joints, and the metacarpophalangeal (MCP) joints (5). The association of hand OA with function and strength measurements has been suggested to be largely mediated by pain (5,6).

The Nor-Hand study was supported by the South-East Norway Regional Health Authority, the Pahles Foundation, the Simon Fougner Hartmanns Family Foundation, and the Trygve Gythfeldt and Wife's Research Foundation.

[1]Ida K. Haugen, MD, PhD: Diakonhjemmet Hospital, Oslo, Norway; [2]Janni Aaserud, MD: University of Oslo, Oslo, Norway; [3]Tore K. Kvien, MD, PhD: Diakonhjemmet Hospital and University of Oslo, Oslo, Norway.

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

Address correspondence to Ida K. Haugen, MD, PhD, Department of Rheumatology, Diakonhjemmet Hospital, Post Box 23 Vinderen, 0319 Oslo, Norway. Email: ida.haugen@diakonsyk.no.

Submitted for publication March 23, 2020; accepted in revised form July 14, 2020.

SIGNIFICANCE & INNOVATIONS

- Individuals with hand osteoarthritis (OA) have lower grip strength than individuals from the general population, especially individuals age <60 years.
- Structural pathology and pain in the thumb base joints are more strongly associated with reduced grip strength than structural pathology and pain in the finger joints. Our results suggest that grip strength should be an outcome measure in clinical trials on patients with thumb base OA but is less relevant in studies focusing on interphalangeal OA.
- Grip strength is not only a marker of hand OA. Female sex, socioeconomic factors, physical fitness, and comorbidities are also negatively associated with grip strength, suggesting that other biopsychosocial factors should be taken into account when interpreting the level of grip strength in individuals with hand OA.

On the other hand, in most medical areas, grip strength is used to define frailty, with weakness considered a key manifestation of sarcopenia. Fried et al included reduced grip strength as one key element of the "frailty phenotype," together with weight loss, exhaustion, slow gait speed, and low activity (7). In women with body mass index (BMI) of ≤26.0 kg/m^2, low grip strength is defined as grip strength of the dominant hand of ≤17.3 kg. Similar values for men are ≤30.0 kg. Reduced grip strength is the most common frailty criterion, after exhaustion, among prefrail older individuals (8).

Few previous studies have investigated the association between general health and grip strength in patients with OA. In the European Project on OA (EPOSA), a cohort of elderly individuals age >65 years, 17% of the participants had clinical hand OA, and associations between several comorbidities and grip strength were identified (9).

Our aim was to compare levels of grip strength in individuals with hand OA with normative values from the general population. In explorative analyses, we were aiming for a better understanding of how grip strength in individuals with hand OA is affected by both their hand OA severity defined by radiographs and joint pain as well as by other biopsychosocial factors.

MATERIALS AND METHODS

Study population. The Nor-Hand study is a large-scale hospital-based observational cohort study, including 300 patients with hand OA ages 40–70 years. The current analyses included cross-sectional data from the baseline examination. Each of the participating patients had hand OA diagnosed either by ultrasound or clinical examination by a rheumatologist. A more detailed description of the inclusion and exclusion criteria can be found in the published protocol (10). Among the exclusion criteria were diagnoses of rheumatoid arthritis, psoriasis, psoriatic arthritis, reactive arthritis, spondyloarthritis, and hemochromatosis. Patients with major comorbidities making them unable to attend the study visit were also excluded.

Grip strength. Maximal isometric grip strength, which is the gold standard for measurement of grip strength, was measured using a Jamar dynamometer. The procedure was executed with the patient sitting in a chair with his/her arm unsupported, keeping the elbow at a 90-degree angle. The dominant hand was tested first, by having the patient squeeze the dynamometer as hard as possible. The test was then repeated 2 times with 15 seconds rest between the assessments, before the same procedure was repeated on the nondominant hand. The results were recorded in kilograms with 1 decimal precision. In our analyses we used the mean value of the 3 measurements of grip strength for each hand. For 2 patients grip strength was measured in 1 hand only.

Markers of hand OA. All patients obtained bilateral frontal hand radiographs (posteroanterior view). A trained reader (IKH) evaluated the 2nd–5th distal interphalangeal (DIP), 1st–5th proximal interphalangeal (PIP), MCP, and CMC1 joints using the Kellgren/Lawrence (K/L) scale (11). In addition, the scaphotrapeziotrapezoidal (STT) joints were scored similarly, although not included in the original scale. Twenty randomly selected hand radiographs were scored twice by the same reader with a mean ± SD interval of 16 ± 4 days between the first and second scoring. The intrareader reliability was excellent, with weighted $\kappa = 0.92$.

Patients were asked to rate the intensity of their pain in each hand after measuring the grip strength. Other questionnaires were completed either prior to or after the clinical examination, including general hand pain and disease activity during the past 24 hours on a numerical rating scale (NRS) (0 = no pain, 10 = worst imaginable pain/disease activity) and the Australian/Canadian Hand index (AUSCAN), which consists of 5 questions addressing pain in rest and during activities in the last 48 hours. Each question is answered on 0–4 scales, leading to a sum score of 0–20 (4). The patients were also asked to assess the presence of pain in individual hand joints during the last 6 weeks on a hand diagram. Whereas NRS pain after testing grip strength and pain in individual hand joints were assessed in both hands separately, the assessment of NRS pain, disease activity, and AUSCAN pain referred to both hands together. Regular use of analgesics was defined as self-reported regular use of oral acetaminophen, nonsteroidal antiinflammatory drugs, or opioid or opioid-like analgesics.

Markers of other biopsychosocial factors. Each patient received an electronic case report form (or paper version if needed) including questions regarding their age, previous injuries of the hands/wrists, physical exercise, and comorbidities. Physical activity was self-reported on 1 question ("how many times you exercise with increased heart beat and respiratory rate for at least 30 minutes") with 4 response categories ("3 or more times/week,"

"1–2 times/week," "1–2 times/month," or "not regularly") (12), of which the first 2 and the 2 latter categories were combined in our analyses. We collected data on comorbidities by having each patient respond to the Self-Administered Comorbidity Questionnaire, which includes 12 of the most prevalent medical conditions in general practice and 3 additional unspecified conditions (13). To illustrate the severity and consequences of the diseases, the questionnaire includes questions about treatment and how the conditions are affecting daily life, giving a total score of 0–45. A trained medical student inspected the answers along with the patients, comparing the comorbidities to their list of medications (10).

To calculate BMI (kg/m^2), a trained medical student measured the height and weight of the patients. Height was measured to the nearest millimeter and weight in kilograms with 1 decimal precision. A trained medical student measured blood pressure and heart rate. The examinations of blood pressure and heart rate were done using an automatic blood pressure machine with the patients in sitting position after 5 minutes of rest in supine position in a quiet room. Repeated measurements were performed until 2 consecutive measurements had a ≤5 mm Hg difference in both systolic and diastolic blood pressure. The mean of these 2 blood pressure and heart rate measurements were used in the analyses.

Social factors included education, use of alcohol, and smoking. Education was divided into 2 categories ("lower than or completed secondary school" versus "higher education at college or university"). Alcohol consumption was examined using the Alcohol Use Disorders Identification Test–Consumption, which includes 3 questions about the frequency and amount of alcohol drinking. Each question has response categories on 0–4 scales, giving a total sum score of 0–12 (14). High alcohol consumption (possible harmful drinking) is defined as a score of 3 or more for women and 4 or more for men. Smoking was reported into 4 categories (never smoker, previous smoker, current daily smoker, and current nondaily smoker), and we dichotomized the smoking variable into current daily/nondaily smokers versus never/previous smokers in the analyses. Psychological health was assessed using the Hospital Anxiety and Depression Scale (HADS), containing 14 questions on 0–3 scales regarding symptoms of depression and anxiety, giving a total score of 0–42 (15).

Statistical analysis. We calculated the mean ± SD grip strength within different age groups in men and women and compared these results with normative values from the general population using 2-sample *t*-tests based on provided information about numbers, mean values, and SDs (16). The severity of radiographic OA in different joint groups (i.e., rays and rows) was calculated using K/L sum scores of the respective joints (11). The association between radiographic hand OA severity (independent variable) and grip strength of the same hand (dependent variable) was examined using linear regression analyses. Unstandardized beta values with 95% confidence intervals (95% CIs) are presented. The analyses were repeated using self-reported pain, disease activity, and other biopsychosocial factors as the independent variables. Regression analyses were performed using generalized estimating equations to account for dependency between the 2 hands within each patient and were adjusted for age, sex, and BMI (SPSS software, version 26). *P* values less than 0.05 were regarded as statistically significant. We performed sex-stratified analyses in case of interactions with sex (*P* less than 0.10).

RESULTS

Demographic and clinical characteristics. Demographic and clinical characteristics are shown in Table 1. The majority of the participants were women. More than half of the participants had at least 1 year of college or university education. The participants demonstrated a wide range of radiographic severity and symptom severity, although a minority of the participants reported regular use of analgesics. The majority fulfilled the American College of Rheumatology criteria for hand OA (17). In total, 56 participants (19.3%) and 25 participants (8.6%) had anxiety and depression scores of ≥8 on the HADS, respectively. There

Table 1. Demographic and clinical characteristics of the study population (n = 300)*

Characteristic	Value
Age, median (IQR) years	61.0 (56.7–65.9)
Women	266 (88.7)
Body mass index, mean ± SD kg/m^2	26.5 ± 5.0
Fulfilling the ACR criteria for hand OA	278 (92.7)
NRS disease activity (range 0–10), mean ± SD†	3.7 ± 2.2
NRS hand pain (range 0–10), mean ± SD†	3.8 ± 2.3
AUSCAN pain (range 0–20), mean ± SD	8.2 ± 4.0
Regular use of oral analgesics	44 (14.7)
K/L sum score all joints right hand (range 0–64), median (IQR)	14.0 (7.0–21.0)
K/L sum score all joints left hand (range 0–64), median (IQR)†	14.0 (7.0–22.0)
Previous right hand/wrist injury†	49 (16.3)
Previous left hand/wrist injury†	49 (16.3)
Infrequent physical exercise†	90 (30.0)
Comorbidity questionnaire sum score (range 0–45), median (IQR)	7.0 (4.0–10.0)
Systolic blood pressure, median (IQR) mm Hg†	129 (120–145)
Diastolic blood pressure, median (IQR) mm Hg†	79 (75–86)
Heart rate at rest, median (IQR) beats/minute†	69 (62–76)
Hospital Anxiety Depression Scale (range 0–42), median (IQR)†	6.0 (3.0–10.0)
Low education†	125 (41.8)
Current daily or nondaily smoking	45 (15.0)
High alcohol consumption (AUDIT-C)†	205 (68.6)

* Values are the number (%) unless indicated otherwise. ACR = American College of Rheumatology; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption; AUSCAN = Australian/Canadian Osteoarthritis Hand Index; IQR = interquartile range; K/L = Kellgren/Lawrence; NRS = numerical rating scale.
† N = 1 missing value for NRS pain, NRS disease activity, systolic blood pressure, diastolic blood pressure, K/L sum score in the left hand, education, and alcohol consumption; n = 2 missing values for heart rate at rest; n = 3 missing values for previous injuries; n = 5 missing values for physical exercise; n = 10 missing values for Hospital Anxiety and Depression Scale.

Table 2. Grip strength in individuals with hand OA in comparison with normative values from the general population*

	Men			Women		
Age group	Normative	Nor-Hand	*P*	Normative	Nor-Hand	*P*
45–49 years	(n = 28)	(n = 2)		(n = 25)	(n = 14)	
Right	49.8 ± 10.4	41.3 ± 1.9	0.27	28.2 ± 6.8	20.9 ± 8.4	0.005
Left	45.7 ± 10.3	38.1 ± 3.1	0.31	25.4 ± 5.7	21.1 ± 5.9	0.03
50–54 years	(n = 25)	(n = 2)		(n = 25)	(n = 35)	
Right	51.5 ± 8.2	41.1 ± 21.5	0.13	29.8 ± 5.2	20.3 ± 9.0	<0.001
Left	46.2 ± 7.7	37.5 ± 14.4	0.16	26.0 ± 4.8	19.2 ± 9.7	0.002
55–59 years	(n = 21)	(n = 7)		(n = 25)	(n = 71)	
Right	45.9 ± 12.1	31.2 ± 7.9	0.01	26.0 ± 5.7	19.3 ± 7.0	<0.001
Left	37.7 ± 10.6	31.1 ± 7.2	0.14	21.5 ± 5.3	18.0 ± 6.9	0.02
60–64 years	(n = 24)	(n = 12)		(n = 25)	(n = 69)†	
Right	40.7 ± 9.2	35.4 ± 15.7	0.21	25.0 ± 4.5	21.7 ± 6.2	0.02
Left	34.8 ± 9.2	33.5 ± 15.2	0.75	20.7 ± 4.5	20.1 ± 6.6	0.68
65–69 years	(n = 27)	(n = 9)		(n = 28)	(n = 69)†	
Right	41.3 ± 9.3	34.5 ± 9.1	0.06	21.3 ± 4.3	19.8 ± 8.3	0.37
Left	34.8 ± 8.9	34.2 ± 12.2	0.87	18.6 ± 3.7	18.8 ± 9.1	0.91

* Values are the mean ± SD unless indicated otherwise. One person age 43 years and 9 persons age 70 years were excluded from analyses. Normative values from Mathiowetz et al (reference 16). Nor-Hand = hospital-based observational cohort study; OA = osteoarthritis.
† n = 1 missing for the left hand.

were few current smokers and many participants reported regular physical activity. On the other hand, the prevalence of possible harmful alcohol drinking was high.

Comparison of grip strength in individuals with hand OA versus the general population. In general, women and men from the Nor-Hand study demonstrated lower grip strength than previously observed in the general population across all age groups, especially in the right hand (dominant hand for 92.3% of the study population) in individuals age <60 years (Table 2). The differences between age categories were numerically larger in the general population than in the Nor-Hand study, where a similar decline in grip strength with increasing age was not observed. In both the general population and the Nor-Hand study, men had considerably higher grip strength than women (Table 2). In our study, 96 (36.1%) and 12 (35.3%) of the women and men, respectively, had reduced grip strength according to the frailty criteria of Fried et al (7).

Associations between hand OA and grip strength. Few participants had undergone prior surgery of the CMC1 joint (left: n = 2 [0.7%], right: n = 2 [0.7%], and bilateral: n = 1 [0.3%]). In analyses of the thumb base and ray 1, these hands were treated as missing.

All markers of self-reported hand pain and disease activity were strongly associated with lower grip strength (Table 3). Having 1 or more painful MCP or thumb base joints during the last 6 weeks was associated with lower grip strength (separate models). When pain in MCP joints and thumb base joints was included in the same model, the strengths of associations were weakened for both joint groups (MCP 1–5: β = –1.05 [95% CI –2.15, 0.04] and thumb base: β = –1.68 [95% CI –2.77, –0.58]). Looking at the different rays, a statistically significant association was found for ray 1 only. When pain in the thumb base, MCP1 and first interphalangeal (IP1) joints were included in the same model, we found a statistically significant association with grip strength for the thumb base only (β = –1.71 [95% CI –2.84, –0.59]) and borderline statistically significant association for the MCP1 joint (β = –1.09 [95% CI –2.34, 0.16]).

Table 3. Associations of hand pain and disease activity with grip strength*

	Beta (95% CI)	*P*
NRS disease activity (range 0–10)	–0.98 (–1.37, –0.60)	<0.001
NRS hand pain (range 0–10)	–1.04 (–1.43, –0.65)	<0.001
NRS hand pain after grip strength (range 0–10)	–0.88 (–1.14, –0.61)	<0.001
AUSCAN hand pain (range 0–20)	–0.48 (–0.69, –0.27)	<0.001
Regular use of oral analgesics (no regular use = ref.)	–3.87 (–6.84, –0.89)	0.01
Painful joints previous 6 weeks		
Any painful joint in joint groups (no pain = ref.)		
DIP 2–5	–0.16 (–1.06, 0.74)	0.73
PIP 1–5	–0.44 (–1.49, 0.61)	0.41
MCP 1–5	–1.60 (–2.59, –0.60)	0.002
Thumb base	–2.15 (–3.15, –1.16)	<0.001
Any painful joint in rays (no pain = ref.)		
Ray 1 (thumb base, MCP1, IP1)	–1.55 (–2.52, –0.58)	0.002
Ray 2 (MCP2, PIP2, DIP2)	–0.27 (–1.20, 0.65)	0.57
Ray 3 (MCP3, PIP3, DIP3)	–0.64 (–1.67, 0.39)	0.22
Ray 4 (MCP4, PIP4, DIP4)	–0.78 (–2.41, 0.86)	0.35
Ray 5 (MCP5, PIP5, DIP5)	–1.09 (–2.52, 0.34)	0.14

* Linear regression analyses with generalized estimating equations, adjusted for age, sex, and body mass index, with separate models for each variable. 95% CI = 95% confidence interval; AUSCAN = Australian/Canadian Osteoarthritis Hand Index; DIP = distal interphalangeal; IP = interphalangeal; MCP = metacarpophalangeal; NRS = numerical rating scale; PIP = proximal interphalangeal; ref. = reference.

Increasing radiographic severity in the PIP and thumb base joints was also associated with lower grip strength (Table 4), but the association remained statistically significant for the thumb base only (β = –0.79 [95% CI –1.10, –0.49]) when both joint groups were included in the same model. Analyses focusing on different rays revealed that OA in ray 1 only was statistically significantly associated with lower grip strength, whereas borderline statistically significant associations were observed for rays 3 and 4. When rays 1, 3, and 4 were included in the same model, the association remained statistically significant for ray 1 only (β = –0.35 [95% CI –0.60, –0.10]). Within ray 1, only OA in the thumb base joints was associated with lower grip strength (CMC1: β = –1.09 [95% CI –1.57, –0.62] and STT: β = –0.56 [95% CI –1.05, –0.08]), whereas no statistically significant associations were found for the IP1 and MCP1 joints (data not shown). When both radiographic OA severity in the CMC1 and STT joints and pain in the previous 6 weeks in the thumb base were included in the same model, the strength of associations got weaker, but remained statistically significant for all 3 variables (data not shown).

Interactions with sex were found for NRS disease activity, several pain outcomes (NRS hand pain, NRS hand pain after grip strength, regular use of oral analgesics, painful DIP joints, and painful joints in ray 5, and radiographic OA severity in PIP joints and ray 5). For these disease activity and pain outcomes, the associations were slightly weaker for women and stronger for men (data not shown). A negative association between OA severity in PIP joints and grip strength was observed in women only. The association between OA severity in ray 5 and grip strength was nonsignificant in both sexes (data not shown).

Associations between other biopsychosocial factors and grip strength. Women in the Nor-Hand study had statistically significantly lower grip strength than men, whereas weak and nonsignificant associations with grip strength were observed for increasing age and BMI. A higher comorbidity index, increasing resting heart rate, and low education were associated with lower grip strength. Patients with symptoms of anxiety and/or depression had lower grip strength than those with no such symptoms, but the association did not reach statistical significance (Table 5).

Interactions with sex were found for infrequent physical exercise, smoking, and low education. Whereas no statistically significant association with grip strength was found for infrequent physical activity in all participants, sex-stratified analyses showed conflicting results, with higher grip strength in women and lower grip strength in men with low levels of physical activity (data not shown). A negative association was found between smoking and grip strength in men only (β = –13.59 [95% CI –26.03, –1.14]), and not in the total study sample or women. The negative association between low education and grip strength was stronger in men than women (data not shown).

Table 4. Association between radiographic hand OA severity and grip strength in the same hand*

	Beta (95% CI)	*P*
K/L sum score in joint groups		
DIP 2–5 (range 0–16)	0.00 (–0.17, 0.17)	1.00
PIP 1–5 (range 0–20)	–0.18 (–0.34, –0.03)	0.02
MCP 1–5 (range 0–20)	0.06 (–0.32, 0.44)	0.74
CMC1/STT (range 0–8)	–0.83 (–1.12, –0.53)	<0.001
K/L sum score in rays		
Ray 1 (STT, CMC1, MCP1, IP1) (range 0–16)	–0.39 (–0.63, –0.15)	0.001
Ray 2 (MCP2, PIP2, DIP2) (range 0–12)	–0.12 (–0.52, 0.28)	0.56
Ray 3 (MCP3, PIP3, DIP3) (range 0–12)	–0.29 (–0.58, 0.01)	0.05
Ray 4 (MCP4, PIP4, DIP4) (range 0–12)	–0.28 (–0.58, 0.02)	0.07
Ray 5 (MCP5, PIP5, DIP5) (range 0–12)	0.03 (–0.31, 0.38)	0.85

* Linear regression analyses with generalized estimating equations, adjusted for age, sex, and body mass index, with separate models for each variable. 95% CI = 95% confidence interval; CMC1 = first carpometacarpal; DIP = distal interphalangeal; IP = interphalangeal; K/L = Kellgren/Lawrence; MCP = metacarpophalangeal; OA = osteoarthritis; PIP = proximal interphalangeal; STT = scaphotrapeziotrapezoidal.

Table 5. Associations of other biopsychosocial factors with grip strength*

	Beta (95% CI)	*P*
Age, per 1 year	–0.05 (–0.19, 0.09)	0.51
Female (male = ref.)	–14.8 (–18.5, –11.1)	<0.001
Body mass index, per 1 kg/m^2	–0.14 (–0.35, 0.06)	0.18
Previous hand/wrist injury (no injury = ref.)	–0.17 (–1.02, 0.69)	0.70
Infrequent physical exercise (frequent exercise = ref.)	1.30 (–0.82, 3.42)	0.23
Comorbidity Questionnaire sum score (range 0–45)	–0.31 (–0.51, –0.10)	0.004
Systolic blood pressure, per 1 mm Hg	0.02 (–0.03, 0.06)	0.51
Diastolic blood pressure, per 1 mm Hg	–0.03 (–0.13, 0.06)	0.50
Heart rate at rest, per 1 heart beat	–0.11 (–0.20, –0.02)	0.02
Hospital Anxiety and Depression Scale (range 0–42)	–0.14 (–0.30, 0.02)	0.08
Low education (high education = ref.)	–2.60 (–4.32, –0.88)	0.003
Current daily or nondaily smoking (no smoking = ref.)	–1.58 (–4.16, 1.01)	0.23
High alcohol consumption (AUDIT-C; little alcohol = ref.)	0.59 (–1.21, 2.40)	0.52

* Linear regression analyses with generalized estimating equations, adjusted for age, sex, and body mass index, with separate models for each variable. 95% CI = 95% confidence interval; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption; ref. = reference.

DISCUSSION

In this study of individuals with hand OA, we found that radiographic OA severity, disease activity, and pain were associated with reduced grip strength in the same hand. Pathology in the CMC1 joint seemed to be most important for the grip strength. Furthermore, other factors such as female sex, higher self-reported comorbidity score, high heart rate at rest, and low education were also related to lower grip strength, suggesting that not only hand OA, but also the general health of the individuals may be important for their grip strength.

Individuals with hand OA tend to have a lower grip strength compared with unaffected individuals in the general population. In the population-based Framingham study, elderly subjects with symptomatic hand OA had a 10% reduced grip strength compared to those without the disease (18). We found a reduction in grip strength in individuals with hand OA compared to the general population in both hands and in all age groups. In the general population, the grip strength in the dominant hand was substantially larger than in the nondominant hand, whereas the difference between the 2 hands was numerically smaller in individuals with hand OA. A possible explanation might be that the loss of grip strength is larger in the dominant hand despite hand OA being equally common in both hands (19). Longitudinal studies are needed to compare the loss of grip strength in the dominant hand versus the nondominant hand and to explore potential explanations for a higher loss in the dominant hand. The difference between our study population and the general population was largest in individuals age <60 years. No difference was found in the oldest age category (65–69 years), which is probably due to a high prevalence of hand OA in this age group in the general population as well (19).

In other medical areas, grip strength is mainly used as a marker of frailty. The clinical phenotype of frailty as defined by Fried et al consists of 5 criteria (7). These are weakness (measured by grip strength in the dominant hand, adjusted for BMI), unintentional weight-loss, self-reported exhaustion, slow walking speed, and low physical activity. Individuals who fulfill 3 or more of the criteria are considered frail. In our study, approximately one-third of the participants had reduced grip strength according to the definition by Fried et al. In the original study of Fried et al, the criterion-specific threshold for weakness was set as the lowest 20% when adjusted for sex and BMI. In the EPOSA cohort, elderly individuals with clinical hand OA had a >2-fold increased risk of frailty when adjusted for sociodemographic and health-related variables (20). The increased risk may be due to reduced grip strength related to the presence of hand OA and/or slow walking speed and low physical activity due to accompanying knee pain in individuals with generalized OA. In the Nor-Hand study, we do not have all the variables that are needed to define frailty. Hence, future longitudinal studies are needed to explore associations between frailty and OA. Hypothetically, treating OA pain may be important to avoid the development of frailty, and conversely frailty may affect the prognosis of OA.

The strengths of the Nor-Hand study are the strict measurement procedure of grip strength and the in-depth assessment of hand OA severity, including both pain and structural pathology. The most important joint affecting grip strength in patients with hand OA was the CMC1. We found a statistically significant association between grip strength and both self-assessed pain and radiographic severity in the CMC1 joint. Previous studies have shown conflicting results. In line with our results, the association between the K/L sum score and grip strength was stronger in the CMC1 joints than for the DIP and PIP joints in the Oslo hand OA cohort (n = 190) (21). Similar results were found in a study conducted by Dominic et al (n = 700), who found that impaired grip strength played a more important role for hand disability among individuals with OA in the CMC1 joints than in the MCP and PIP joints (5). On the other hand, Spacek et al (n = 116) found similar levels of grip strength in patients with predominantly thumb base symptoms as in those with predominantly digital symptoms (22).

Dominic et al suggested that the association between CMC1 OA and grip strength could not be fully explained by concomitant hand pain (5). However, they did not have information about pain in the thumb base specifically. In our study, the association between radiographic severity in the CMC1 joint and grip strength was not fully explained by accompanying pain in the thumb base, suggesting that structural pathology in the joint in itself may also affect the grip strength. The OMERACT report recommends hand strength as part of the core domains for all hand OA studies (1). However, current evidence suggests that measurement of grip is most useful in studies focusing on OA in the thumb base joint.

According to treatment guidelines, exercises should be considered in every patient with hand OA (23), because better muscle strength may lead to better function and less pain. Our results support an association between pain and reduced grip strength, but due to the cross-sectional study design we can not draw any conclusions about causality.

In the Nor-Hand study, we thoroughly mapped variables of other biopsychosocial factors to get an impression of patients' general health, making us able to explore the associations of those factors with grip strength. None of the above-mentioned studies looking into the associations between hand OA and grip strength has explored associations between other biopsychosocial factors and grip strength (5,21,22). A higher comorbidity index, increasing resting heart rate (as a potential marker of poor physical fitness), and low education were associated with lower grip strength. Our results are in line with the EPOSA study, in which strong associations with grip strength were found for comorbidities, particularly for cognitive impairment, anxiety, depression, cardiovascular disease, peripheral artery disease, and osteoporosis (9).

In the Nor-Hand study, a borderline statistically significant association was found between anxiety/depression and grip strength, which is in line with a previous study on patients in primary care age >60 years (24). The observed association may be due to more comorbidities and lower physical fitness in individuals with such symptoms. Our results suggest that the patient's general health is important to take into account when evaluating the grip strength in individuals with hand OA.

There are a few limitations to our study. There may be a limited generalizability because most patients with OA are managed in primary care, while this study had a hospital-based design with patients from secondary care. Joint mobility was not assessed, and may contribute to loss of grip strength in individuals with hand OA.

In conclusion, our study found that pathology in the thumb base joint is the most important factor to influence grip strength in individuals with hand OA, and clinical trials focusing on the thumb base joints should include grip strength as an outcome measure. However, other factors than joint-related pathology, such as female sex, physical fitness (i.e., a high resting heart rate), comorbidities, and socioeconomic factors (i.e., low education), might also affect grip strength, and the patient's general health is therefore important to consider when evaluating the grip strength in individuals with hand OA.

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. Haugen 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. Haugen.
Acquisition of data. Haugen.
Analysis and interpretation of data. Haugen, Aaserud, Kvien.

REFERENCES

1. Kloppenburg M, Boyesen P, Smeets W, Haugen IK, Liu R, Visser W, et al. Report from the OMERACT Hand Osteoarthritis Special Interest Group: advances and future research priorities. J Rheumatol 2014;41:810–8.
2. Stamm T, van der Giesen F, Thorstensson C, Steen E, Birrell F, Bauernfeind B, et al. Patient perspective of hand osteoarthritis in relation to concepts covered by instruments measuring functioning: a qualitative European multicentre study. Ann Rheum Dis 2009;68:1453–60.
3. Dreiser RL, Maheu E, Guillou GB, Caspard H, Grouin JM. Validation of an algofunctional index for osteoarthritis of the hand. Rev Rhum Engl Ed 1995;62(6 Suppl 1):43–53s.
4. Slatkowsky-Christensen B, Kvien TK, Bellamy N. Performance of the Norwegian version of AUSCAN: a disease-specific measure of hand osteoarthritis. Osteoarthritis Cartilage 2005;13:561–7.
5. Dominick KL, Jordan JM, Renner JB, Kraus VB. Relationship of radiographic and clinical variables to pinch and grip strength among individuals with osteoarthritis. Arthritis Rheum 2005;52:1424–30.
6. Jones G, Cooley HM, Bellamy N. A cross-sectional study of the association between Heberden's nodes, radiographic osteoarthritis of the hands, grip strength, disability and pain. Osteoarthritis Cartilage 2001;9:606–11.
7. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56:M146–56.
8. Drey M, Pfeifer K, Sieber CC, Bauer JM. The Fried frailty criteria as inclusion criteria for a randomized controlled trial: personal experience and literature review. Gerontology 2011;57:11–8.
9. Siviero P, Zambon S, Limongi F, Castell MV, Cooper C, Deeg DJ, et al. How hand osteoarthritis, comorbidity, and pain interact to determine functional limitation in older people: observations from the European Project on OSteoArthritis Study. Arthritis Rheumatol 2016;68:2662–70.
10. Gløersen M, Mulrooney E, Mathiessen A, Hammer HB, Slatkowsky-Christensen B, Faraj K, et al. A hospital-based observational cohort study exploring pain and biomarkers in patients with hand osteoarthritis in Norway: the Nor-Hand protocol. BMJ Open 2017;7:e016938.
11. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
12. Kurtze N, Rangul V, Hustvedt BE, Flanders WD. Reliability and validity of self-reported physical activity in the Nord-Trøndelag Health Study: HUNT 1. Scan J Public Health 2008;36:52–61.
13. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 2003;49:156–63.
14. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP): alcohol use disorders identification test. Arch Intern Med 1998;158:1789–95.
15. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983;67:361–70.
16. Mathiowetz V, Kashman N, Volland G, Weber K, Dowe M, Rogers S. Grip and pinch strength: normative data for adults. Arch Phys Med Rehabil 1985;66:69–74.
17. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. Arthritis Rheum 1990;33:1601–10.
18. Zhang Y, Niu J, Kelly-Hayes M, Chaisson CE, Aliabadi P, Felson DT. Prevalence of symptomatic hand osteoarthritis and its impact on functional status among the elderly: the Framingham Study. Am J Epidemiol 2002;156:1021–7.
19. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis 2011;70:1581–6.
20. Castell MV, van der Pas S, Otero A, Siviero P, Dennison E, Denkinger M, et al. Osteoarthritis and frailty in elderly individuals across six European countries: results from the European Project on OSteoArthritis (EPOSA). BMC Musculoskelet Disord 2015;16:359.
21. Haugen IK, Slatkowsky-Christensen B, Bøyesen P, van der Heijde D, Kvien TK. Cross-sectional and longitudinal associations between radiographic features and measures of pain and physical function in hand osteoarthritis. Osteoarthritis Cartilage 2013;21:1191–8.
22. Spacek E, Poiraudeau S, Fayad F, Lefèvre-Colau MM, Beaudreuil J, Rannou F, et al. Disability induced by hand osteoarthritis: are patients with more symptoms at digits 2–5 interphalangeal joints different from those with more symptoms at the base of the thumb? Osteoarthritis Cartilage 2004;12:366–73.
23. Kloppenburg M, Kroon FP, Blanco FJ, Doherty M, Dziedzic KS, Greibrokk E, et al. 2018 update of the EULAR recommendations for the management of hand osteoarthritis. Ann Rheum Dis 2019;78:16–24.
24. Lino VT, Rodrigues NC, O'Dwyer G, Andrade MK, Mattos IE, Portela MC. Handgrip strength and factors associated in poor elderly assisted at a primary care unit in Rio de Janeiro, Brazil. PLoS One 2016;11:e0166373.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 801–809
DOI 10.1002/acr.24169

American College of Rheumatology
Empowering Rheumatology Professionals

Development and Underlying Structure of a Second-Generation Appropriateness Classification System for Total Knee Arthroplasty

Antonio Escobar Martinez,[†] Robert A. Perera, and Daniel L. Riddle

Objective. Evidence supports the validity of total knee arthroplasty (TKA) appropriateness classification based on the RAND Corporation and University of California Los Angeles system.The system, however, is ~2 decades old, and the science and clinical application of TKA has changed dramatically. We undertook this study to describe the methods used to develop a second-generation system and to examine the structure of the system to determine the extent to which each of the indication criteria informed appropriateness.

Methods. Multivariable multinomial regression analyses determined the extent to which each of the 8 individually analyzed indication criteria informed appropriateness judgments. Classification tree analysis illustrates how the indication criteria, in combination, led to judgments of appropriate, inappropriate, or uncertain.

Results. An expert panel selected 8 indication criteria (i.e., age, knee pain, function, radiographic osteoarthritis severity, osteoarthritis location, psychological factors, pain catastrophizing, and comorbidities). A total of 1,008 clinical scenarios were written, based on the criteria. Regression analyses indicated that age, knee pain, function, and radiographic severity dominated prediction of appropriateness, while the other criteria played a smaller role. Classification tree analysis confirmed the regression findings.

Conclusion. Our second-generation classification system, which incorporates contemporary indicators of TKA prognosis and risk, demonstrated preliminary evidence for utility in clinical practice.

INTRODUCTION

The RAND Corporation and University of California Los Angeles (RAND/UCLA) system for classifying the extent of appropriate use of surgical procedures (1) has been applied to many surgical procedures, including total knee arthroplasty (TKA), for more than 3 decades (2). The basic premise that drives the need for RAND/UCLA appropriateness criteria is that randomized clinical trials comparing surgical to nonsurgical treatments (or placebo surgery) for many surgical interventions are lacking or are not feasible. While this has recently begun to change for some surgical interventions (3,4), other surgical treatments lack multiple large-scale trials to inform practice decisions, and TKA is one example. We found 1 randomized trial of 100 participants comparing TKA to a nonsurgical approach (5). The RAND/UCLA appropriateness criteria also have the potential to improve general clinical practice and practice guidelines for knee OA.

One of the earliest applications of the RAND/UCLA system to TKA appropriateness classification was by Escobar and colleagues as shown in the study by Katz (6). Despite evidence suggesting that the Escobar system has predictive utility using contemporary data (7,8), concerns about limited shelf life have been expressed by our team (9) and others (10). The system developed by Escobar and colleagues was based on an evidence synthesis that is ~2 decades old and, given dramatic growth in TKA technologies, evidence, and indications (11), an updated RAND/UCLA system is needed. The shortcomings of a recently developed RAND/UCLA–based system for TKA (12), and the continued high utilization of TKA (13) in the face of a 20% rate of persistent knee pain following surgery (14), amplify the need for an alternative.

The content herein is solely the responsibility of the authors and does not necessarily represent the official views of the Department of Health of the Basque Country.

Supported by the Fondo de Investigación Sanitaria (grant PI15/01105), the Department of Health of the Basque Country (grant 2015111137), the Research Commission of the Bilbao-Basurto Integrated Health Organization, and the European Regional Development Fund.

Robert A. Perera, PhD, Daniel L. Riddle, PT, PhD, FAPTA: Virginia Commonwealth University, Richmond.

[†]Dr. Escobar Martinez is deceased.

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

Address correspondence to Daniel L. Riddle, PT, PhD, FAPTA, Departments of Physical Therapy, Orthopaedic Surgery and Rheumatology, Virginia Commonwealth University Richmond, VA 23298-0224. Email: dlriddle@vcu.edu.

Submitted for publication October 18, 2019; accepted in revised form February 11, 2020.

SIGNIFICANCE & INNOVATIONS

- This second-generation RAND–based knee arthroplasty appropriateness classification system is the first to incorporate contemporary indicators of both prognosis and risk.
- Preliminary evidence suggested strong potential for clinical application if prospective application on a large sample of patients undergoing knee arthroplasty demonstrates prediction of outcome.
- The proposed system improves upon previously reported systems based on the RAND system of classification.

The RAND/UCLA system for determining appropriateness of medical procedures requires a resource-intensive process with multiple steps. First, an appointed panel of content experts (i.e., panel 1) conducts an exhaustive literature review for purposes of identifying and defining key patient-related variables (i.e., indication criteria) that drive appropriateness for the intervention of interest. Next, the panel writes a comprehensive set of patient scenarios capturing all permutations of key patient-related variables. A second independent panel of experts (i.e., panel 2) then rates the appropriateness of each scenario during a series of Delphi-like sessions. A 1–9 appropriateness rating scale is used by panel 2 for each scenario with a median score of 1–3 indicating an inappropriate scenario, a median rating of 4–6 indicating an uncertain scenario, and a 7–9 median rating indicating an appropriate scenario.

In 2016, our team began a revision to the RAND/UCLA TKA appropriateness classification system first published by members of our team in 2003 (15). The purposes of the present study are to describe the methods used to develop the revised version of our TKA appropriateness classification system, and to quantitatively examine the structure of the new system. We hypothesized that in addition to traditional variables of pain, function, age, and knee osteoarthritis (OA) severity, prognostic variables that are not traditionally included in appropriateness systems (15,16) (including psychological health, pain catastrophizing, and comorbidity) would associate with appropriateness classification.

MATERIALS AND METHODS

We followed recommendations originally developed by Brook et al (17) and fully described in the RAND/UCLA user's manual (1) to develop the second-generation appropriateness criteria for TKA. Our second-generation system addresses poor outcome risk more directly than the first-generation system (15) and was built independently of the first-generation system.

Assumptions for the expert panels. Prior to the initiating of the RAND/UCLA process, the expert panels were instructed to assume the following: 1) only unilateral primary TKA for OA for unilateral symptoms would be considered, 2) if hip OA was also present, it would be appropriately managed, 3) prior surgical treatment and prior and current medications for knee OA were appropriately managed, 4) active infection, quadriceps tendon rupture, or severe peripheral vascular disease had effectively been ruled out, 5) expectations regarding outcome following TKA was appropriately addressed, 6) social support (e.g., caregiver or spouse/partner) for the TKA recipient would be available, and 7) body mass index was <40 kg/m^2.

Expert panel 1. Expert panel 1 included 6 orthopedic surgeons, 1 psychologist, 1 statistician, 1 social worker, and 1 epidemiologist. This panel was identified by 1 author (AE-M) and his research team in Spain. The first step was to conduct a comprehensive review of the scientific literature to synthesize available evidence about the use, efficacy, opinions, outcomes, and risks of TKA. The panel synthesized the available scientific evidence about TKA from January 1, 2000 to February 1, 2016 in EMBASE (OVID) and Medline (PubMed) databases (for search strategy, see Supplementary Appendix A, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24169/abstract). The purpose of the literature review was to provide members of expert panel 1 with an up-to-date source of scientific content to identify key indication criteria that would serve as the basis for writing the scenarios.

There were a total of 8 indication criteria identified by expert panel 1 based on the literature review (see Table 1). The criteria were chosen based on the literature review and the experience of the expert panel 1 clinicians involved in the treatment of these patients. The final decision on criteria inclusion and exclusion was unanimous. Of these criteria, 5 had dichotomous responses, 2 had trichotomous responses, and 1 had 4 responses. There was discussion about other factors such as body mass index, social support, impact on sleep, mental preparation, or patient expectations; however, the expert panel 1 did not find these variables to be as prognostically important as the 8 selected criteria. The definitions for the response cut points for the 8 indication criteria were developed by expert panel 1 prior to writing the scenarios. Cut point thresholds were based on the literature review.

For the pain and functional limitation criteria, the Spanish-validated Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain and function subscales (18), respectively, were chosen by expert panel 1. The WOMAC scales each ranged from 0 to 100, with higher scores equating to worse pain or function. Each had 3 categories (i.e., slight [scores <35], moderate [scores 35–50 for pain and 35–54 for function], and severe [scores >50 for pain and >54 for function]). For psychological factors, the panel chose to use the recommended cut point of 10 to differentiate between those patients with clinical anxiety and depression versus those without, as measured by the Spanish-validated version (19) of the Hospital Anxiety and Depression Scale (20). For pain catastrophizing, the panel chose a cut point of 30 to dichotomize low versus high catastrophizing based

Table 1. Indication criteria measures and appropriateness ratings of 1,008 vignettes for the revised total knee arthroplasty appropriateness rating system*

Indication criteria measurement scale	Scenario sample size
Age group, years	
<55	288 (28.6)
55–65	288 (28.6)
>66–85	288 (28.6)
>85	144 (14.3)
Radiology (K/L grade)	
≤2	504 (50.0)
≥3	504 (50.0)
Knee OA localization	
Unicompartmental†	504 (50.0)
Multiple compartment†	504 (50.0)
WOMAC pain‡	
Slight (<35)	336 (33.3)
Moderate (35–50)	336 (33.3)
Severe (>50)	336 (33.3)
WOMAC function‡	
Slight (<35)	336 (33.3)
Moderate (35–54)	336 (33.3)
Severe (>54)	336 (33.3)
Psychological factors (HADS anxiety or depression)	
≤10 in anxiety and depression	504 (50.0)
>10 in anxiety or depression	504 (50.0)
Pain catastrophizing	
≤30	504 (50.0)
>30	504 (50.0)
Comorbidities	
No	504 (50.0)
≥1§	504 (50.0)
Appropriateness rating	
Inappropriate	671 (66.6)
Uncertain	256 (25.4)
Appropriate	81 (8.0)

* Values are the number (%). HADS = Hospital Anxiety and Depression Scale; K/L = Kellgren/Lawrence; OA = osteoarthritis; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† For vignettes with age >85 years (n = 144), the expert panel indicated that unicompartmental or multicompartmental disease would be treated in the same way.
‡ Scores range from 0 (best) to 100 (worst).
§ Relevant comorbidities included osteoporosis, asthma, chronic obstructive pulmonary disease, heart failure, Parkinson's disease, multiple sclerosis, stroke, and chronic back pain.

on the Spanish-validated Pain Catastrophizing Scale (21–23). For comorbidity, the panel chose a cut point of 0 to differentiate between persons with ≥1 comorbidity and those without (Table 1). For the radiology criterion, a Kellgren/Lawrence (K/L) grade cut point of 2 was used. Knee OA localization was dichotomized to either unicompartmental or multicompartmental disease.

Expert panel 1 used the 8 criteria to write a series of brief clinical vignettes that incorporated the 8 criteria. The scenarios had to be mutually exclusive and exhaustive as well as clinically feasible for persons considering TKA. The intention was to cover virtually all clinical patterns of knee problems for which TKA might be considered. Combining these variables in a factorial design would result in n = 1,152 scenarios ($2^5 \times 3^2 \times 4^1$). The expert panel concluded that for scenarios with patients age >85 years, the knee OA location (i.e., number of compartments affected) did not influence classification, and therefore 1,008 scenarios were written.

Expert panel 2. A national panel of clinical experts from Spain was selected. The research team asked the Spanish Knee Society to nominate nationally recognized specialists in TKA care. After the initial contact, 12 specialists, 10 orthopedic surgeons, and 2 rheumatologists agreed to participate. Panelists were provided with the literature review, the list of indication criteria, and the 1,008 scenarios. Each member of expert panel 2 was asked to rate each scenario for the appropriateness of TKA, taking into consideration the average patient and average physician in 2016.

Appropriate was defined as when the expected health benefit of TKA exceeded the expected negative consequences by a sufficiently wide margin to make TKA worth performing. Inappropriate was defined as the risk of negative consequences outweighing expected benefits, and uncertain was defined either as the situation in which benefit and risk could not be estimated or where benefits and risks were about equally balanced (17). The indication criteria and scenarios were reviewed and approved by expert panel 2 prior to scenario rating.

Scenario ratings by expert panel 2. Ratings took place during 2 rounds using a modified Delphi method (1). The first round was performed at the individual level, and the second round during a one-day in-person panel meeting in which each panelist received the results of his/her own scores for each scenario and the anonymized ratings made by the other members. After discussion of scenarios with disagreement during round 1 (n = 27), and briefly, all other scenarios as well as a review of the scenario scoring method, the panelists were able to revise their ratings. The aim was not to reach consensus but to identify the level of agreement among the participants after discussion.

Ratings were scored on a 9-point scale. The medians for each scenario could lie within the ranges of 1–3 for a rating of inappropriate, 4–6 for a rating of uncertain, and 7–9 for appropriate. Use of TKA was considered appropriate if the panel's median rating was between 7 and 9 without disagreement (defined below), uncertain if the panel's median rating was between 4–6 without disagreement. A scenario also was scored as uncertain for any median score in which there was disagreement (defined below). A rating of inappropriate was given if the median rating was between 1 and 3 without disagreement. Agreement was established when one-third or less of the panelists' scores (i.e., ≤4) occurred in any 3-point range outside the median categorical appropriateness rating. Disagreement occurred when ≥4 ratings for a scenario fell at the extremes of the scale (i.e., 4 ratings of 1, 2, or 3, and 4 ratings of 7, 8, or 9). Scenarios were rated as indeterminate when criteria for either agreement or disagreement were not met. Please see Fitch et al for a more detailed description of appropriateness rating for scenarios (1).

Data analysis. To determine the relative strength of association for indication criteria with appropriateness ratings, multivariable ordinal logistic regression was used. All indication variables were included in the model, with estimates of the coefficients and their corresponding odds ratios being reported. To confirm findings of the ordinal logistic model, an ordinary least squares (OLS) regression was conducted using the median expert rating. Given that the results of the expert panel classifications represent the entire population of scenarios, all estimates reported are the population values. Therefore, no SEs, confidence intervals, or *P* values are required. Due to the age category of >85 years not being crossed with the number of knee OA compartments, a sensitivity analysis was completed for each model excluding scenarios where the age was >85 years. In every model, the independent variables were treated as categorical. This accounts for potential nonlinear relationships and allows for direct comparison of estimates between variables. All models were fit using SAS software, version 9.4.

To determine the best combination of indication criteria for predicting each of the appropriateness ratings, we also used a decision tree approach (i.e., Exhaustive Chi-Squared Automatic Interaction Detection [Exhaustive CHAID]). This approach allows for polytomous predictor variables (e.g., age, WOMAC scores) unlike classification and regression trees, which require dichotomous predictor variables. Exhaustive CHAID is a nonparametric approach that iteratively tests each of the 8 predictor variables to find the variable that most strongly associates with appropriateness classification. Once the most predictive variable (i.e., variable with the highest chi-square result) is found, all possible splits for this branch of the tree are examined, and the next most predictive variable is identified. Only variables that improve prediction relative to the more proximal branch are included. Variable inclusion in the tree is based on Bonferroni corrected chi-square estimates with $P < 0.05$. The goal with decision trees is to create pure terminal nodes (i.e., child nodes) for each branch of the tree in a parsimonious way. Branches of a tree that lead to additional branches are termed parent nodes.

The approach is iterative to the extent that the goal is to build a tree that both enhances the ability to correctly predict appropriateness classification and that is parsimonious. We studied the entire population of scenarios, which eliminated the need for cross validation. Because of the small number of scenarios classified as appropriate (n = 81), we set the tree depth at 5 levels, with each parent node set to a minimum of 25 scenarios and each child node at a minimum of 15 scenarios. We were purposely liberal in setting the child node limit to create the best chance of building pure nodes without creating nodes with a very small number of

Table 2. Ordinal logistic regression model with complete sample with inappropriate as the reference group*

Prognostic variable	Estimate	Odds ratio
Intercept		
Uncertain	–24.48	
Appropriate	–34.45	
Age group, years		
<55	Ref.	95.80
55–65	4.56	>999.99
66–85	9.21	>999.99
>85	6.97	
WOMAC pain		
Slight	Ref.	
Moderate	8.10	>999.99
Severe	16.88	>999.99
WOMAC function		
Slight	Ref.	
Moderate	5.22	185.18
Severe	10.76	>999.99
K/L grade		
≤2	Ref.	
3 or 4	6.79	884.68
Compartment		
Unicompartment OA	Ref.	
Multicompartment OA	4.32	75.26
Comorbidities		
None	Ref.	
≥1	–3.05	0.05
Anxiety and/or depression		
No	Ref.	
Yes	–1.84	0.16
Pain catastrophizing scale		
≤30	Ref.	
>30	–2.97	0.05

* K/L = Kellgren/Lawrence; OA = osteoarthritis; Ref. = reference; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

Table 3. Ordinary least squares regression using expert panel median scores as the dependent variable*

Prognostic variable, ratings	Estimate
Intercept	–24.87
Age group, years	
<55	Ref.
55–65	0.78
66–85	1.56
>85	1.04
WOMAC pain	
Slight	Ref.
Moderate	1.18
Severe	3.12
WOMAC function	
Slight	Ref.
Moderate	0.76
Severe	2.01
K/L grade	
2 or less	Ref.
3 or 4	1.26
Compartment	
Unicompartment OA	Ref.
Multicompartment OA	0.74
Comorbidities	
None	Ref.
≥1	–0.52
Anxiety and/or depression	
No	Ref.
Yes	–0.23
Pain catastrophizing scale	
≤30	Ref.
>30	–0.42

* See Table 2 for definitions.

scenarios (i.e., n < 15). We used IBM SPSS, version 25, for all decision tree analyses.

RESULTS

A total of 1,008 scenarios were classified as appropriate, uncertain, or inappropriate by expert panel 2. A total of 81 scenarios (8%) were rated as appropriate, 256 (25.4%) as uncertain, and 671 (66.6%) as inappropriate for TKA. There were no scenario scores with disagreement or with indeterminate agreement. Scenario median scores were used in the analyses.

Associations between indication criteria and appropriateness ratings. Multivariable ordinal logistic regression coefficients demonstrated a similar pattern when comparing appropriate to inappropriate ratings and uncertain to inappropriate ratings (Table 2). The odds ratios indicated that older age, moderate and severe WOMAC pain, severe WOMAC function, and grade 3 or 4 K/L grade were positively associated with appropriate and uncertain ratings as compared to inappropriate ratings, with odds ratios all exceeding 880. The number of knee compartments affected had relatively less impact on classification. In contrast, odds ratios for comorbidities, psychological factors, and pain catastrophizing were negatively associated with ratings of appropriate and uncertain. However, the magnitude of these odds ratios indicates a smaller relative impact on classification as compared to age, WOMAC scores, and K/L grade. A sensitivity analysis with the age >85 years scenarios removed essentially replicated findings for the full sample of scenarios (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24169/abstract). Furthermore, the OLS regression model using median

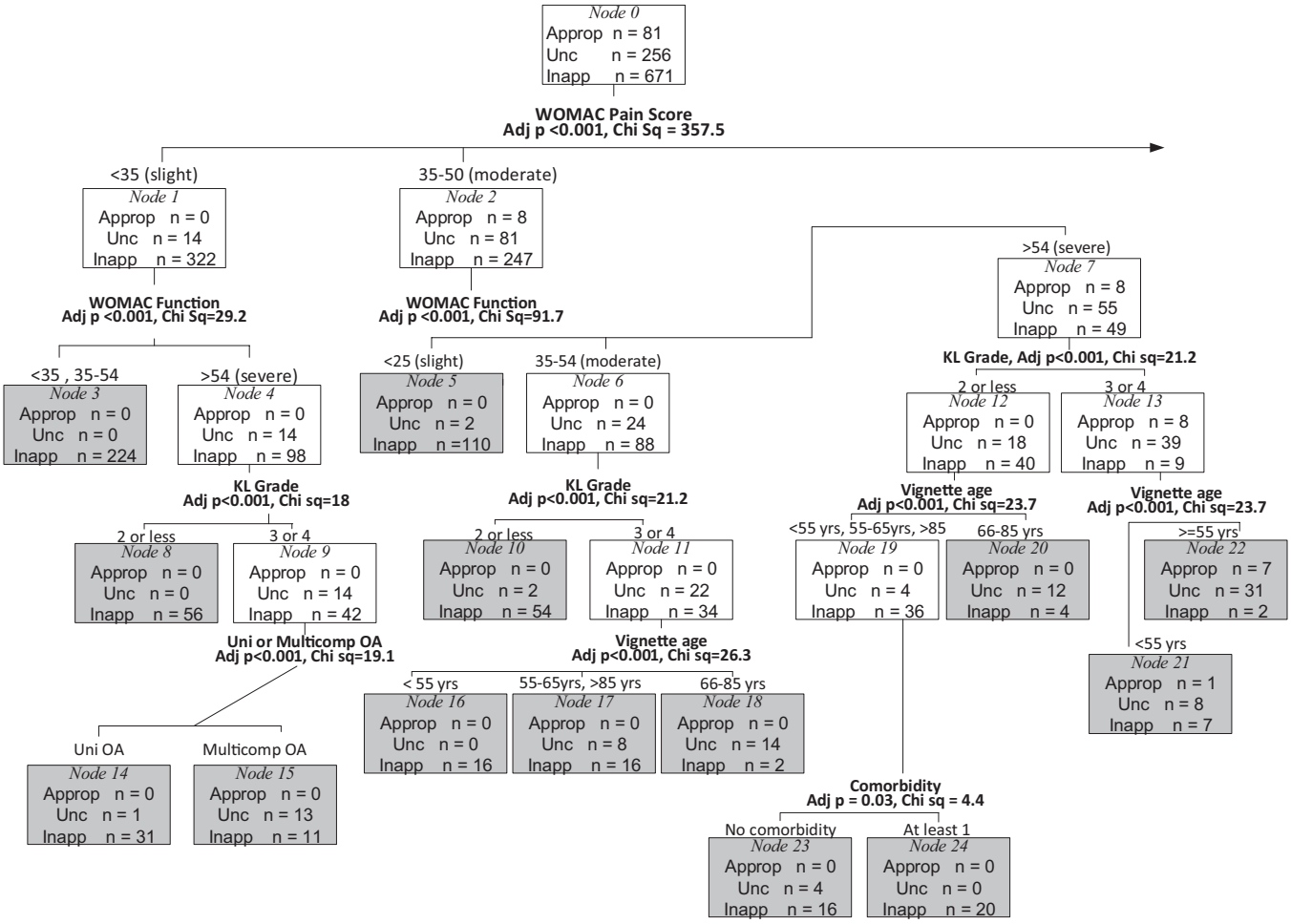

Figure 1. Results for the left side of the classification tree analysis for trichotomous judgments of appropriate, inappropriate, or uncertain. The branches of the tree are labeled based on the key variables forming the tree branches, which are, in order, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain, WOMAC function, Kellgren/ Lawrence (K/L) grade, age, knee osteoarthritis (OA) compartment, pain catastrophizing, and comorbidity. The terminal nodes of each branch (e.g., nodes 3, 8, 14) indicate the final distributions of expert panel ratings of appropriate, inappropriate, and uncertain. Scenario sample sizes are reported in each box. The ratings with the largest sample size in the terminal nodes (shaded) reflect the predicted category for each terminal node. Adj p = adjusted *P* value; Approp = appropriate; Chi sq = chi-square; Unc = uncertain; Inapp = inappropriate; Uni OA = unicompartmental knee OA; Multicomp OA = multicompartmental knee OA.

panel ratings demonstrated comparable findings to the ordinal logistic models (Table 3) with similar patterns of associations between the indication criteria and classification. The corresponding sensitivity analysis with the age >85 years scenarios removed confirmed these findings (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24169/abstract).

Decision tree findings. The decision tree for trichotomized ratings of appropriate, uncertain, and inappropriate appear in Figures 1–3. Because the tree was large, the figures illustrate the left side (i.e., Figure 1), center (i.e., Figure 2), and right side (i.e., Figure 3) of the entire tree. The indication criteria variables, in order of importance, were WOMAC pain, WOMAC function, K/L grade, and age, followed by several other indication criteria. These findings generally echo the findings from the regression analyses. The overall accuracy of the decision tree was 86.7%, with 90.6% of inappropriate scenarios, 86.3% of uncertain scenarios, and 55.6% of appropriate scenarios correctly classified. The extent of agreement with the predicted as compared to actual classification was weighted $\kappa = 0.77$ (95% confidence interval 0.74–0.81). As can be seen in Figure 1, some terminal nodes are pure, suggesting perfect prediction of the tree, while other nodes are mixed. For example, node 3 is a pure node with all 224 scenarios classified as inappropriate when WOMAC pain is <35 and WOMAC function is ≤54. Node 15 is a mixed terminal node with 11 inappropriate and 13 uncertain scenarios.

DISCUSSION

Previously published methods for determining TKA appropriateness either do not include variables that increase poor outcome risk (12,15,24) or have not been evaluated prospectively to determine if classification influenced the outcome (12,25). Our second-generation RAND/UCLA-based system was designed to address these former limitations by incorporating contemporary measures of poor TKA outcome risk, such as the presence of clinical depression (26), high pain catastrophizing (27), and comorbidity (28). The original appropriateness system (15) and the current system are

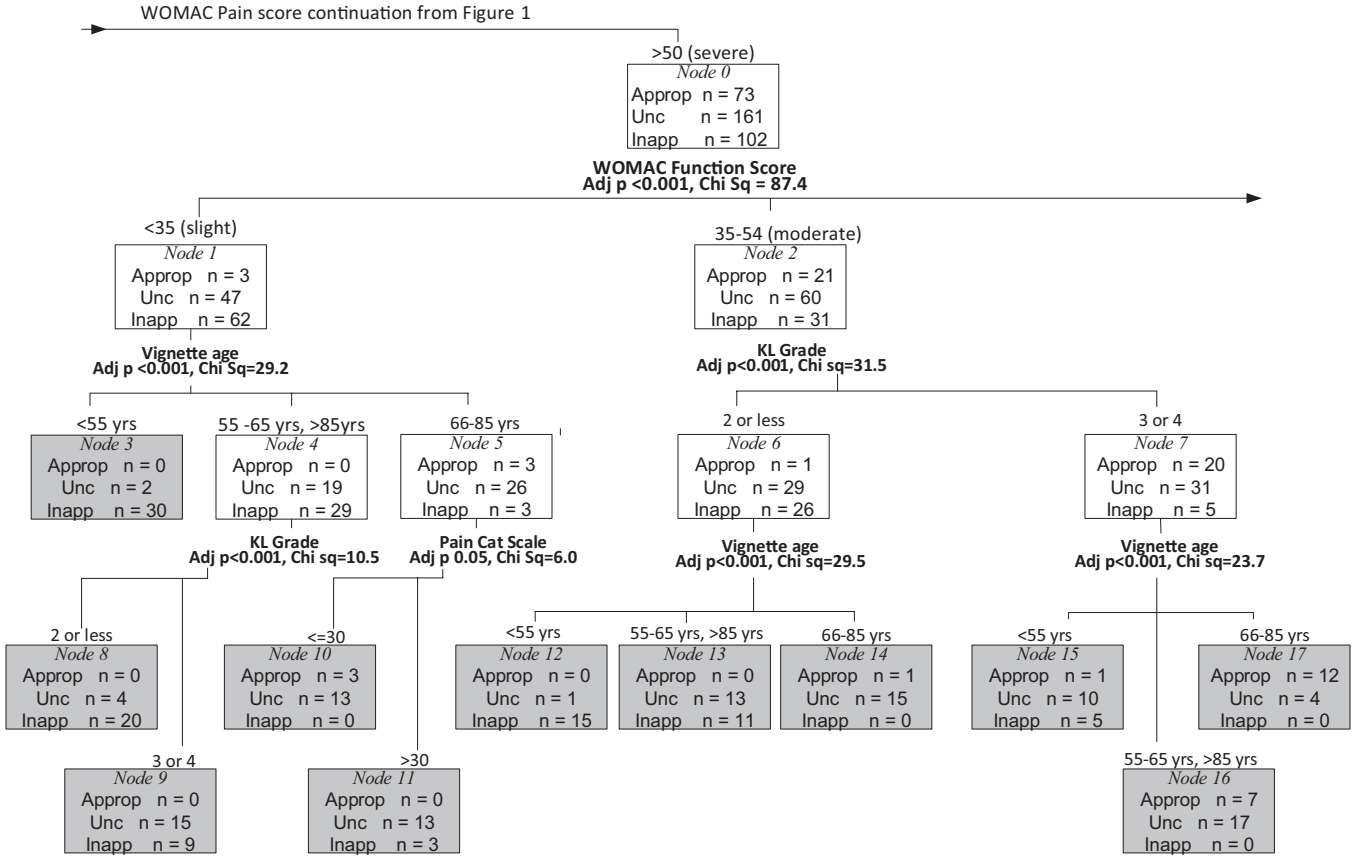

Figure 2. Results for the center of the classification tree analysis for trichotomous judgments of appropriate, inappropriate, or uncertain. The branches of the tree are labeled based on the key variables forming the tree branches, which are, in order, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain, WOMAC function, Kellgren/Lawrence (K/L) grade, age, knee osteoarthritis (OA) compartment, pain catastrophizing, and comorbidity. The terminal nodes of each branch (e.g., nodes 3, 8, 10) indicate the final distributions of expert panel ratings of appropriate, inappropriate, and uncertain. Scenario sample sizes are reported in each box. The ratings with the largest sample size in the terminal nodes (shaded) reflect the predicted category for each terminal node. See Figure 1 for definitions.

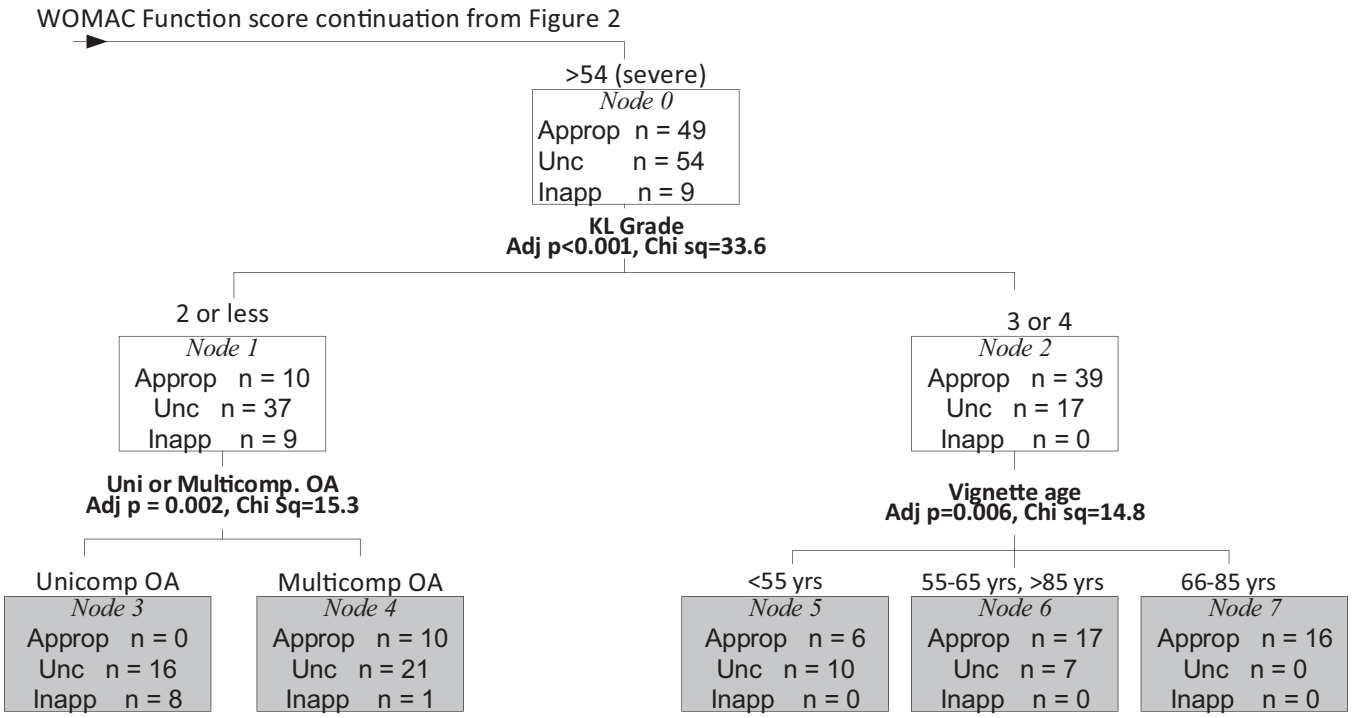

Figure 3. Results for the right side of the classification tree analysis for trichotomous judgments of appropriate, inappropriate, or uncertain. The branches of the tree are labeled based on the key variables forming the tree branches, which are, in order, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain, WOMAC function, Kellgren/Lawrence (K/L) grade, age, knee osteoarthritis (OA) compartment, pain catastrophizing, and comorbidity. The terminal nodes of each branch (e.g., nodes 3, 4, 5) indicate the final distributions of expert panel ratings of appropriate, inappropriate, and uncertain. Scenario sample sizes are reported in each box. The ratings with the largest sample size in the terminal nodes (shaded) reflect the predicted category for each terminal node. See Figure 1 for definitions.

compared in Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24169/abstract.

The development of our system closely followed recommendations described in the RAND/UCLA user's manual. We recruited 2 independent sets of experts who were skilled in the diagnosis and treatment of persons potentially eligible for TKA. Our set of 1,008 scenarios was comprehensive, but most were found by expert panel 2 to be classified as inappropriate (n = 671 [66.6%]), while only 81 (8%) were deemed appropriate for TKA. Of the 624 scenarios in our previous RAND/UCLA study of TKA appropriateness (15), 167 (26.8%) were rated as appropriate, 304 (48.7%) were rated as inappropriate, and the remaining 153 (24.5%) were rated as uncertain. Additionally, a larger number of scenarios in the 2003 study, relative to the current study, were eliminated because they were considered to be clinically implausible. We suspect that the greater percentage of scenarios rated as inappropriate in the current analysis relative to the 2003 study was also due to the inclusion of additional criteria associated with poor outcome (i.e., psychological factors, pain catastrophizing, and comorbidity). A total of 882 scenarios in our study had ≥1 of these criteria scored as positive, and 504 had ≥2 of these criteria scored as positive. For scenarios with 2 positive criteria associated with a poor outcome, 4.6% of these scenarios were rated as appropriate. In 126 scenarios, all 3 poor outcome criteria were positive, and 2.4% of these scenarios were rated as appropriate.

The incorporation of criteria that were associated with poor outcome likely led to a lower percentage of scenarios rated as appropriate relative to prior work that did not incorporate these criteria. These data suggest that when clinical experts are provided with data that inform both the benefits and risks of TKA surgery, appropriateness ratings are likely to be influenced, particularly when multiple poor outcome risk criteria are included in the scenarios. These data also suggest that contemporary appropriateness criteria may actually have stronger utility for identifying cases classified as inappropriate. Given that 20% of patients have a poor outcome following TKA (14,29), we suggest that appropriateness criteria should be particularly suited toward identification of patients inappropriate for TKA and at risk for poor outcome.

The most important predictors of appropriateness ratings in the primary analyses were advanced age, more severe pain and functional limitation, and more severe knee OA. These findings were replicated in the sensitivity analyses. Clinical experts continue to place the greatest emphasis on these traditionally important indicators of knee arthroplasty appropriateness, much like the findings in the 2003 study by Escobar et al (15) and in more contemporary appropriateness studies (12,25). Newer additions to

the criteria, including comorbidity, psychological factors, and pain catastrophizing also contributed but in a minor way. These variables, when positive, were associated with classifications of inappropriate in the regression analyses but were much less important in driving classification as compared to the more traditional variables. In the decision tree analyses, comorbidity and pain catastrophizing entered the tree but only in the final branching, while psychological factor criterion did not contribute to the trees.

It is likely that the decision trees generated in these analyses oversimplified the relationships among the criteria, and this was likely driven by the relatively small number of criteria classified as appropriate. When considering the classification tree in total, only 55.6% of scenarios rated as appropriate were correctly predicted, while 90.6% of inappropriate scenarios were correctly predicted in the decision tree. Of the 81 scenarios classified by the expert panel as appropriate, the tree correctly predicted 45 of these cases, while the other 36 were classified as uncertain by the tree method. These data suggest that the decision tree shown in Figures 1–3 may not be useful for application to clinical practice. External validation, ideally on multiple clinical samples, is a necessary step prior to clinical application.

In a previous study (9), we argued for updating appropriateness systems and for external validation prior to clinical application of any system proposed for judging appropriateness. For implementation, we proposed that any appropriateness system could inform the surgeon and patient regarding risk and benefit, but in the end there may be strong justification for deviating from the appropriateness system finding (9). For example, a patient may be found to be inappropriate for TKA using an appropriateness system, but the patient may provide compelling reasons for the surgery. For example, the patient may argue that they can no longer work and provide family financial support because of limitations in knee function. After a thorough discussion of benefits and risks with the surgeon, both may endorse the need for surgery.

Future testing of our second-generation appropriateness system on patient samples could be done using a less parsimonious but more accurate approach than the tree (Figures 1–3). By expanding the major branches of the trees (Figures 1–3) and setting the tree depth to 6 levels, with each parent node set to a minimum of 10 scenarios and each child node to a minimum of 5 scenarios, we generated new trees. The overall accuracy of this approach is 90.3% in contrast to the larger tree with an overall accuracy of 86.7%. More importantly, the exploding tree approach led to no child nodes with both appropriate and inappropriate scenarios. The expanded trees for each major branch along with classification accuracy tables are illustrated in Supplementary Figures 1–6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24169/abstract.

Our study has some notable limitations. First and most importantly, we only examined the ratings by clinical experts using case scenarios. From these data alone, it is unclear how the findings generalize to patients considering TKA and whether outcome is influenced by classification. Second, our decision tree analysis was limited by the relatively small number of cases rated as appropriate for TKA. The false classification rate for these cases in the decision tree using trichotomous ratings was high, at 44.4%. Third, expert panel 2 had mostly orthopedic surgeons (i.e., 10 of 12 panelists), and single specialty panels are not recommended because it is believed that they increase risk of biased ratings (1). Finally, this system was developed in Spain, and the extent to which the data might generalize to clinicians from other countries with different health care systems is unknown.

In conclusion, we used the well-established RAND/UCLA system to develop an appropriateness classification system that, in theory, overcomes limitations of prior systems. Specifically, our system included evidence-based indicators of poor outcome risk along with more traditional indicators of TKA candidacy. Using 1,008 clinical scenarios and 8 indication criteria, we found that more traditional indicators of knee pain, functional limitation, OA severity, and age played major roles in determining appropriateness. Comorbidity, depression, anxiety, and pain catastrophizing, while contributing to classification, played minor roles relative to more traditional indicators.

ACKNOWLEDGMENTS

We thank the following orthopedic surgeons for their panel 2 participation: Drs. Elvira Montáñez Heredia, Hospital Universitario Virgen de la Victoria, Malaga, Spain; Pedro Hinarejos Gomez, Parc de Salut Mar, Barcelona, Spain; Vicente J. León Muñoz, Hospital de la Vega Lorenzo Girau, Murcia, Spain; María Luisa Bertrand Garcia, Hospital Costa del Sol, Malaga, Spain; Oscar Ares Rodriguez, Hospital Clinic, Barcelona, Spain; Francisco Javier Montilla Jimenez, Hospital Viamed Santa Angela de la Cruz, Sevilla, Spain; Joan Leal Blanquet, Parc de Salut Mar, Barcelona, Spain; Javier Martinez Martin, Fundacion Hospital Alcorcon, Madrid, Spain; Jose Luis Pais Brito, Hospital Universitario de Canarias Tenerife, Spain; José Antonio Guerrero Molina, Hospital Universitario Cruces, Vizcaya, Spain. We also thank the rheumatologists, Drs. Jordi Monfort Faure, Parc de Salut Mar, Barcelona, Spain and Elena Garmendia Sánchez, Hospital Universitario Cruces, Vizcaya, Spain.

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. Riddle 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. Escobar Martinez, Riddle, Perera.
Acquisition of data. Escobar Martinez, Riddle, Perera.
Analysis and interpretation of data. Escobar Martinez, Riddle, Perera.

REFERENCES

1. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. The RAND/UCLA Appriateness Method User's Manual. 2001. URL: http://www.rand.org/pubs/monograph_reports/MR1269.html.
2. Lawson EH, Gibbons MM, Ko CY, Shekelle PG. The appropriateness method has acceptable reliability and validity for assessing overuse and underuse of surgical procedures. J Clin Epidemiol 2012;65:1133–43.

3. 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.
4. 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.
5. Skou ST, Roos EM, Laursen MB, Rathleff MS, Arendt-Nielsen L, Simonsen O, et al. A randomized, controlled trial of total knee replacement. N Engl J Med 2015;373:1597–606.
6. Katz JN. Appropriateness of total knee arthroplasty [editorial]. Arthritis Rheumatol 2014;66:1979–81.
7. Riddle DL, Jiranek WA, Hayes CW. Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: a multicenter longitudinal cohort study. Arthritis Rheumatol 2014;66:2134–43.
8. Riddle DL, Perera RA, Jiranek WA, Dumenci L. Using surgical appropriateness criteria to examine outcomes of total knee arthroplasty in a United States sample. Arthritis Care Res 2015;67:349–57.
9. Riddle DL, Ghomrawi H, Jiranek WA, Dumenci L, Perera RA, Escobar A. Appropriateness criteria for total knee arthroplasty: additional comments and considerations. J Bone Jt Surg Am 2018;100:e22.
10. Katz JN, Winter AR, Hawker G. Measures of the appropriateness of elective orthopaedic joint and spine procedures. J Bone Jt Surg Am 2017;99:e15.
11. Price AJ, Alvand A, Troelsen A, Katz JN, Hooper G, Gray A, et al. Knee replacement. Lancet 2018;392:1672–82.
12. Riddle DL, Perera RA. Appropriateness and total knee arthroplasty: an examination of the American Academy of Orthopaedic Surgeons appropriateness rating system. Osteoarthritis Cartilage 2017;25:1994–8.
13. Inacio MC, Paxton EW, Graves SE, Namba RS, Nemes S. Projected increase in total knee arthroplasty in the United States – an alternative projection model. Osteoarthritis Cartilage 2017;25:1797–803.
14. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. BMJ Open 2012;2:e000435.
15. Escobar A, Quintana JM, Arostegui I, Azkarate J, Guenaga JI, Arenaza JC, et al. Development of explicit criteria for total knee replacement. Int J Technol Assess Health Care 2003;19:57–70.
16. Riddle DL, Perera RA. Appropriateness and total knee arthroplasty: an examination of the American Academy of Orthopaedic Surgeons appropriateness rating system. Osteoarthr Cartil 2017;25:1994–8. URL http://www.ncbi.nlm.nih.gov/pubmed/28888903.
17. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. Int J Technol Assess Health Care 1986;2:53–63.
18. Escobar A, Quintana J, Bilbao A, Azkárate J, Güenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster Universities Osteoarthritis Index. Clin Rheumatol 2002;21:466–71.
19. Herrero MJ, Blanch J, Peri JM, De Pablo J, Pintor L, Bulbena A. A validation study of the hospital anxiety and depression scale (HADS) in a Spanish population. Gen Hosp Psychiatry 2003;25:277–83.
20. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983;67:361–370.
21. Dave AJ, Selzer F, Losina E, Klara KM, Collins JE, Usiskin I, et al. Is there an association between whole-body pain with osteoarthritis-related knee pain, pain catastrophizing, and mental health? Clin Orthop Relat Res 2015;473:3894–902.
22. Scott W, Wideman TH, Sullivan MJL. Clinically meaningful scores on pain catastrophizing before and after multidisciplinary rehabilitation: a prospective study of individuals with subacute pain after whiplash injury. Clin J Pain 2014;30:183–90.
23. Miró J, Nieto R, Huguet A. The Catalan version of the pain catastrophizing scale: a useful instrument to assess catastrophic thinking in whiplash patients. J Pain 2008;9:397–406.
24. Ghomrawi HM, Alexiades M, Pavlov H, Nam D, Endo Y, Mandl LA, et al. An evaluation of two appropriateness criteria for total knee replacement. Arthritis Care Res (Hoboken) 2014;66:1749–53.
25. 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.
26. Vissers MM, Bussmann JB, Verhaar JA, Busschbach JJ, Bierma-Zeinstra SM, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. Semin Arthritis Rheum 2012;41:576–88.
27. Dumenci L, Perera RA, Keefe FJ, Ang DC, Slover J, Jensen MP, Riddle DL. Model-based pain and function outcome trajectory types for patients undergoing knee arthroplasty: a secondary analysis from a randomized clinical trial. Osteoarthritis Cartilage 2019;27:878–84.
28. Dowsey MM, Smith AJ, Choong PF. Latent class growth analysis predicts long term pain and function trajectories in total knee arthroplasty: a study of 689 patients. Osteoarthritis Cartilage 2015;23:2141–9.
29. Wylde V, Dieppe P, Hewlett S, Learmonth ID. Total knee replacement: is it really an effective procedure for all? Knee 2007;14:417–23.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 810–817
DOI 10.1002/acr.24177

Racial Differences in Pain and Function Following Knee Arthroplasty: A Secondary Analysis From a Multicenter Randomized Clinical Trial

Daniel L. Riddle,[1] James Slover,[2] Francis J. Keefe,[3] Dennis C. Ang,[4] Levent Dumenci,[5] and Robert A. Perera[1]

Objective. The assessment of racial differences in pain and function outcome following knee arthroplasty (KA) has received little attention despite very substantial literature exploring a variety of other prognostic factors. The present study was undertaken to determine whether race was associated with KA outcome after accounting for potential confounding factors.

Methods. We conducted a secondary analysis of a randomized clinical trial of 384 participants with moderate-to-high pain catastrophizing who underwent KA. Preoperative measures included race/ethnicity status as well as a variety of potential confounders, including socioeconomic status, comorbidity, and bodily pain. Outcome measures were Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain and function scales as well as performance measures. Linear mixed-effects models compared outcomes over a 1-year follow-up period for African American versus non–African American participants.

Results. WOMAC pain scores differences for African American versus non–African American participants averaged ~2 points in unadjusted analyses and 1–1.5 points in adjusted analyses. In adjusted analyses, follow-up WOMAC function scores differed by 6 points for African Americans compared to non–African Americans ($P = 0.002$).

Conclusion. African Americans generally had worse pain, function, and performance prior to KA and worse scores after surgery, but differences were small and attenuated by ~25–50% after adjustment for potential confounding. Only WOMAC function scores showed clinically important postsurgical differences in adjusted analyses. Clinicians should be aware that after adjustment for potential confounders, African Americans have approximately equivalent outcomes compared to others, with the exception of WOMAC function score.

INTRODUCTION

Racial differences in a variety of preoperative and postoperative characteristics have been found for patients undergoing knee arthroplasty (KA) (1,2). A recently published systematic review suggested that African Americans have worse preoperative and postoperative pain and function with KA as compared to White patients (3). While the investigators found that pain was generally worse preoperatively and postoperatively for African Americans as compared to White patients, adjustment for key covariates such as socioeconomic and psychological status was generally not considered. Despite a comprehensive literature search of multiple databases from the years January 2000 to April 2015, only 7 of 346 longitudinal cohort studies assessed for racial differences in outcome (3). Goodman et al (3) concluded that race and socioeconomic and psychological status are not generally taken into account in studies of KA outcomes and that this is an important gap in the literature on racial disparities.

We found 2 studies published after the systematic review by Goodman et al (3) that examined outcome differences between African Americans and non–African Americans. Lavernia and Villa, in a retrospective study of 1,905 White and 105 African American patients, found that group difference in pain and function outcomes was statistically different but clinically insignificant 1 year postsurgery

ClinicalTrials.gov identifier: NCT01620983.

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

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant UM1-AR-062800) and in part by the National Center for Advancing Translational Sciences (grant UL1-TR-000058).

[1]Daniel L. Riddle, PT, PhD, FAPTA, Robert A. Perera, PhD: Virginia Commonwealth University, Richmond; [2]James Slover, MD: New York University Medical Center, New York, New York; [3]Francis J. Keefe, PhD: Duke University, Durham, North Carolina; [4]Dennis C. Ang, MD: Wake Forest School of Medicine, Winston-Salem, North Carolina; [5]Levent Demenci, PhD: Temple University, Philadelphia, Pennsylvania.

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

Address correspondence to Daniel L. Riddle, PT, PhD, FAPTA, Departments of Physical Therapy, Orthopaedic Surgery, and Rheumatology, Virginia Commonwealth University, Richmond, VA 23298. Email: dlriddle@vcu.edu.

Submitted for publication September 9, 2019; accepted in revised form February 25, 2020.

SIGNIFICANCE & INNOVATIONS

- African Americans had worse pain, function, and physical performance prior to knee replacement and worse scores following surgery, but differences between African Americans and non–African Americans were small and, in most cases, not likely to be clinically relevant, particularly after adjustment for confounding.
- Adjustment for potential confounding attenuated score differences over time by ~25–50%.
- Only Western Ontario and McMaster Universities Osteoarthritis Index Function scores at 12 months postsurgery demonstrated both statistically and clinically significant differences between African American and non–African Americans.
- Clinicians should be aware that most differences in outcome following knee replacement among African American and non–African American patients were small and attenuated by confounding factors; race itself does not contribute to clinically meaningful differences for most pain and functional status outcomes.

(4). However, substantial loss to follow-up (e.g., 67 of 105 [63%] of African Americans) likely led to biased estimates of group differences. Goodman et al examined Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) outcomes 2 years after KA and found in a multivariable analysis that race, education, presurgical expectations, and baseline WOMAC pain, as well as census tract-based poverty level status, were associated with 2-year WOMAC pain scores (5). These data suggest that greater community-based poverty status is associated with worse outcome, but the investigators did not collect annual income or psychological variables, and 43% (i.e., 83 of 194) of African Americans in the sample were lost to follow-up. This newer evidence does not, in our view, answer the question of whether socioeconomic and psychological variables explain racial differences in KA outcomes.

African Americans tend to rely on social networks of family and friends in their community when opting for or against KA (6). African Americans are also known to be less likely to undergo KA when compared to either a similarly aged White population (7,8) or a White population at risk for KA with similar levels of symptomatic knee osteoarthritis (OA) (1). If adjustment for socioeconomic and psychological status attenuates or eliminates outcome differences between African Americans and non–African Americans, patients hesitant to undergo KA who are African American could be assured that outcomes are likely influenced by a variety of factors but that race is not likely to be one of them. Additionally, clinicians may be influenced by either implicit bias (i.e., an unconscious belief) (9) or explicit bias (i.e., an overt belief) (10) regarding race and KA outcomes or fitness for surgery. If evidence demonstrated that outcome differences are explained by psychological and socioeconomic differences between African Americans and non–African Americans and not by race, these data could encourage African Americans who qualify to undergo KA.

We recently completed a multicenter, 3-arm randomized clinical trial of a pain-coping skills training intervention as compared to usual care or arthritis education for individuals undergoing KA (11). For our no-effect trial, we obtained consent from 384 subjects who underwent KA, 135 (35.2%) of whom were African Americans. Because we collected a comprehensive spectrum of socioeconomic and psychological variables, the data were well suited for a comparison of outcomes for African American and non–African American patients.

The 2 objectives were to determine whether preoperative and postoperative pain and function, as well physical performance, differed for African Americans as compared to the remainder of the sample, and to determine if outcomes differed between African Americans and the remainder of the sample after adjustment for variables that potentially confound the relationship between race and outcome. We hypothesized that differences in pain and function over time between African Americans and non–African Americans would be attenuated after adjustment for psychological and sociodemographic variables, much like that seen for individuals with medically treated lower extremity OA (12,13).

PATIENTS AND METHODS

The study sample was taken from a 3-arm randomized clinical trial, funded by the National Institutes of Health, with a 1-year follow-up period (11). The trial examined effects of physical therapist–delivered pain coping skills relative to arthritis education or usual care for patients with at least moderate pain catastrophizing and scheduled for KA (the Knee Arthroplasty Skills Training Pain study [KASTPain]). Primary outcome measures for the trial were measured at baseline (preoperative visit, and 2 months, 6 months, and 12 months postoperatively). The protocol for the trial (14) and the final results have been published (11). The study was approved by the institutional review boards (IRBs) of the 5 participating sites (Virginia Commonwealth University [the central coordinating site (IRB HM14326)], Duke University, Wake Forest University, New York University, and Northern Illinois University). All subjects read and signed an IRB-approved consent form.

Study sample. To be eligible for our KASTPain trial, all participants were ≥45 years of age. They successfully completed cognitive screening (15), had a diagnosis of knee OA, and were scheduled for KA between 1 and 8 weeks following consent. Participants also scored ≥16 on the Pain Catastrophizing Scale (PCS) (16) and ≥5 on the WOMAC pain scale (17). The WOMAC pain threshold of ≥5 ensured our ability to measure improvement over time if it occurred. Exclusion criteria were revision or bilateral KA, inflammatory arthritis, established infection, fracture, and malignancy or plans to undergo additional hip or KA within 6 months of

the scheduled KA surgery. Because our trial design was pragmatic (18), we did not control clinical site surgical analgesic approaches, operative and pharmacologic or perioperative pain management, or preoperative or postoperative care delivery.

Key outcome variables. The WOMAC pain and function scales were collected in person at baseline and by telephone at all follow-up time points to quantify the extent of self-reported function-limiting pain and difficulty with activity at the baseline visit, 2 months, 6 months, and 12 months following KA. The WOMAC scales have clearly established psychometric properties for individuals undergoing KA (19). Results from the Six-Minute Walk Test (20) and the Short Physical Performance Battery (21) were obtained at baseline and the 12-month follow-up and were used to quantify the participants' physical performance. Both scales have strong validity evidence for individuals with knee OA (22,23).

Potential confounders of the relationship between race and outcome. We used prior evidence to select likely confounding variables (3,24–26). The previously validated Bodily Pain Questionnaire was used to identify which of 16 body regions had been painful for at least the prior 3 months (27). We also used the

Table 1. Baseline characteristics and medication use over the study period for African American participants and the remaining sample of participants who underwent knee arthroplasty surgery*

Baseline characteristics	African Americans (n = 135)	Remaining sample (n = 249)	*P*†	No. missing
Age, mean ± SD years	64.5 ± 8.1	60.7 ± 7.3	<0.001	–
Sex, female	94 (69.6)	163 (65.5)	0.407	–
Body mass index, mean ± SD kg/m^2	34.2 ± 6.4	31.2 ± 5.8	<0.001	7
Current income			<0.001	–
Less than $10,000	24 (17.8)	11 (4.4)		
$10,000–$24,999	39 (28.9)	39 (15.7)		
$25,000–$49,999	34 (25.2)	54 (21.7)		
$50,000–$99,999	21 (15.6)	72 (28.9)		
$100,000 or more	3 (2.2)	50 (20.1)		
Declined	14 (10.4)	23 (9.2)		
Education			<0.001	–
Less than high school	17 (12.6)	5 (2.0)		
High school graduate	37 (27.4)	49 (19.7)		
Some college	43 (31.9)	58 (23.3)		
College degree or higher	38 (28.1)	137 (55.0)		
Modified Charlson comorbidity score, mean ± SD‡	9.5 ± 4.3	8.2 ± 3.9	0.002	–
Knee pain duration, median years (IQR)	7 (3–15)	6 (3–14)	0.630	–
Patient Health Questionnaire score, mean ± SD§	6.3 ± 4.8	5.6 ± 4.9	0.230	–
Generalized Anxiety Scale score, mean ± SD¶	5.9 ± 5.0	5.1 ± 4.9	0.115	–
WOMAC pain score, mean ± SD#	12.4 ± 3.3	10.8 ± 3.3	<0.001	–
WOMAC physical function score, mean ± SD**	40.4 ± 11.3	35.3 ± 11.2	<0.001	–
Short Physical Performance Battery score, mean ± SD††	8.5 ± 3.2	9.7 ± 2.6	<0.001	15
Six-Minute Walk Test, mean ± SD meters	326.8 ± 97.3	372.4 ± 102.9	0.001	112
Pain Catastrophizing Scale score, mean ± SD‡‡	32.6 ± 9.1	28.5 ± 9.1	<0.001	–
Bodily pain sites, mean ± SD	6.5 ± 4.3	5.0 ± 3.8	0.001	–
Opioid medication use				
Baseline	62 (45.9)	58 (23.3)		–
2 months postoperative	68 (57.1)	79 (34.8)		38
6 months postoperative	33 (28.2)	19 (8.8)		50
12 months postoperative	21 (17.4)	14 (6.2)		38
Antiinflammatory medication use				
Baseline	73 (54.1)	153 (61.4)		–
2 months postoperative	20 (16.8)	65 (28.6)		38
6 months postoperative	20 (16.8)	44 (20.3)		50
12 months postoperative	21 (17.4)	28 (12.4)		38

* Values are the number (%) unless indicated otherwise. IQR = interquartile range; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† Independent sample *t*-tests were used for continuous variable comparisons, and chi-square tests were used for categorical variable comparisons.
‡ Modified Charlson comorbidity score range is 0–45. Higher scores equate to greater comorbidity burden.
§ Patient Health Questionnaire 8 score range is 0 to 24. Higher scores equate to more depressive symptoms.
¶ Generalized Anxiety Scale 7 score range is 0–21. Higher scores equate to more anxiety.
# WOMAC pain scale score range is 0–20. Higher scores equate to more function-limiting pain.
** WOMAC function scale range is 0–68. Higher scores equate to more difficulty with functional activities.
†† Short Physical Performance Battery score range is 0–12. Higher scores equate to better performance.
‡‡ Pain Catastrophizing Scale range is 0–52. Higher scores equate to more pain catastrophizing.

validated Patient Health Questionnaire 8 (PHQ-8) (28) to quantify depressive symptoms. Comorbidity was assessed using a validated comorbidity questionnaire (29). PCS is a validated measure that quantifies the extent to which patients catastrophize about their pain (30,31). For all scales, higher scores equated to a worse status for that scale.

We also measured yearly income and education as indicators of socioeconomic status (SES). Yearly income was measured using a 5-item ordinal scale (i.e., <$10,000, $10,000–$24,999, $25,000–$49,999, $50,000–$99,999, ≥$100,000), while education was measured using a 4-item ordinal scale (i.e., more than high school, high school graduate, some college, college degree or higher). Based on recommendations from Goodman et al (3) and substantial disparities in both income and education between African Americans and non–African American populations considering KA (32), we included income and education as indicators of SES in our analyses. Medication use during the postoperative period was assessed at the 2-month, 6-month, and 12-month time points. Participants were provided a comprehensive list of opioid and antiinflammatory medications and were systematically asked whether they were taking each of the medications on the list. For the purposes of this study, we only utilized preoperative opioid medication data dichotomized to indicate whether the participant was taking an opioid (yes or no) preoperatively (33). The demographic variables of sex, body mass index (BMI) (kg/m^2), and age also were included as potential confounders.

Data on race and ethnicity were collected in the trial (i.e., American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian or other Pacific Islander, White, or declined). For purposes of the current study, we dichotomized race/ethnicity to either African American/Black or other.

Data analysis. Linear mixed-effects models were used to determine the effect of potential confounders on differences in pain and function at baseline and at 2, 6, and 12 months postsurgery based on race. This allowed us to account for the within-subjects nature of the data and the nesting for participant within surgeon within site. Parallel models were fit for each outcome. Unadjusted models included a race main effect, time main effect, and race-by-time interaction. Treatment group from the main trial was not included in the analyses as there is no evidence for differences between treatment groups, as previously reported (11). The models were compared to adjusted models that included each proposed confounder (i.e., preoperative painful body regions, PHQ-8 score, comorbidity, PCS score, SES, annual income, age, sex, BMI, as well as treatment with opioids). F tests of the main effect of race and interaction between race and time are reported, as well as the estimated means and their corresponding 95% confidence intervals at each time point for each racial group. Missing data were handled directly during estimation using maximum likelihood estimation, which assumes missing at random (MAR) data and is appropriate when data are either MAR or missing completely at random (34). All analyses were completed using the lme4 package in the R statistical software (35).

RESULTS

Sample characteristics prior to surgery for African Americans and all other participants in the study are displayed in Table 1, and a Consolidated Standards of Reporting Trials (CONSORT) flow chart illustrating the patient flow and loss to follow-up appears in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24177/abstract. For the sample of 384 participants, 135 self-reported as African American or Black, and 240 self-reported as White. The remainder (n = 9) were distributed among the other race/ethnicity categories. African Americans were older, had a higher BMI, and a lower educational level and income as compared to the non–African Americans.

The main effects of race and race-by-time interactions are shown in Table 2. Generally, there were statistically significant differences between African Americans and non–African Americans for all outcomes in the unadjusted analyses of main effects, but in the adjusted analyses, only WOMAC pain and function continued to be significant for race in the main effects analysis. When considering race-by-time interactions, the physical performance comparisons were not significant, and when considering the race-by-time interaction for the adjusted analyses, only WOMAC function scores were significantly different. For example, the race-by-time interaction for WOMAC function was $F_{3,920.5} = 4.8$, $P = 0.002$. Follow-up comparisons were completed, testing for differences in WOMAC function at each time point using a

Table 2. Results of main effects and race-by-time interactions for the race and outcome comparisons*

Outcome	Race main effect				Race-by-time interaction			
	Unadjusted	*P*	Adjusted	*P*	Unadjusted	*P*	Adjusted	*P*
WOMAC pain score†	$F_{1,389.1} = 50.0$	<0.001	$F_{1,357.9} = 20.4$	<0.001	$F_{3,1046.8} = 2.7$	0.043	$F_{3,1022.7} = 2.3$	0.071
WOMAC function score‡	$F_{1,380.1} = 62.8$	<0.001	$F_{1,344.4} = 28.2$	<0.001	$F_{3,934.4} = 5.5$	0.001	$F_{3,923.0} = 4.9$	0.002
Six-Minute Walk Test score	$F_{1,362.1} = 12.2$	<0.001	$F_{1,329.0} = 2.8$	0.095	$F_{1,310.5} = 0.3$	0.603	$F_{1,303.9} = 0.1$	0.802
SPPB score§	$F_{1,392.6} = 13.7$	<0.001	$F_{1,349.9} = 3.1$	0.081	$F_{1,215.7} = 0.2$	0.646	$F_{1,214.1} = 0.1$	0.775

* SPPB = Short Physical Performance Battery; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† WOMAC pain scale range is 0–20. Higher scores equate to more function-limiting pain.
‡ WOMAC function scale range is 0–68. Higher scores equate to more difficulty with functional activities.
§ SPPB score range is 0–12. Higher scores equate to better performance.

Bonferroni-corrected α of 0.05/4 = 0.0125. Statistically significant differences were found at 2 (P = 0.004), 6 (0.004), and 12 months (P = 0.001), but not preoperatively (P = 0.152), indicating insufficient evidence for a preoperative difference in groups but diverging function scores following surgery. These data indicate that the improvement for African Americans was less over the early postoperative period compared to the improvement seen in the comparator group. Table 3 summarizes the mean differences between African Americans and non–African Americans for both the unadjusted analyses and the analyses adjusted for potential confounders. For example, unadjusted point estimates for baseline WOMAC pain scores were 12.3 for African Americans and 11.0 for non–African Americans, while adjusted estimates were 11.5 for African Americans and 11.0 for non–African Americans. Figures 1 and 2 illustrate both the unadjusted and adjusted WOMAC pain (Figure 1) and WOMAC function (Figure 2) scores over time for the 2 groups.

DISCUSSION

Research examining pain and function outcomes following KA for African Americans versus non–African Americans has traditionally lacked adjustment for key potential socioeconomic and psychological confounders (3–5). Evidence that accounts for potential confounders is critical in determining whether race-based differences are attributable specifically to race or to other variables that confound the race–outcome relationship. We found that prior to surgery, African Americans reported significantly more pain with activity as well as lower levels of function and physical performance, and this finding is consistent with other research (3). One novel contribution of our study is that preoperative differences were attenuated by ~75% for pain, function, and performance measures in the adjusted analyses. All adjusted preoperative differences for the pain, function, and performance measures fall below minimum clinically important difference (MCID) estimates for each measure (e.g., ~2.0 points for WOMAC pain and 7 points for WOMAC function scores) (36–39). These data suggest that prior assertions (3) that African Americans have more pain and functional loss at baseline as compared to others is likely to be true. However, when confounding factors are accounted for in the analytic approach, these baseline differences become substantially smaller (see Figures 1 and 2). Preoperative differences attributed only to race are not likely to be clinically meaningful.

After surgery, adjusted analyses of 2-month, 6-month, and 12-month postoperative outcomes comparing African Americans to all others in the sample were attenuated on the order of 25% to 50% relative to unadjusted estimates. The adjusted analyses indicated, for example, that for postoperative WOMAC pain scores, mean differences were ~1.5 WOMAC pain points, less than MCID estimates of 2.0 (36,37). In adjusted analyses, only WOMAC function showed a statistically significant interaction. These later findings indicated that only the trajectories of adjusted WOMAC function score differences between the 2 race groups became greater over time. The performance measures were nonsignificant for both the adjusted and unadjusted analyses.

Taken as a whole, our findings indicate that once confounders are accounted for, differences in presurgical and postsurgical

Table 3. Mean estimates and 95% confidence intervals (95% CIs) from linear mixed-effects models for each outcome over time for African Americans as compared to all other races and ethnicities*

	Unadjusted analyses		Adjusted analyses	
	African American	Other	African American	Other
WOMAC pain score†				
Baseline	12.3 (11.6–12.9)	11.0 (10.4–11.5)	11.5 (10.8–12.2)	11.0 (10.3–11.6)
2 months	7.8 (7.1–8.6)	5.4 (4.9–6.0)	7.4 (6.7–8.2)	5.8 (5.2–6.4)
6 months	5.5 (4.8–6.2)	3.4 (2.8–4.0)	4.9 (4.2–5.6)	3.6 (3.0–4.2)
12 months	4.5 (3.8–5.2)	2.4 (1.8–2.9)	4.2 (3.4–4.9)	2.8 (2.1–3.4)
WOMAC function score‡				
Baseline	40.3 (38.1–42.4)	35.9 (34.2–37.6)	37.8 (35.8–39.9)	36.0 (34.3–37.7)
2 months	27.1 (24.7–29.4)	18.1 (16.3–20.0)	25.0 (22.8–27.2)	19.0 (17.1–20.8)
6 months	20.9 (18.6–23.3)	12.2 (10.4–14.1)	18.5 (16.3–20.7)	12.5 (10.6–14.3)
12 months	17.6 (15.3–19.9)	8.4 (6.6–10.2)	15.6 (13.5–17.8)	9.3 (7.4–11.1)
Six-Minute Walk Test score				
Baseline	269.6 (247.4–291.8)	304.6 (288.1–321.2)	292.4 (269.3–315.4)	307.9 (288.3–327.4)
12 months	327.0 (303.2–350.9)	369.9 (352.0–387.7)	350.4 (326.1–374.6)	371.6 (351.2–392.0)
SPPB§				
Baseline	7.2 (6.7–7.6)	8.0 (7.6–8.4)	7.4 (7.0–7.8)	7.8 (7.4–8.1)
12 months	7.7 (7.2–8.3)	8.6 (8.2–9.0)	8.0 (7.5–8.5)	8.4 (8.0–8.8)

* Values are the mean (95% CI). Analyses were adjusted for bodily pain, depressive symptoms, comorbidity, pain catastrophizing, income, education, sex, body mass index, opioid and antiinflammatory medication, and age. SPPB = Short Physical Performance Battery; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† WOMAC pain scale score range is 0–20. Higher scores equate to more function-limiting pain.
‡ WOMAC function scale range is 0–68. Higher scores equate to more difficulty with functional activities.
§ SPPB score range is 0–12. Higher scores equate to better performance.

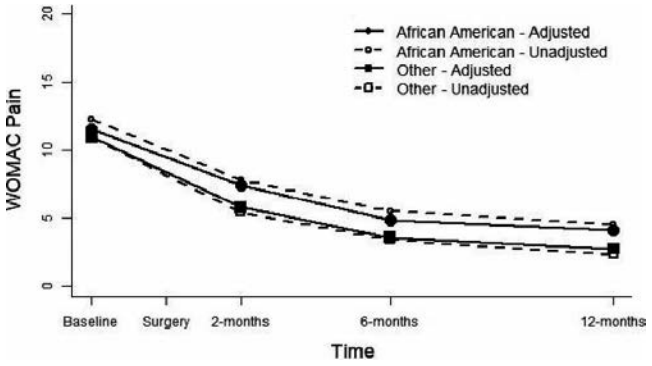

Figure 1. Unadjusted Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain trajectories and trajectories adjusted for potential confounders are illustrated from prior to surgery to 1 year following knee arthroplasty for African American participants and all others.

outcomes for African Americans as compared to all others are small and unlikely to be clinically relevant with an exception. Differences in adjusted WOMAC function scores from 2 months to 12 months postsurgery were on the order of 6–7 points higher (worse) for African Americans compared to all others. This difference is statistically significant and clinically important as well (36,37).

African American patients are known to underutilize KA at a substantially higher rate as compared to White patients (1,40). Underutilization among African Americans is driven primarily by values and preferences for nonsurgical care, an expectation that KA is not effective, and a lack of trust in the health care system (6,41). Additionally, clinicians may be contributing to underutilization because of implicit or explicit bias toward African Americans (9,10,42).

Underutilization of KA by African Americans is clearly multifactorial, and the current findings address some of these factors. For example, our finding that preoperative and postoperative differences among African Americans are generally below MCID estimates (with the exception of postoperative WOMAC function) after accounting for non-race confounders indicates that race itself likely does not contribute to a difference in clinically important outcomes for pain and physical performance. Educationally based decision aids developed specifically for African Americans have been shown to lead to increased utilization of KA (43). Our findings suggest that decision aids also should include information indicating that preoperative and postoperative pain and performance outcomes are not likely to be clinically different for African Americans as compared to non–African Americans. This information may help to change expectations of inferior outcomes following KA.

Because preoperative scores are the most powerful predictors of postsurgical outcomes (44), we would expect that, on average, African Americans would have slightly worse outcomes compared to all others. Only differences in WOMAC function scores between African Americans and non–African Americans meet the MCID (36,45). Importantly, we found that adjusted differences between African Americans and non–African Americans are most pronounced during the period from surgery to 2 months postsurgery. Our analyses suggested that this period may be the critical time to focus on exercise and activity enhancement for African Americans to optimize function during this critical period.

Our findings are consistent with research conducted on individuals receiving nonsurgical treatment for knee OA. For example, in a meta-analysis designed to detect baseline pain and function differences in African Americans compared to White patients, Vaughn and colleagues found worse physical performance, self-reported pain, and function scores for African Americans (46). Flowers et al found that African Americans had worse physical performance test scores than White patients, but similar to our study, these differences were attenuated, and in some cases, nonsignificant after adjustment for demographic, psychological, and SES variables (47). Our findings appear to extend the findings of the 2 studies mentioned above regarding KA preoperative and follow-up measures obtained on African Americans as compared to the non–African Americans.

Lavernia and Villa compared preoperative and postoperative scores on WOMAC and visual analog scale pain measures among a registry-based sample of 2010 patients with total KA, 5% of whom were African American. Baseline and follow-up differences between African Americans and non–African Americans were small and not clinically meaningful. However, substantial loss to follow-up (i.e., 63% of African Americans and 36% of non–African Americans) raises concerns about the generalizability of the findings. Similarly, Goodman et al used registry-based data on total KA and reported more substantial differences in baseline and 2-year WOMAC outcomes but noted a 43% loss to follow-up for African Americans. Loss to follow-up for African Americans in our study ranged from 10% to 13% for the 2-month to 12-month follow-up visits.

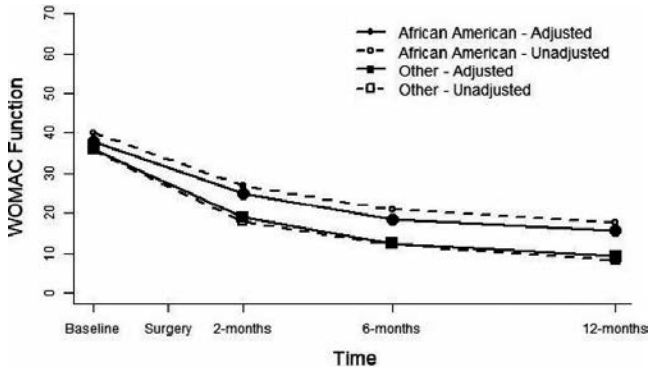

Figure 2. Unadjusted Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) function trajectories and trajectories adjusted for potential confounders are illustrated from prior to surgery to 1 year following knee arthroplasty for African American participants and all others.

Our study had several strengths, including the relatively large proportion of participants who were African American (i.e., 35%) and the comprehensive assessment of potential confounding variables. There are also some important limitations. Because we conducted a secondary analysis, our study was not powered a priori for comparing outcomes among African Americans and all others in the sample. With this said, we had 135 African American participants with 90% follow-up along with 249 additional participants, also with 90% follow-up. Our analyses indicated that we were powered to detect unadjusted differences between African Americans and non–African Americans of 1 to 2 WOMAC pain points. Our measures of SES only included annual income and educational level. A more sophisticated measure of SES that included occupation, neighborhood, and other assets (48) may have more accurately reflected the construct of interest and may have better explained race–outcome relationships. Additionally, other variables that explain the small differences in outcome scores after confounder adjustment in our study could not be determined and require further research.

All of our participants had at least moderate levels of pain catastrophizing, and while our 1-year pain and function outcomes were very similar to outcomes reported for patients who were not recruited based on pain catastrophizing scores (44,49), it is possible that our finding may not generalize to heterogeneous samples of patients. We did not measure actual physical activity before and following KA, and this outcome may have been useful for determining whether differences existed among African Americans and non–African Americans. African Americans were well represented in our sample, but other races and ethnicities were poorly represented, with only 9 additional subjects, and this lack of representation reduces generalizability for these populations. Finally, our study did not account for preoperative knee OA severity, a known predictor of poor outcome (50). It is possible that African Americans had milder knee OA than the non–African Americans, although we found no data to suggest that this is the case.

In conclusion, we found that prior to surgery, African Americans have slightly worse self-reported preoperative pain, function, and physical performance as compared to other participants, but these differences are attenuated by ~50% after adjustment for confounding variables and are likely not clinically relevant. Follow-up measures showed a similar pattern, with 1 exception. Small differences were attenuated after adjustment and likely became clinically insignificant. The WOMAC function scale was the 1 exception that demonstrated differences among African Americans versus non–African Americans that were likely to be clinically relevant.

ACKNOWLEDGMENTS

The authors thank the study participants for their willingness to enroll and continue participation over the 1-year period of study. We also thank the physical therapists and nurses for their role in providing the interventions under study, as well as study staff and study psychologists, at Duke University, New York University, Southern Illinois University, Virginia Commonwealth University, and Wake Forest University.

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. Riddle 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. Riddle.
Acquisition of data. Riddle, Slover, Keefe, Ang.
Analysis and interpretation of data. Riddle, Dumenci, Perera.

REFERENCES

1. Macfarlane LA, Kim E, Cook NR, Lee IM, Iversen MD, Katz JN, et al. Racial variation in total knee replacement in a diverse nationwide clinical trial. J Clin Rheumatol 2018;24:1–5.
2. Zhang W, Lyman S, Boutin-Foster C, Parks ML, Pan TJ, Lan A, et al. Racial and ethnic disparities in utilization rate, hospital volume, and perioperative outcomes after total knee arthroplasty. J Bone Joint Surg Am 2016;19:1243–52.
3. Goodman SM, Parks ML, McHugh K, Fields K, Smethurst R, Figgie MP, et al. Disparities in outcomes for African Americans and Whites undergoing total knee arthroplasty: a systematic literature review. J Rheumatol 2016;43:765–770.
4. Lavernia CJ, Villa JM. Does race affect outcomes in total joint arthroplasty? Clin Orthop Relat Res 2015;473:3535–41.
5. Goodman SM, Mandl LA, Parks ML, Zhang M, McHugh KR, Lee YY, et al. Disparities in TKA outcomes: census tract data show interactions between race and poverty. Clin Orthop Relat Res 2016;474: 1986–95.
6. Parks ML, Hebert-Beirne J, Rojas M, Tuzzio L, Nelson CL, Boutin-Foster C. A qualitative study of factors underlying decision making for joint replacement among African Americans and Latinos with osteoarthritis. J Long Term Eff Med Implants 2014;24:205–12.
7. Singh JA, Lu X, Rosenthal GE, Ibrahim S, Cram P. Racial disparities in knee and hip total joint arthroplasty: an 18-year analysis of national Medicare data. Ann Rheum Dis 2014;73:2107–15.
8. Skinner J, Weinstein JN, Sporer SM, Wennberg JE. Racial, ethnic, and geographic disparities in rates of knee arthroplasty among Medicare patients. N Engl J Med 2003;349:1350–9.
9. Fitzgerald C, Hurst S. Implicit bias in health care professionals: a systematic review. BMC Med Ethics 2017;18:19.
10. Salles A, Awad M, Goldin L, Krus K, Lee JV, Schwabe MT, et al. Estimating implicit and explicit gender bias among health care professionals and surgeons. JAMA Netw Open 2019;2:e196545.
11. Riddle DL, Keefe FJ, Ang DC, Slover JD, Jensen MP, Bair MJ, et al. Pain coping skills training for patients who catastrophize about their pain prior to knee arthroplasty: a multisite randomized clinical trial. J Bone Joint Surg Am 2018;101:218–27.
12. Allen KD, Helmick CG, Schwartz TA, DeVellis RF, Renner JB, Jordan JM. Racial differences in self-reported pain and function among individuals with radiographic hip and knee osteoarthritis: the Johnston County Osteoarthritis Project. Osteoarthritis Cartilage 2009;17:1132–6.
13. Allen KD, Oddone EZ, Coffman CJ, Keefe FJ, Lindquist JH, Bosworth HB. Racial differences in osteoarthritis pain and function: potential explanatory factors. Osteoarthritis Cartilage 2010; 18:160–7.
14. Riddle DL, Keefe FJ, Ang D, Saleh KJ, Dumenci L, Jensen MP, et al. A phase III randomized three-arm trial of physical therapist delivered pain coping skills training for patients with total knee arthroplasty: the KASTPain protocol. BMC Musculoskelet Disord 2012;13:149.
15. Callahan CM, Unverzagt FW, Hui SL, Perkins AJ, Hendrie HC. Six-item screener to identify cognitive impairment among potential subjects for clinical research. Med Care 2002;40:771–81.

16. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. Psychol Assess 1995;7:524–32.
17. Bellamy N. The WOMAC knee and hip osteoarthritis indices: development, validation, globalization and influence on the development of the AUSCAN Hand Osteoarthritis Indices. Clin Exp Rheumatol 2005;23 Suppl S:148S–53S.
18. Riddle DL, Johnson RE, Jensen MP, Keefe FJ, Kroenke K, Bair MJ, et al. The pragmatic-explanatory continuum indicator summary (PRECIS) instrument was useful for refining a randomized trial design: experiences from an investigative team. J Clin Epidemiol 2010;63:1271–5.
19. McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties [review]. Arthritis Rheum 2001;45:453–61.
20. Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. BMC Musculoskelet Disord 2005;6:3.
21. Freire AN, Guerra RO, Alvarado B, Guralnik JM, Zunzunegui MV. Validity and reliability of the short physical performance battery in two diverse older adult populations in Quebec and Brazil. J Aging Health 2012;24:863–78.
22. Cecchi F, Molino-Lova R, Di IA, Conti AA, Mannoni A, Lauretani F, et al. Measures of physical performance capture the excess disability associated with hip pain or knee pain in older persons. J Gerontol A Biol Sci Med Sci 2009;64:1316–24.
23. Ko V, Naylor JM, Harris IA, Crosbie J, Yeo AE. The six-minute walk test is an excellent predictor of functional ambulation after total knee arthroplasty. BMC Musculoskelet Disord 2013;14:145.
24. Dave AJ, Selzer F, Losina E, Usiskin I, Collins JE, Lee YC, et al. The association of pre-operative body pain diagram scores with pain outcomes following total knee arthroplasty. Osteoarthritis Cartilage 2016;25:667–75.
25. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. Clin Orthop Relat Res 2010;468:798–806.
26. Vissers MM, Bussmann JB, Verhaar JA, Busschbach JJ, Bierma-Zeinstra SM, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. Semin Arthritis Rheum 2012;41:576–88.
27. Arnold LM, Stanford SB, Welge JA, Crofford LJ. Development and testing of the fibromyalgia diagnostic screen for primary care. J Womens Health (Larchmt) 2012;21:231–9.
28. 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.
29. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 2003;49:156–63.
30. Sullivan M, Tanzer M, Stanish W, Fallaha M, Keefe FJ, Simmonds M, et al. Psychological determinants of problematic outcomes following total knee arthroplasty. Pain 2009;143:123–9.
31. Dumenci L, Perera RA, Keefe FJ, Ang DC, Slover J, Jensen MP, et al. Model-based pain and function outcome trajectory types for patients undergoing knee arthroplasty: a secondary analysis from a randomized clinical trial. Osteoarthritis Cartilage 2019;27:878–84.
32. Dunlop DD, Song J, Manheim LM, Chang RW. Racial disparities in joint replacement use among older adults. Med Care 2003;41:288–98.
33. 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.
34. Little RJ, Rubin DB. Statistical analysis with missing data. Vol. 793. Hoboken (NJ): Wiley Publishers; 2019.
35. Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015;67:1–48.
36. 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.
37. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis 2005;64:29–33.
38. Dobson F, Hinman RS, Hall M, Marshall CJ, Sayer T, Anderson C, et al. Reliability and measurement error of the Osteoarthritis Research Society International (OARSI) recommended performance-based tests of physical function in people with hip and knee osteoarthritis. Osteoarthritis Cartilage 2017;25:1792–6.
39. Latham NK, Mehta V, Nguyen AM, Jette AM, Olarsch S, Papanicolaou D, et al. Performance-based or self-report measures of physical function: which should be used in clinical trials of hip fracture patients? Arch Phys Med Rehabil 2008;89:2146–55.
40. Singh JA, Lu X, Rosenthal GE, Ibrahim S, Cram P. Racial disparities in knee and hip total joint arthroplasty: an 18-year analysis of national Medicare data. Ann Rheum Dis 2014;73:2107–15.
41. Shahid H, Singh JA. Racial/ethnic disparity in rates and outcomes of total joint arthroplasty. Curr Rheumatol Rep 2016;18:20.
42. Haider AH, Schneider EB, Sriram N, Dossick DS, Scott VK, Swoboda SM, et al. Unconscious race and social class bias among acute care surgical clinicians and clinical treatment decisions. JAMA Surg 2015;150:457–64.
43. Ibrahim SA, Blum M, Lee GC, Mooar P, Medvedeva E, Collier A, et al. Effect of a decision aid on access to total knee replacement for black patients with osteoarthritis of the knee a randomized clinical trial. JAMA Surg 2017;152:e164225.
44. Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. J Bone Joint Surg Am 2004;86:2179–86.
45. Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. J Rheumatol 2002;29:131–8.
46. Vaughn IA, Terry EL, Bartley EJ, Schaefer N, Fillingim RB. Racial-ethnic differences in osteoarthritis pain and disability: a meta-analysis. J Pain 2019;20:629–44.
47. Flowers PP, Schwartz TA, Arbeeva L, Golightly YM, Pathak A, Cooke J, et al. Racial differences in performance-based function and potential explanatory factors among individuals with knee osteoarthritis. Arthritis Care Res (Hoboken) 2020;72:1196–204.
48. Shavers VL. Measurement of socioeconomic status in health disparities research. J Natl Med Assoc 2007;99:1013–23.
49. Collins JE, Donnell-Fink LA, Yang HY, Usiskin IM, Lape EC, Wright J, et al. Effect of obesity on pain and functional recovery following total knee arthroplasty. J Bone Joint Surg Am 2017;99:1812–8.
50. Youlden DJ, Dannaway J, Enke O. Radiographic severity of knee osteoarthritis and its relationship to outcome post total knee arthroplasty: a systematic review. ANZ J Surg 2020;90:237–42.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 818–827
DOI 10.1002/acr.24486

American College of Rheumatology
Empowering Rheumatology Professionals

Willingness to Undergo Joint Surgery Following a First-Line Intervention for Osteoarthritis: Data From the Better Management of People With Osteoarthritis Register

Andrea Dell'Isola,[1] Thèrèse Jönsson,[1] Ola Rolfson,[2] Anna Cronström,[3] Martin Englund,[1] and Leif Dahlberg[1]

Objective. To assess the proportion of participants reconsidering their willingness to undergo surgery after 3 and 12 months. Secondary aims were to analyze and compare the characteristics of individuals willing and unwilling to undergo joint surgery for osteoarthritis (OA) before a first-line intervention, and to study the association between pain intensity, walking difficulties, self-efficacy, and fear of movement with the willingness to undergo surgery.

Methods. This was an observational study based on Swedish register data. We included 30,578 individuals with knee or hip OA who participated in a first-line intervention including education and exercise.

Results. Individuals willing to undergo surgery at baseline showed a higher proportion of men (40% versus 27%) and more severe symptoms and disability. Respectively, 45% and 30% of the individuals with knee and hip OA who were willing to undergo surgery at baseline became unwilling after the intervention. At the end of the study period (12 months), 35% and 19% of those with knee and hip OA, respectively, who were willing to undergo surgery at baseline became unwilling. High pain intensity, walking difficulties, and fear of movement were associated with higher odds of being willing to undergo surgery at both follow-ups, while increased self-efficacy showed the opposite association.

Conclusion. A first-line intervention for OA is associated with reduced willingness to undergo surgery, with a greater proportion among patients with knee OA than hip OA. Due to its temporal variability, willingness to undergo surgery should be used with care to deem surgery eligibility.

INTRODUCTION

In individuals with long-standing and severe knee or hip osteoarthritis (OA), total joint replacement (TJR) is an effective intervention to reduce pain and disability (1). In the last decades, the use of joint replacement for OA has dramatically increased and its growth is expected to continue, partially driven by the rising prevalence of OA (2–4).

Despite the fact that TJR is a common procedure, there appears to be little consensus regarding the indication for TJR (5,6). Decision-making is complex and based on the interaction of multiple factors, such as patient willingness to undergo surgery and disease severity, but also based on social factors and previous experiences as well as availability (6–12). Patients' willingness to undergo surgery is the strongest predictor for TJR and has been hypothesized to be in part responsible for the high number of TJR procedures deemed as inappropriate and also responsible for the residual pain and disability observable in 1 of 5 patients with a TJR for OA (13,14).

Exercise in combination with education (and weight loss if indicated) is the first-line intervention for hip and knee OA, and both national and international guidelines recommend it. Randomized controlled trials (RCTs) have shown that first-line interventions can postpone surgery for up to 2 years in patients on a waiting list for TJR (15,16). Similarly, observational studies have shown that first-line interventions can shift patients' willingness to undergo surgery in the short term, raising further questions on the use of preferences for TJR in the surgical decision process (17,18). However, very little is known about how often patients with OA reconsider their willingness to undergo surgery after a first-line intervention

[1]Andrea Dell'Isola, PT, PhD, Thèrèse Jönsson, PT, Martin Englund, MD, PhD, Leif Dahlberg, MD, PhD: Lund University, Lund, Sweden; [2]Ola Rolfson, MD, PhD: The Swedish Hip Arthroplasty Register, Centre of Registers Västra Götaland, and Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; [3]Anna Cronström, PT, PhD: Lund University, Lund, and Umeå University, Umeå, Sweden.

Leif Dahlberg is the cofounder and Chief Medical Officer of Joint Academy, a company that provides web-based nonsurgical interventions for patients with hip and knee osteoarthritis. No other disclosures relevant to this article were reported.

Address correspondence to Andrea Dell'Isola, PT, PhD, Lund University, Faculty of Medicine, Department of Clinical Sciences, Orthopaedics, Clinical Epidemiology Unit, Lund, Sweden. Email: andrea.dellisola@med.lu.se.

Submitted for publication February 28, 2020; accepted in revised form October 8, 2020.

SIGNIFICANCE & INNOVATIONS

- Results from this large cohort show a reduction of 30–45% in willingness to undergo surgery after completion of a first-line intervention and indicate that this reduction is partially maintained up to 12 months. Patients with osteoarthritis undergoing a first-line intervention tend to reconsider their willingness to undergo surgery multiple times, suggesting that the willingness to undergo surgery may not be an optimal indicator to deem eligibility for total joint replacement, especially if the person seeking surgery has not yet undertaken a first-line intervention.
- Pain was significantly associated with the willingness to undergo surgery at both follow-ups, suggesting that a reduction of 1 unit in the pain intensity (measured on a 0–10 numeric rating scale) can lead to 60–80% lower odds of being willing to undergo surgery.
- Walking difficulties appear to be central in determining a person's willingness to undergo surgery and may, in certain cases, be more important than pain, especially when pain is measured on a quantitative scale.

delivered in a clinical context, and whether the change in willingness is maintained in the long term. In fact, RCTs provide useful information to establish causality but are often limited by stringent selection criteria and cannot account for the large variability that characterizes clinical settings. On the other hand, existing observational studies often have small sample sizes and short follow-ups, somewhat limiting the generalizability of results.

To better understand patients' preferences and to improve the decision process leading to surgery, it is fundamental to understand how willingness to undergo surgery may shift at different time points after a first-line intervention for OA and to understand which factors are associated with the shift in willingness to undergo surgery. Therefore, the main aim of this study was to assess the proportion of participants reconsidering their willingness to undergo surgery at 3 and 12 months after taking part in a first-line intervention. Our secondary aims were to compare the characteristics of individuals willing and unwilling to undergo surgery before taking part in a first-line self-management intervention provided nationwide in Sweden and to study the association of symptoms, quality of life, and psychological factors with the willingness to undergo surgery after 3 and 12 months.

MATERIALS AND METHODS

Study design. This was an observational register-based study, and it adhered to the Strengthening the Reporting of Observational Studies in Epidemiology guidelines for observational studies (19). The study was approved by the Regional Ethical Review Board in Gothenburg (1059-16).

Intervention. The Better Management of People with Osteoarthritis (BOA) is a national quality register collecting data of a first-line management program for individuals with hand, hip, and knee OA that started in Sweden in 2008 and is provided in primary care settings. All individuals taking part in the program receive a minimum of 2 theoretical group sessions led by a physical therapist focusing on the disease pathophysiology, on the effectiveness and indication of OA treatments (including surgery, pharmacologic, and nonpharmacologic treatments), and on the benefit of exercise, including self-management advice and strategy to incorporate exercise into daily life (20).

Between 1 and 3 weeks after the education, all the participants are offered the possibility to take part in the exercise phase of the intervention, which consists of a face-to-face session with a physical therapist. In this session, the patients receive a personalized intervention program and the necessary instructions to perform it independently at home (21,22). Thereafter, participants are given the possibility to perform their exercise program on their own or to participate in up to 12 supervised group exercise sessions with a physical therapist, provided 2 times a week for 6 weeks. Further details on the program delivered to the participants recorded in BOA can be found elsewhere (20).

Study sample. The study sample consisted of patients with knee and/or hip OA with data recorded in BOA between September 2008 and December 2016. These patients sought treatment in primary health care in Sweden for knee and/or hip pain and were referred for a standardized core treatment (education and exercises) after a confirmed clinical or radiographic OA diagnosis as recommended by the Swedish National Board of Health and Welfare (23). These guidelines are in line with internationally accepted diagnostic criteria for OA, suggesting that a radiographic examination should only be used in uncertain cases, if the patient is not responding to treatment or when surgical intervention is planned (24,25). Patients were excluded from the program if 1) the joint pain was caused by other diseases or conditions (e.g., sequel hip fractures, chronic widespread pain, inflammatory joint diseases, or cancer), 2) they had received a TJR in 1 of their knees or hips within the previous 12 months, 3) they had other surgery of the knee or hip within the past 3 months, or 4) they were not able to read or understand Swedish. The index joint for the treatment was identified by a physical therapist and based on the patient's medical history and complaints. If >1 joint was affected, the most symptomatic joint was used as the index joint.

For this study, we included all participants who completed the 2 mandatory education sessions, were willing to undergo surgery, as recorded at baseline, and attended both follow-ups at 3 months (between 60 and 150 days from enrollment) and 12 months (between 360 and 450 days from enrollment), or did not attend 1 of the follow-ups because they received a TJR in the index joint. Participants who underwent TJR before the 3-month follow-up were considered as willing to undergo surgery

at the 3-month follow-up. Similarly, patients who underwent TJR after the 3-month follow-up were considered as being willing to undergo surgery at 12 months. For pragmatic reasons, we allowed a 3-month window at the follow-ups to ensure that all the participants were able to attend the follow-ups.

Variables. Willingness to undergo surgery was assessed by the question "Are your knee/hip symptoms so severe that you wish to undergo surgery? (Yes/No)," asked at baseline and both follow-ups. Mean pain intensity during the last week in the most affected joint was evaluated at baseline and follow-ups on a numeric rating scale ranging from 0 (no pain) to 10 (maximum pain) (21). The presence of perceived walking difficulties was assessed by the question "Do you have problems walking as results of your joint problems (Yes/No)" at baseline and both follow-ups. Participants reported their age, sex, and level of education. Body weight and height were self-reported at the first visit, from which the body mass index (BMI) was calculated as kg/m^2.

Participants rated their general health status using the 5-level EuroQol 5-domain (EQ-5D-5L) instrument. For this study, we used the EQ-5D-5L visual analog scale, with a score ranging from 0 (worst imaginable health state) to 100 (best imaginable health state) as a measure of overall health-related quality of life (26,27).

Self-efficacy was assessed by the Arthritis Self-Efficacy Scale (ASES), designed to assess participants' confidence in their ability to manage symptoms of arthritis. The final score ranges 10–100 in 10-point increments, with higher values representing higher self-efficacy. ASES has previously been used to evaluate patient education programs for individuals with arthritis and is validated in Swedish (22,23). In BOA, only the scales assessing self-efficacy for pain and other symptoms have been included. For this study, we used only the scale assessing pain self-efficacy, which was converted into a 1–10 scale with a 1-point increment to facilitate interpretation of the results.

Comorbidities were measured using the Charnley classification, which categorizes individuals into 3 categories: A (unilateral OA of knee or hip), B (bilateral OA in both knees or both hips), or C (OA in multiple joint sites, e.g., hip and knee) and/or the presence of any other disease that affects walking ability (28,29). No other measure of comorbidity was available. Fear of movement was assessed by the question "Are you afraid your joints will be injured by physical training/activity? (Yes/No)."

Any kind of surgery was considered, e.g., meniscectomy, osteotomy, partial joint replacement, and TJR. Considering that the joint that received replacement could not be considered as the index joint for the treatment, none of the reported index joints had received TJR before enrolling in the program. However, participants who received TJR in the contralateral joint >12 months prior to enrollment could be included in the program. At baseline, participants were asked if they had previously sought care for their joint problem, but no information on the treatment sought was collected.

Self-reported radiographs prior to enrollment were recorded and divided into 4 categories: no radiographs, radiographs received >6 months before enrollment, radiographs received <6 months before enrollment, or does not know. Previous consultations with a physical therapist for the problems in the index joint were self-reported by the participants (yes/no). No information regarding treatments received was recorded. Participants were also asked to report whether they were on a waiting list for receiving joint surgery at the time of enrollment (yes/no). No data were available for the 3- and 12-month follow-up.

Statistical analysis. To account for differences linked to the affected joint, all the analyses were performed separately for patients with hip and knee OA. Normality was assessed through visual inspection of histograms and assessment of Q-Q plots. We used the independent *t*-test to compare characteristics of patients willing and unwilling to undergo surgery at baseline. The chi-square test and Z test with Bonferroni correction were used for categorical variables and to assess the proportion of crossovers from willing to undergo surgery at baseline to unwilling to undergo surgery at the various follow-ups. Alpha level was set at 0.05.

Logistic regression models were used to study the association between pain intensity, walking difficulties, self-efficacy, and fear of movement with the willingness to undergo surgery. Separate models were built to study the association between the independent variables and the dependent variable at 2 time points, 3 months and 12 months. In the first model, we studied the association at 3 months between pain, walking difficulties, self-efficacy, and fear of movement with the participant's willingness to undergo surgery. In the second model, we studied the association at 12 months between pain, walking difficulties, self-efficacy and fear of movement with the participant's willingness to undergo surgery. The analyses were adjusted for sex, age, BMI, willingness to undergo surgery at the previous follow-up, pain intensity at the previous follow-up,

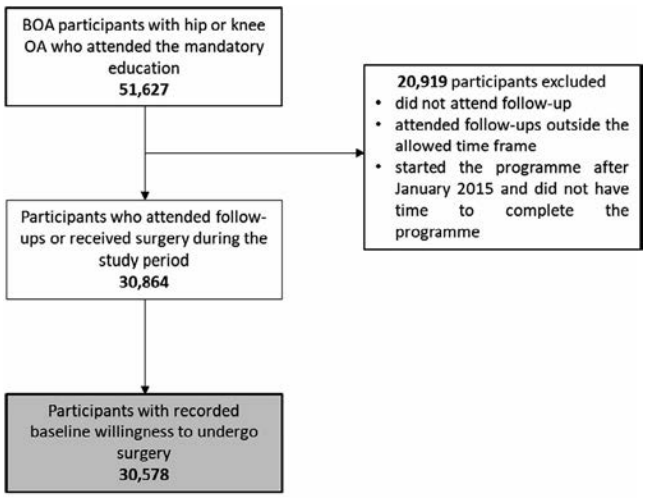

Figure 1. Participant selection flow chart. BOA = Better Management of People with Osteoarthritis; OA = osteoarthritis.

previous surgery (index and contralateral), level of education, previous care sought for the joint problem, previous consultation with a physical therapist, previous radiographs, and whether the participant was on a waiting list for joint surgery. Results are presented as odds ratios with 95% confidence intervals. All statistical analyses were conducted using SPSS software, version 25.0.

RESULTS

From 2008 to 2016, 51,627 individuals with hip or knee OA were recorded in the BOA register and were eligible for this study, of whom 30,578 filled the inclusion criteria (knee OA: 20,649; hip OA: 9,929) (Figure 1 and Table 1).

Table 1. Characteristics of the BOA participants*

Characteristic	Included (n = 30,578)		Excluded (n = 21,049)	
	No.	Value	No.	Value
Age, years	30,578	66.7 ± 9.0	20,987	65.4 ± 9.7
Body mass index	30,081	27.8 ± 4.8	20,515	28.2 ± 4.9
Pain (0–10)	30,509	5.3 ± 1.9	20,897	5.3 ± 2.0
ASES pain (1–10)	29,947	6.3 ± 1.9	19,886	6.3 ± 1.9
EQ-5D-5L VAS (1–100)	24,605	67.1 ± 19.0	18,083	65.8 ± 19.5
Sex, %				
Men	9,274	30.3	6,424	30.6
Women	21,304	69.7	14,563	69.4
Willingness to undergo surgery, %				
No	22,700	74.2	15,567	76.1
Yes	7,878	25.8	4,884	23.9
Comorbidities, %				
Charnley class A	11,761	38.5	7,830	37.3
Charnley class B	6,036	19.8	3,809	18.1
Charnley class C	12,743	41.7	9,320	44.5
Fear of movement, %				
No	25,895	84.4	16,925	82.3
Yes	4,798	15.6	3,629	17.7
Walking difficulties, %				
No	5,939	19.2	3,910	18.8
Yes	24,482	80.8	16,930	81.2
Education, %				
Primary school	10,622	34.4	7,060	33.8
High school	10,997	36.9	7,982	38.2
University	8,871	29.1	5,845	28.0
Previous care sought for the joint problem, %				
No	877	2.9	682	3.3
Yes	29,502	97.1	20,188	96.7
Previous radiographs index joint, %				
No	6,345	20.8	4,540	21.6
Yes, >6 months	9,756	32.0	6,570	31.3
Yes, <6 months	14,132	46.4	9,675	46.1
Does not know	225	0.7	186	0.9
Previous physical therapist consultation (index joint), %				
No	16,053	52.7	11,593	55.3
Yes	14,433	47.3	9,363	44.7
Previous surgery index joint, %				
No	26,717	87.6	18,203	86.7
Yes	3,799	12.4	2,789	13.3
Surgery contralateral, %				
No	27,218	89.4	18,791	89.7
Yes	3,227	10.6	2,147	10.3
Waiting list for joint surgery (index joint), %				
No	29,697	97.6	20,549	98.2
Yes	721	2.4	374	1.8
Surgery during study period, %				
No	26,314	86.1	–	–
Before 3 months	78	0.3	–	–
Between 3 and 12 months	4,186	13.7	–	–

* Values are the mean ± SD unless indicated otherwise. ASES = Arthritis Self Efficacy Score; BOA = Better Management of People with Osteoarthritis; EQ-5D-5L = 5-level EuroQol 5-domain instrument; VAS = visual analog scale.

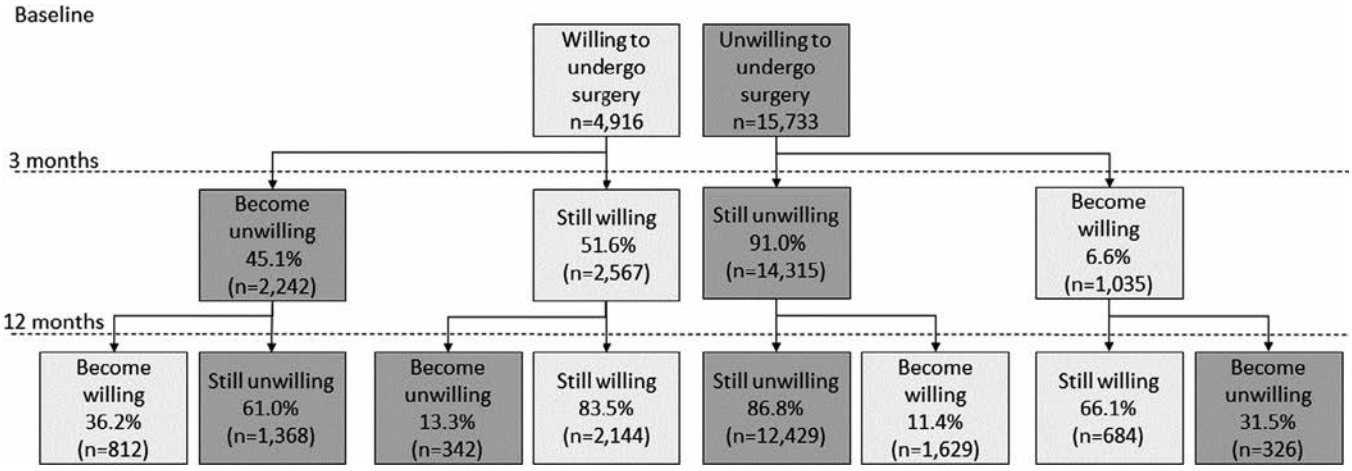

Figure 2. Proportion of individuals with knee osteoarthritis reconsidering their willingness to undergo surgery at 3 and 12 months from the intervention. Percentages are reported separately for individuals willing to consider surgery at baseline and for those unwilling to consider surgery. Percentages represent the proportion of individuals from the level above who either changed or did not change their mind at the follow-up. A percentage that does not reach 100% indicates individuals without willingness to undergo surgery recorded at follow-up.

Shift in willingness to undergo surgery. At baseline, 4,916 participants (24%) with knee OA and 2,962 (30%) with hip OA were willing to undergo surgery. Of these, 45.1% of those with knee OA (n = 2,242) and 30 % of those with hip OA (n = 901) became unwilling after 3 months. At 12 months, 61% of those with knee OA (n = 1,368) and 45.3% of those with hip OA (n = 408) who became unwilling were still unwilling. Among the individuals unwilling to undergo surgery at baseline, 6.6% of those with knee OA (n = 1,035) and 11.8% of those with hip OA (n = 820) became willing at 3 months. At 12 months, 66.1% of those with knee OA (n = 684) and 80.5% of those with hip OA (n = 660) who became willing were still willing to undergo surgery.

Overall, 34.8% of the individuals with knee OA (n = 1,710) and 19.0% of those with hip OA (n = 564) who were willing to undergo surgery at baseline became willing during the study period and were unwilling at 12 months. On the other hand, 14.7% of those with knee OA (n = 2,313) and 26.8% of those with hip OA (n = 1,871) changed from being unwilling to consider surgery at baseline to be willing at the 12-month follow-up. Levels of pain, self-efficacy, and quality of life at the 3 follow-ups in relation to willingness are reported in Supplementary Tables 1 and 2 (available on the *Arthritis Care Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24486/abstract) and Figures 2 and 3.

Baseline comparison and factors associated with willingness to undergo surgery. Baseline characteristics of the included participants in Table 2 show that more individuals with hip OA than knee OA were willing to consider surgery at baseline.

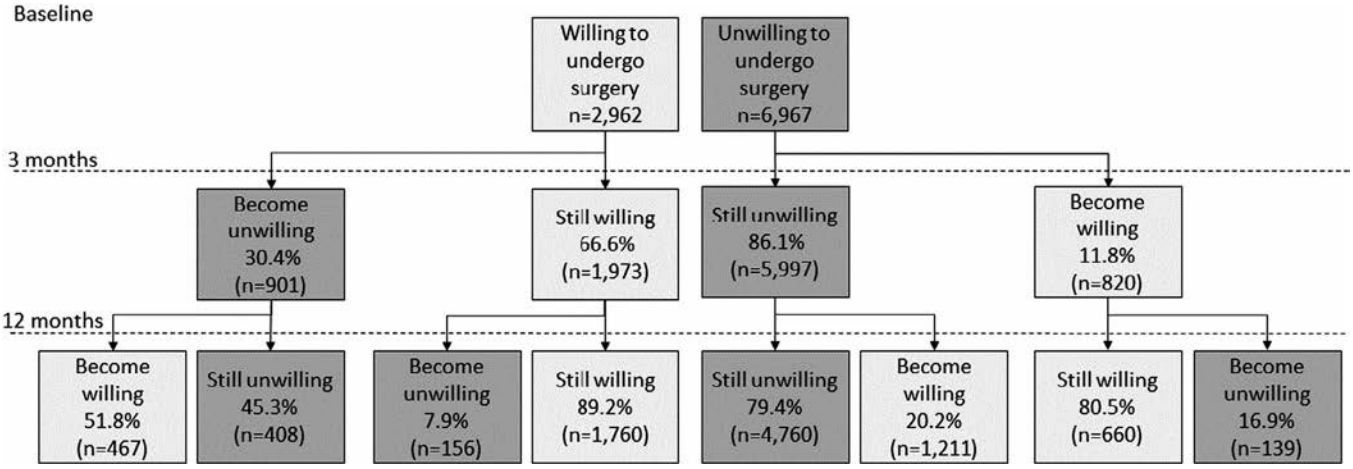

Figure 3. Proportion of individuals with hip osteoarthritis reconsidering their willingness to undergo surgery at 3 and 12 months from the intervention. Percentages are reported separately for individuals willing to consider surgery at baseline and for those unwilling to consider surgery. Percentages represent the proportion of individuals from the level above who either changed or did not change their mind at the follow-up. A percentage that does not reach 100% indicates individuals without willingness to undergo surgery recorded at follow-up.

Table 2. Baseline characteristics of BOA participants based on the affected joint and willingness to undergo surgery*

	Knee OA, unwilling (n = 15,733)		Knee OA, willing (n = 4,916)			Hip OA, unwilling (n = 6,967)		Hip OA, willing (n = 2,962)		
	No.	Value	No.	Value	*P*	No.	Value	No.	Value	*P*
Age	15,733	66.6 ± 8.9	4,916	66.0 ± 9.2	<0.001†	6,967	67.3 ± 9.0	2,962	67.3 ± 9.0	0.941
Body mass index	15,467	27.9 ± 4.8	4,837	29.3 ± 4.9	<0.001†	6,862	26.6 ± 4.3	2,915	27.6 ± 4.5	<0.001†
Pain (0–10)	15,696	4.9 ± 1.9	4,907	6.4 ± 1.6	<0.001†	6,953	4.9 ± 1.9	2,953	6.5 ± 1.5	<0.001†
ASES pain (1–10)	15,402	6.8 ± 1.7	4,823	5.4 ± 1.8	<0.001†	6,814	6.5 ± 1.7	2,908	5.0 ± 1.8	<0.001†
EQ-5D-5L VAS (0–100)	12,677	70.3 ± 17.8	3,860	60.0 ± 19.7	<0.001†	5,712	69.0 ± 17.9	2,356	56.9 ± 19.8	<0.001†
Sex, %										
Men	4,196	26.7	1,985	40.4	<0.05†	1,892	27.2	1,201	40.5	<0.05†
Women	11,537	73.3	2,931	59.6	<0.05†	5,075	72.8	1,761	59.5	<0.05†
Charnley class, %										
A	6,354	40.4	1,624	33.3	<0.05†	2,719	39.0	1,064	35.9	<0.05†
B	3,759	23.9	1,192	24.2	<0.05†	798	11.5	287	9.7	<0.05†
C	5,604	35.6	2,092	42.6	<0.05†	3,437	49.3	1,610	54.4	<0.05†
Fear of movement, %										
No	13,445	85.8	3,727	76.3	<0.05†	6,117	88.2	2,394	81.3	<0.05†
Yes	2,221	14.2	1,156	23.7	<0.05†	819	11.8	550	18.7	<0.05†
Walking difficulties, %										
No	3,879	24.8	268	5.5	<0.05†	1,679	24.2	113	3.8	<0.05†
Yes	11,763	75.2	4,631	94.5	<0.05†	5,248	75.8	2,840	96.2	<0.05†
Education, %										
Primary school	5,175	33.0	1,945	39.7	<0.05†	2,310	33.3	1,192	40.3	<0.05†
High school	5,683	36.2	1,852	37.8	<0.05†	2,420	34.9	1,042	35.3	>0.05
University	4,831	30.8	1,105	22.5	<0.05†	2,213	31.9	722	24.4	<0.05†
Previous care sought for joint problem, %										
No	488	3.1	51	1.0	<0.05†	296	4.3	42	1.4	<0.05†
Yes	15,134	96.9	4,939	99.0	<0.05	6,604	95.7	2,907	98.6	<0.05†
Previous radiographs index joint, %										
No	3,631	23.2	486	9.9	<0.05†	1,899	27.4	329	11.1	<0.05†
Yes, >6 months	5,204	33.2	1,687	34.3	>0.05	1,992	28.8	873	29.5	>0.05
Yes, <6 months	6,711	42.8	2,705	55.0	<0.05†	2,976	43.0	1,740	58.9	<0.05†
Does not know	119	0.8	36	0.7	>0.05	57	0.8	13	0.4	<0.05†
Previous physical therapist consultation (index joint), %										
No	8,432	53.7	2,392	48.8	<0.05†	3,756	54.1	1,473	49.9	<0.05†
Yes	7,257	46.3	2,512	51.2	<0.05†	3,284	45.9	1,480	50.1	<0.05†
Previous surgery index joint, %										
No	13,306	84.7	3,692	75.2	<0.05†	6,833	98.4	2,886	97.6	<0.05†
Yes	2,398	15.3	1,217	24.8	<0.05†	113	1.6	71	2.4	<0.05†
Surgery contralateral, %										
No	14,117	90.1	3,991	81.6	<0.05†	6,531	94.1	2,579	87.4	<0.05†
Yes	1,548	9.9	899	18.4	<0.05†	407	5.9	373	12.6	<0.05†
Waiting list for joint surgery (index joint), %										
No	15,555	99.4	4,526	92.4	<0.05†	6,879	99.4	2,737	92.8	<0.05†
Yes	96	0.6	370	7.6	<0.05†	44	0.6	211	7.2	<0.05†
Surgery during study period, %										
No	15,028	95.5	3,667	74.6	<0.05†	6,102	87.6	1,517	51.2	<0.05†
Yes	705	4.5	1,249	25.4	<0.05†	865	12.4	1,445	48.8	<0.05†

* Values are the mean ± SD unless indicated otherwise. Knee OA participants total: n = 20,649; hip OA participants total: n = 9,929. ASES = Arthritis Self Efficacy Score; BOA = Better Management of People with Osteoarthritis; EQ-5D-5L = 5-level EuroQol 5-domain instrument; OA = osteoarthritis; VAS = visual analog scale.
† Statistically significant.

Table 3. Factors associated with the willingness to undergo surgery at the 3-month and 12-month follow-ups*

Factors measured	Knee OA (n = 20,649)	Hip OA (n = 9,929)
3-month follow-up		
Pain (0–10)	1.61 (1.54, 1.67)†	1.70 (1.61, 1.80)†
EQ-5D-5L VAS (0–100)	1.00 (0.99, 1.01)	0.99 (0.99, 1.00)
Fear of movement (y/n)	1.49 (1.21, 1.82)†	1.56 (1.15, 2.10)†
ASES pain (1–10)	0.73 (0.70, 0.75)†	0.69 (0.66, 0.73)†
Walking difficulties (y/n)	3.46 (2.85, 4.20)†	4.03 (3.04, 5.35)†
12-month follow-up		
Pain (0–10)	1.79 (1.71, 1.88)†	1.86 (1.74, 1.99)†
EQ-5D-5L VAS (0–100)	0.99 (0.99, 1.00)	0.99 (0.99, 1.00)
Fear of movement (y/n)	1.51 (1.23, 1.85)†	1.43 (1.02, 2.00)†
ASES pain (1–10)	0.79 (0.76, 0.83)†	0.76 (0.71, 0.80)†
Walking difficulties (y/n)	3.55 (2.83, 4.47)†	5.25 (3.61, 87.63)†

* Values are the odds ratio (95% confidence interval). Analyses are adjusted for sex, age, body mass index, willingness to undergo surgery at the previous visit (baseline/3-month), pain at the previous visit (baseline/3-month), previous surgery (either knee or either hip), and education. Values for dichotomous variables are reported for the presence of the factor (absence used as reference category). ASES = Arthritis Self Efficacy Score; EQ-5D-5L = 5-level EuroQol 5-domain instrument; OA = osteoarthritis; VAS = visual analog scale.
† Statistically significant.

Individuals willing to consider surgery were on average younger and had higher BMI, higher baseline pain, lower self-efficacy, and lower quality of life compared to those who did not consider surgery. In addition, patients willing to undergo surgery more often had a Charnley score of C and walking difficulties, more often had consulted a physical therapist in the past, more often had received radiographs in the index knee in the last 6 months, more often had received surgery in the index or contralateral joint, and were more often on a waiting list to receive surgery in the index joint.

Table 3 shows that, regardless of the joint affected (hip or knee), patients with higher pain, fear of movement, and walking difficulties at 3 months had higher odds of being willing to undergo surgery at 3 months. By contrast, individuals with a higher level of self-efficacy at 3 months had lower odds of becoming willing to undergo surgery. Similarly, regardless of the joint affected (hip or knee), patients with a higher level of pain, fear of movement, and walking difficulties at 12 months had higher odds of becoming willing to undergo surgery at 12 months. Having a higher level of self-efficacy and quality of life at 12 months was associated with lower odds of becoming willing to undergo surgery at 12 months (Table 3).

DISCUSSION

In this study conducted in >30,000 individuals with knee or hip OA, we showed that 30% to 45% of the patients willing to undergo surgery no longer considered surgery as a therapeutic option after receiving a first-line intervention (3 months) including exercise and education. On the other hand, 7% to 12% of those who were not willing at baseline changed their mind. Overall, 35% of the patients with knee OA and 19% of those with hip OA willing to undergo surgery at baseline were no longer considering surgery at 12 months. Overall, >90% of the individuals included in this study had previously sought care for the joint problem, with roughly 50% who consulted a physical therapist in the past and 50% who received radiographs in the 6 months before the intervention. However, <3% of the participants were on a waiting list for surgery, suggesting that the sample may represent a population with moderate symptomatology accessing a first-line intervention after having previously undergone other treatments.

As expected, those individuals willing to undergo surgery at baseline appeared to have an overall worse disease severity, lower quality of life, and lower self-efficacy, regardless of the joint affected by OA. Nonetheless, many of these participants reconsidered their position about surgery. Due to the design of the study, little can be said about the influence of the provided intervention on the desire for surgery. However, our results show that improvement in OA symptoms and walking difficulties are associated with higher odds of becoming unwilling to undergo surgery. Worsening pain may instead explain the reconsiderations shown between the 3- and 12-month follow-ups, when no treatment was delivered, and the effect of the intervention is expected to subside (see Supplementary Tables 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24486/abstract) (30,31). Prolonged adherence to an exercise regime, longer interventions, booster sessions, or digitally delivered programs may thereby minimize reconsideration about surgery in the months following the intervention (32,33).

Individuals unwilling to consider surgery at baseline appear to be less prone to become willing after receiving the intervention. Differences in baseline characteristics may help to explain the results. Individuals unwilling to undergo surgery at baseline appeared to have less severe symptoms and showed higher levels of self-efficacy, which was associated with lower odds of desiring surgery and has been linked to better outcomes from self-management interventions (34,35). Nevertheless, certain individuals unwilling to consider surgery changed their mind after the treatment or at 12 months. These individuals seem to have experienced an overall worsening of the symptoms, which may have led them to consider surgery as a therapeutic option for their joint problems (see Supplementary Tables 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24486/abstract) (36).

Overall, our results are in line with a study showing a similar rate of reconsideration (30% became unwilling, 6% became willing) after 6 weeks of participation in a digital self-management program based on the BOA program (17). Despite the fact that both of the studies are observational and cannot establish causality, results from an RCT have shown that >60% of the patients eligible for surgery who received a 12-week first-line intervention decided not to undergo surgery up to the 2-year follow-up (1,15). A recent study analyzing the willingness to undergo surgery in individuals

with knee OA from the Osteoarthritis Initiative cohort showed a lower rate of reconsideration, with 16% of the participants becoming more willing to undergo surgery and 14% becoming less willing over a 2-year period where no structured intervention was provided (37). However, analyzed together, these results suggest that first-line interventions have the potential to reduce willingness to undergo surgery in a large part of the OA population, including individuals eligible for TJR and those with more moderate symptoms. Caution in the interpretation is needed, considering that the individuals who decide to undergo treatment for their problems may be more prone to reassess therapeutic options than those who are not seeking care.

Pain was significantly associated with the willingness to undergo surgery at both follow-ups, suggesting that a reduction of 1 unit in the pain intensity can lead to 60–80% lower odds of being willing to undergo surgery. However, the presence of walking difficulties also showed a strong association, increasing by 3–5-fold the odds of being willing to undergo surgery. Despite the fact that pain is often considered the most important factor driving care-seeking behavior, qualitative evidence showed that patients with OA often consider the use of quantitative measures of pain to deem eligibility to surgery to be inappropriate due to the inability of these scales to capture the real impact of pain on a person's life (38). Thus, measures of walking difficulties and physical disabilities may help to capture the experience of a person with OA, explaining their strong association with the willingness to undergo surgery. Thus, addressing perceived walking difficulties may lead to less surgery consideration.

Among the other factors analyzed, quality of life was not associated with the willingness to undergo surgery. Factors external to the joint disease may influence the quality of life without necessarily impacting the willingness to undergo surgery. On the other hand, higher levels of self-efficacy reduced the odds of being willing to undergo surgery. Focusing on function and participation rather than solely on symptom reduction may further reduce surgery consideration.

Patients with hip OA appear to benefit less from first-line interventions when compared to patients with knee OA (30,33,39,40). In this study, 30% of the individuals with OA who were willing to undergo surgery changed their mind after the treatment. However, those with hip OA were less likely to become unwilling to undergo surgery and more often received TJR during the study time than patients with knee OA. In addition, 26% of the individuals with hip OA who were unwilling to undergo surgery at baseline became willing by the end of the study period, while only 19% made the opposite shift. This trend is reverted in individuals with knee OA. Despite the differences in the rate of reconsideration, all the analyzed factors showed a similar association with the willingness to undergo surgery across the joints, suggesting that differences in the rate of surgery reconsideration are likely due to joint-specific differences in pain, symptom reduction, and surgery indication.

Some limitations need to be discussed. First, this was an observational study and, therefore, the effect of the treatment on the willingness for surgery cannot be asserted. In addition, several individuals did not have data recorded for 1 or both of the follow-ups and could not be included in the study. The exclusion of these individuals may have influenced our findings and should be taken into account when interpreting the results. Second, we do not know whether the patients who were willing to undergo surgery would be deemed eligible for surgery at the end of the intervention. This lack of information on surgery eligibility implies that reconsideration may not result in a direct change in the number of surgical procedures. However, individuals unwilling to undergo surgery have been shown to be less likely to receive TJR than those who are willing (14). Third, individual decision-making on important health care concerns such as surgery is complex and cannot be explained solely by the factors investigated in this study. Individuals at a later stage of the disease may have different expectations from an intervention than individuals at earlier stages. The limited information regarding the stage of the diseases (e.g., disease duration, date of diagnosis) may thus limit the applicability of these results. Finally, cultural differences between countries may exist and may somewhat limit the generalizability of the results outside Sweden.

Results from this large cohort show reduced willingness to undergo surgery by 30% to 45% after completion of a first-line intervention and show that this reduction is partially maintained for up to 12 months. Walking difficulties appear to be central in determining a patient's willingness to undergo surgery and may be as important as pain, especially when pain is measured on a quantitative scale. Finally, individuals' preferences are key in the care process of every disease and should always be considered. However, willingness to undergo surgery should be used with care in the decision process leading to surgery, in light of its temporal variability, especially if the patient seeking surgery has not yet undertaken a first-line intervention.

ACKNOWLEDGMENTS

The authors thank all participating patients and physical therapists reporting data to the BOA registry and others involved in the BOA program.

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. Dell'Isola 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. Dell'Isola, Jönsson, Rolfson, Cronström, Englund, Dahlberg.

Acquisition of data. Jönsson, Rolfson.

Analysis and interpretation of data. Dell'Isola.

REFERENCES

1. Skou ST, Roos EM, Laursen MB, Rathleff MS, Arendt-Nielsen L, Simonsen O, et al. A randomized, controlled trial of total knee replacement. N Engl J Med 2015;373:1597–606.

2. Turkiewicz A, Petersson IF, Bjork J, Hawker G, Dahlberg LE, Lohmander LS, et al. Current and future impact of osteoarthritis on health care: a population-based study with projections to year 2032. Osteoarthritis Cartilage 2014;22:1826–32.
3. Ackerman IN, Bohensky MA, Zomer E, Tacey M, Gorelik A, Brand CA, et al. The projected burden of primary total knee and hip replacement for osteoarthritis in Australia to the year 2030. BMC Musculoskelet Disord 2019;20:90.
4. Tian W, DeJong G, Brown M, Hsieh CH, Zamfirov ZP, Horn SD. Looking upstream: factors shaping the demand for postacute joint replacement rehabilitation. Arch Phys Med Rehabil 2009;90:1260–8.
5. Dreinhofer KE, Dieppe P, Sturmer T, Grober-Gratz D, Floren M, Gunther KP, et al. Indications for total hip replacement: comparison of assessments of orthopaedic surgeons and referring physicians. Ann Rheum Dis 2006;65:1346–50.
6. Mandl LA. Determining who should be referred for total hip and knee replacements. Nat Rev Rheumatol 2013;9:351–7.
7. Mota RE, Tarricone R, Ciani O, Bridges JF, Drummond M. Determinants of demand for total hip and knee arthroplasty: a systematic literature review. BMC Health Serv Res 2012;12:225.
8. Barlow T, Griffin D, Barlow D, Realpe A. Patients' decision making in total knee arthroplasty: a systematic review of qualitative research. Bone Joint Res 2015;4:163–9.
9. McHugh GA, Campbell M, Silman AJ, Kay PR, Luker KA. Patients waiting for a hip or knee joint replacement: is there any prioritization for surgery? J Eval Clin Pract 2008;14:361–7.
10. Hawker GA. Who, when, and why total joint replacement surgery? The patient's perspective. Curr Opin Rheumatol 2006;18:526–30.
11. Hawker GA, Wright JG, Badley EM, Coyte PC for the Toronto Arthroplasty Health Services Research Consortium. 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.
12. Hawker GA, Wright JG, Glazier RH, Coyte PC, Harvey B, Williams JI, et al. The effect of education and income on need and willingness to undergo total joint arthroplasty. Arthritis Rheum 2002;46:3331–9.
13. Riddle DL, Jiranek WA, Hayes CW. Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: a multicenter longitudinal cohort study. Arthritis Rheumatol 2014;66:2134–43.
14. 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.
15. Skou ST, Roos EM, Laursen MB, Rathleff MS, Arendt-Nielsen L, Rasmussen S, et al. Total knee replacement and non-surgical treatment of knee osteoarthritis: 2-year outcome from two parallel randomized controlled trials. Osteoarthritis Cartilage 2018;26:1170–80.
16. Svege I, Nordsletten L, Fernandes L, Risberg MA. Exercise therapy may postpone total hip replacement surgery in patients with hip osteoarthritis: a long-term follow-up of a randomised trial. Ann Rheum Dis 2015;74:164–9.
17. Cronström A, Nero H, Dahlberg LE. Factors associated with patients' willingness to consider joint surgery after completion of a digital osteoarthritis treatment program: a prospective cohort study. Arthritis Care Res (Hoboken) 2019;71:1194–201.
18. Teoh LS, Eyles JP, Makovey J, Williams M, Kwoh CK, Hunter DJ. Observational study of the impact of an individualized multidisciplinary chronic care program for hip and knee osteoarthritis treatment on willingness for surgery. Int J Rheum Dis 2017;20:1383–92.
19. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007;335:806–8.
20. Thorstensson CA, Garellick G, Rystedt H, Dahlberg LE. Better management of patients with osteoarthritis: development and nationwide implementation of an evidence-based supported osteoarthritis self-management programme. Musculoskeletal Care 2015;13:67–75.
21. Ageberg E, Roos EM. Neuromuscular exercise as treatment of degenerative knee disease. Exerc Sport Sci Rev 2015;43:14–22.
22. Ageberg E, Nilsdotter A, Kosek E, Roos EM. Effects of neuromuscular training (NEMEX-TJR) on patient-reported outcomes and physical function in severe primary hip or knee osteoarthritis: a controlled before-and-after study. BMC Musculoskelet Disord 2013;14:232.
23. Socialstyrelsen. Nationella riktlinjer för rörelseorganens sjukdomar. Reumatoid artrit, axial spondylartrit, psoriasisartrit, artros och osteoporos. Socialstyrelsen 2021. URL: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/nationella-riktlinjer/2021-1-7137.pdf.
24. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum 1991;34:505–14.
25. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis Rheum 1986;29:1039–49.
26. Bilbao A, Garcia-Perez L, Arenaza JC, Garcia I, Ariza-Cardiel G, Trujillo-Martin E, et al. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. Qual Life Res 2018;27:2897–908.
27. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001;33:337–43.
28. Bjorgul K, Novicoff WM, Saleh KJ. Evaluating comorbidities in total hip and knee arthroplasty: available instruments. J Orthop Traumatol 2010;11:203–9.
29. Charnley J, Halley DK. Rate of wear in total hip replacement. Clin Orthop Relat Res 1975;112:170–9.
30. Dell'Isola A, Jönsson T, Ranstam J, Dahlberg LE, Ekvall Hansson E. Education, home exercise, and supervised exercise for people with hip and knee osteoarthritis as part of a nationwide implementation program: data from the Better Management of Patients With Osteoarthritis Registry. Arthritis Care Res (Hoboken) 2020;72:201–7.
31. Jonsson T, Eek F, Dell'Isola A, Dahlberg LE, Ekvall Hansson E. The Better Management of Patients with Osteoarthritis Program: outcomes after evidence-based education and exercise delivered nationwide in Sweden. PLoS One 2019;14:e0222657.
32. Nero H, Dahlberg J, Dahlberg LE. A 6-week web-based osteoarthritis treatment program: observational quasi-experimental study. J Med Internet Res 2017;19:e422.
33. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. Fransen M, ed. Chichester (UK): John Wiley; 2015.
34. McAuley E, Szabo A, Gothe N, Olson EA. Self-efficacy: implications for physical activity, function, and functional limitations in older adults. Am J Lifestyle Med 2011;5:361–9.
35. 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.
36. Cronstrom A, Dahlberg LE, Nero H, Hammarlund CS. "I was considering surgery because I believed that was how it was treated": a qualitative study on willingness for joint surgery after completion of a digital management program for osteoarthritis. Osteoarthritis Cartilage 2019;27:1026–32.

37. Bendich I, Halvorson RT, Ward D, Nevitt M. Predictors of a change in patient willingness to have total knee arthroplasty: insights from the Osteoarthritis Initiative. Knee 2020;27:667–75.
38. 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.
39. Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. Cochrane Database Syst Rev 2014:CD007912.
40. Dell'Isola A, Jönsson T, Nero H, Eek F, Dahlberg L. Factors associated with the outcome of a first-line intervention for patients with hip or knee osteoarthritis or both: data from the BOA Register. Phys Ther 2020;100:1771–81.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 828–832
DOI 10.1002/acr.24494

BRIEF REPORT

Association of Hydroxychloroquine Use With Decreased Incident Atrial Fibrillation in Systemic Lupus Erythematosus

Alisha Gupta,[1] Kelly J. Shields,[2] Susan Manzi,[3] Mary Chester Wasko,[3] and Tarun S. Sharma[3]

Objective. To study the relationship between hydroxychloroquine (HCQ) use and new-onset atrial fibrillation in patients with systemic lupus erythematosus (SLE).

Methods. A retrospective cohort of adult patients with SLE was constructed from December 1, 2014 to May 30, 2017. Patients were categorized as either HCQ users or nonusers. The primary outcome was incident atrial fibrillation. Secondary outcomes included incident ventricular arrhythmias (composite of ventricular tachycardia, ventricular fibrillation, or torsades de pointes). Outcomes were adjudicated by review of the electronic health record. Statistical analyses included simple and multivariable logistic regression tests to estimate the association between HCQ use and incident atrial fibrillation after adjusting for relevant confounders. Propensity score matching analysis was completed.

Results. Our study included 1,647 patients with SLE, of which 917 were HCQ users and 730 were nonusers. A total of 23 atrial fibrillation events occurred, including 3 in HCQ users and 20 in nonusers. Logistic regression analysis showed an odds ratio (OR) of 0.12 (95% confidence interval [95% CI] 0.034–0.39, $P = 0.0005$) for incident atrial fibrillation and 2.39 (95% CI 0.25–23.0, $P = 0.45$) for ventricular arrhythmias. Results remained significant in the fully adjusted and propensity score–matched models.

Conclusion. In this exploratory study, HCQ use was associated with an 88% decrease in the risk of incident atrial fibrillation in patients with SLE. Considering the increased cardiovascular risk in SLE, incorporation of HCQ into the regimen may be beneficial for both disease manifestations and reducing the risk of atrial fibrillation. Further studies would be needed to confirm the antifibrillatory benefit of this relatively safe and low-cost medication.

INTRODUCTION

Cardiac manifestations of systemic lupus erythematosus (SLE) can include myocarditis, pericarditis, valvular disease, thrombosis, and cardiac conduction defects (1,2). Rhythm abnormalities occur in >15% of lupus patients, often in association with other cardiac comorbidities or active lupus (2). A recent case–control study demonstrated higher incidence of atrial fibrillation in patients with SLE, a 4.5-fold higher incidence in men with SLE as compared to women with SLE, and a correlation with disease activity (3). Another recent study showed that the prevalence of resting electrocardiogram (EKG) changes, including atrial fibrillation, that are predictive of future clinical cardiovascular disease was greater among patients with SLE than the general population (prevalence of up to 24% versus 17%) (4).

The antimalarial drug, hydroxychloroquine, has been a cornerstone therapy in the management of SLE. Early studies of its parent molecules, quinidine and chloroquine, demonstrated their antiarrhythmic benefit; however, their use declined over the years due to concerns of toxicity (5). Hydroxychloroquine has been shown to prevent shortening of atrial effective refractory period, a primary pathophysiologic mechanism promoting atrial fibrillation (6,7). Despite its potential antiarrhythmic effects and known cardiovascular benefits (8), the relationship between hydroxychloroquine usage and incident atrial fibrillation has not been described. The purpose of this study was to examine the effect of

[1]Alisha Gupta, MD: Emory University, Atlanta, Georgia; [2]Kelly J. Shields, PhD: Highmark Health, Pittsburgh, Pennsylvania; [3]Susan Manzi, MD, MPH, Mary Chester Wasko, MD, MSc, Tarun S. Sharma, MD: Allegheny Health Network, Pittsburgh, Pennsylvania.

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

Address correspondence to Tarun S. Sharma, MD, Division of Rheumatology, Allegheny Health Network, Suite 2615 North Tower, 4800 Friendship Avenue, Pittsburgh, PA 15217. Email: tarun.sharma@ahn.org.

Submitted for publication May 31, 2020; accepted in revised form October 20, 2020.

SIGNIFICANCE & INNOVATIONS

- A few in vitro and animal model studies have explored the antiarrhythmic effect of hydroxychloroquine. Chloroquine, a closely related antimalarial, has been shown to have more potent antifibrillatory properties compared to its parent molecule quinidine, a class Ia antiarrhythmic agent. Our study is the first to demonstrate hydroxychloroquine's protective relationship against incident atrial fibrillation in patients with systemic lupus erythematosus (SLE).
- While hydroxychloroquine is a cornerstone therapy in management of SLE, with a favorable safety profile, antithrombotic effects, and cardiovascular benefits, its protective association against atrial fibrillation in our exploratory study is a significant finding in light of the high cardiovascular risk in this population.

hydroxychloroquine use on incident atrial fibrillation in a retrospective cohort of patients with SLE.

PATIENTS AND METHODS

Patients. Institutional review board approval was obtained, and informed consent was waived for this observational study. From December 1, 2014 to May 30, 2017, a retrospective cohort of adult patients with SLE was assembled at Allegheny Health Network, a tertiary care academic center in western Pennsylvania with a fully implemented electronic health record (EHR) system. A lupus diagnosis was based on the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code M32, entered at least once by a rheumatologist. Patients with a previous history of atrial fibrillation were excluded from the study. In addition, patients with an incident atrial fibrillation event occurring in the first year of observation (December 1, 2014 through December 1, 2015) were excluded in order to allow for a run-in period and enhance validity.

Outcome measures. Eligible patients were divided into hydroxychloroquine users and nonusers according to hydroxychloroquine use during the observation period of December 1, 2014 to May 30, 2017. The primary outcome was incident atrial fibrillation, defined as ICD-10 code I48, and was subsequently validated via EHR review and EKG confirmation. For the incident event to be attributed to the hydroxychloroquine user group, hydroxychloroquine use had to be current at the time of the event. Duration of hydroxychloroquine use prior to and dose at the time of incident atrial fibrillation were recorded via manual EHR review.

The secondary outcome was a composite of incident ventricular arrhythmias, namely, ventricular fibrillation (VF), ventricular tachycardia (VT), or torsades de pointes. Secondary end points were also validated via manual EHR review. Eligible patients were then matched to prescription claims data provided by a single, common integrated delivery and finance system in order to determine immunosuppressive medication and glucocorticoid use during the study period for patients with pharmacy benefits coverage. Baseline characteristics were compared between groups, including demographic characteristics (age, sex, and ethnicity), risk factors for atrial fibrillation and medical comorbidities (body mass index, smoking status, alcohol use, diabetes mellitus, hypertension, coronary artery disease, heart failure, stroke or transient ischemic attack, peripheral vascular disease, thyroid disease, chronic kidney disease, end-stage renal disease, liver dysfunction, obstructive sleep apnea, and chronic obstructive pulmonary disease), and medication use (immunosuppressive, glucocorticoids, and antiarrhythmic medications).

Statistical analyses. Patient follow-up started on December 1, 2014 and continued until the incident atrial fibrillation event, death, or end of the observation period. Patient characteristics were summarized as mean ± SD for normally distributed variables, median (interquartile range) for skewed variables, and frequency (%) for categorical variables. A standard *t*-test was used for continuous, normally distributed variables, and the nonparametric Wilcoxon's rank sum test (Mann-Whitney U test) was used for continuous variables that did not meet normality requirements. Chi-square or Fisher's exact tests were used for categorical variables.

Simple logistic regression (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24494/abstract) and multivariable logistic regression models with simple *P* values less than 0.2 were used to assess the relationship between hydroxychloroquine use and other covariates (independent variables) and incident atrial fibrillation or composite ventricular arrhythmias (dependent variable). The odds ratio (OR) and 95% confidence interval (95% CI) of incident atrial fibrillation for hydroxychloroquine users compared with nonusers was calculated, adjusting for covariates. Additionally, to control for confounding by indication, a propensity score analysis (1:1 greedy match) for the probability of each patient receiving hydroxychloroquine was calculated using a logistic regression model; adjusting for age, sex, body mass index, ethnicity, and smoking status was completed as an alternative approach. A box plot was used to investigate differences in the propensity score between hydroxychloroquine users and nonusers, and if balancing was achieved, then distribution was considered to be similar. Subsequently, unadjusted and fully adjusted (age, sex, race, alcohol, any antiarrhythmic medication use, hypertension, coronary artery disease, heart failure, cerebrovascular accident/transient ischemic attack, diabetes mellitus, obstructive sleep apnea, peripheral vascular disease, and chronic kidney disease) logistic regression models, including the propensity score matched data, were used to assess the relationship between hydroxychloroquine use and other covariates (independent variables) and incident atrial

fibrillation or composite ventricular arrhythmias (dependent variables). All analyses were performed using the Alteryx Designer x64 2018.3.4 (Alteryx) and SAS Enterprise Guide 7.15 HF3.

RESULTS

A total of 1,647 patients with SLE met the inclusion criteria for this study, of which 917 were hydroxychloroquine users and 730 were nonusers. Overall, the cohort was predominantly female (92%) and White (84%), with a mean ± SD age of 53.6 ± 14 years. The median (interquartile range) dose of hydroxychloroquine was 400 mg/day (300–400) (n = 859 of 917).

Patient demographic and clinical characteristics for the hydroxychloroquine users and nonusers are detailed in Table 1. Patients in the nonuser group were older and more likely to have hypertension and coronary artery disease as compared to hydroxychloroquine users. Prescriptions claims data for immunosuppressive and glucocorticoid medication use were available for 865 patients. Of the 23 incident atrial fibrillation events that occurred during the observation period, 3 were in the hydroxychloroquine user group, compared to 20 events in the nonuser group. All 3 atrial fibrillation events in the hydroxychloroquine user group were paroxysmal atrial fibrillation, whereas there were 14 paroxysmal, 1 persistent, 1 permanent, and 4 unspecified atrial fibrillation events identified in the nonuser group based on ICD coding. The median duration of follow-up in hydroxychloroquine users and nonusers with incident atrial fibrillation was 566 days (529–624 days) and 830 days (533–943 days), respectively ($P = 0.32$). The mean duration of hydroxychloroquine use prior to incident atrial fibrillation in the 3 hydroxychloroquine users was 420 days, and the median dose at the time of incident atrial fibrillation was 400 mg/day (individual doses of 200 mg/day, 400 mg/day, and 400 mg/day, and cumulative doses of 84 grams, 177 grams, and 159 grams, respectively, in the 3 patients). Based on manual review of echocardiograms of the patients with incident atrial fibrillation, the median left atrium size in the hydroxychloroquine users and nonusers was 4.8 cm (4.6–4.9) and 4.5 cm (4.3–4.7), respectively ($P = 0.37$). Additionally, EKG changes in the hydroxychloroquine users and nonusers included (in baseline EKGs during study period) median P wave duration of 50 msec (40–60) and 60 msec (60–80), respectively ($P = 0.38$), and median PR interval of 160 msec (154–204) and 156 msec (132–180), respectively ($P = 0.55$).

Four incident ventricular arrhythmias occurred during the study period, of which 1 occurred in the nonusers (VF) and 3 in

Table 1. Baseline characteristics according to hydroxychloroquine (HCQ) use*

Variable	HCQ users (n = 917)	HCQ nonusers (n = 730)	*P*
Demographic characteristic			
Age, years mean ± SD	51.8 ± 13.8	55.9 ± 14.6	<0.0001
Women	851 (93)	672 (92)	0.57
White	765 (84)	596 (84)	0.86
Atrial fibrillation risk factors and comorbidities			
Age ≥65 years	164 (18)	194 (27)	<0.0001
Median (IQR) body mass index, kg/m^2	27.7 (23–33)	27.9 (23–34)	0.19
Smoking, ever	366 (40)	325 (45)	0.06
Alcohol use	381 (42)	260 (36)	0.014
Hypertension	252 (27)	253 (35)	0.0017
Chronic obstructive pulmonary disease	0 (0)	0 (0)	–
Obstructive sleep apnea	48 (5.2)	34 (4.7)	0.59
Coronary artery disease	37 (4.0)	61 (8.4)	0.0002
Heart failure	27 (2.9)	36 (4.9)	0.037
Diabetes mellitus	68 (7.4)	89 (12)	0.001
CVA/TIA	10 (1.1)	9 (1.2)	0.79
Chronic kidney disease	38 (4.1)	42 (5.8)	0.13
End-stage renal disease	11(1.2)	11(1.5)	0.59
Peripheral vascular disease	9 (0.98)	19 (2.6)	0.012
Thyroid disease	4 (0.44)	4 (0.55)	0.74
Medications			
Any antiarrhythmic use	92 (10)	98 (13)	0.032
Beta blockers (metoprolol, propranolol, atenolol)	79 (8.6)	78 (11)	0.16
Calcium channel blockers (diltiazem, verapamil)	16 (1.7)	20 (2.7)	0.17
Other antiarrhythmics (amiodarone, digoxin, flecainide)	6 (0.65)	11 (1.5)	0.089
Immunosuppressives (methotrexate, azathioprine, belimumab, and/or cyclophosphamide), no./total no. (%)†	92/520 (18)	42/345 (12)	0.028
Glucocorticoids, no./total no. (%)	286/520 (55)	186/345 (54)	0.75

* Values are the number (%) unless indicated otherwise. CVA/TIA = cerebrovascular accident or transient ischemic attack; IQR = interquartile range.
† For methotrexate: n = 59 of 520 HCQ users, n = 31 of 345 HCQ nonusers; azathioprine: n = 44 of 520 HCQ users, n = 14 of 345 HCQ nonusers; belimumab: n = 0; cyclophosphamide: n = 0.

the hydroxychloroquine users. The 3 events in the hydroxychloroquine users included 2 VT and 1 torsades de pointes, of which none were fatal. One of the 4 patients with ventricular arrhythmia had a positive SSA antibody test result (hydroxychloroquine user with VT).

In simple logistic regression analysis, hydroxychloroquine users had a protective effect against atrial fibrillation (OR 0.117 [95% CI 0.034–0.39], $P = 0.0005$). Hydroxychloroquine use was not associated with ventricular arrhythmias (OR 2.39 [95% CI 0.25–23.0], $P = 0.45$). The protective relationship with incident atrial fibrillation was sustained in the fully adjusted logistic regression model adjusting for demographics (age, sex, race), comorbidities (alcohol, hypertension, diabetes mellitus, coronary artery disease, heart failure, cerebrovascular accident/transient ischemic attack, obstructive sleep apnea, peripheral vascular disease, chronic kidney disease), and medications (any antiarrhythmic use) (Table 2).

After the propensity score analysis, hydroxychloroquine remained protective against incident atrial fibrillation in a simple, logistic regression analysis (OR 0.154 [95% CI 0.045–0.52], $P = 0.003$). Additionally, hydroxychloroquine remained protective against incident atrial fibrillation in the multivariable, logistic regression (OR 0.151 [95% CI 0.04–0.63], $P = 0.009$).

DISCUSSION

Atrial fibrillation is the most common cardiac arrhythmia in the general population, and patients with SLE are at higher risk (2–4). Hydroxychloroquine is an integral component of care for cutaneous and musculoskeletal manifestations of SLE. The antiarrhythmic effects and impact on heart rate of hydroxychloroquine and its parent molecule chloroquine were documented in small studies in the late 1950s (5). From a molecular structure standpoint, hydroxychloroquine and chloroquine are derivatives of quinine and its stereoisomer quinidine, a class Ia antiarrhythmic agent. From a pathophysiology standpoint, more recently, several in vitro and animal model studies have illustrated hydroxychloroquine and chloroquine's inward-rectifier K+ channel blockade profile as the primary mechanism for their effectiveness in atrial fibrillation (7,9,10). This is also the mechanism of action of several Food and Drug Administration–approved antiarrhythmic agents. In addition, in patients with SLE, hydroxychloroquine use has been associated with a reduced risk of arterial thrombosis, the most important cause of cardioembolic stroke in atrial fibrillation (11,12).

In our cohort of patients with SLE, hydroxychloroquine use was independently associated with a protective effect (88% decreased risk) against incident atrial fibrillation events. To our knowledge, this is the first study to report such an association.

In our cohort, hydroxychloroquine nonusers were found to be older with higher frequency of hypertension and coronary artery disease, all of which are independent risk factors for atrial fibrillation. These variables were adjusted in the multivariable regression model. In our univariable analysis (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24494/abstract), we found a reverse association between alcohol use and atrial fibrillation, which is at odds with published data that suggest higher risk with heavy alcohol consumption. This could be an incidental finding as this was not a study aim, and we did not have data on quantity of alcohol consumption.

The 3 ventricular arrhythmia events in the hydroxychloroquine users were 2 VTs and 1 torsades de pointes, and none were fatal. We did not include sudden death as part of our secondary outcome, as sudden death could occur as a result of nonventricular arrhythmia reasons as well, and none of the 4 ventricular arrhythmias in our study were fatal events. Few published case reports have reported the occurrence of torsades de pointes with hydroxychloroquine use, and few studies have noted QT prolongation to be more prevalent among SSA-positive patients with SLE (13–15). Due to the small number of ventricular arrhythmia events in our cohort, we consider analysis of our secondary outcome to be limited. This area remains to be studied in larger cohorts.

Our study has several methodologic advantages, including exclusion of patients with history of atrial fibrillation by manual EHR adjudication, utilization of a 1-year run-in period, manual EHR plus

Table 2. Logistic regression analyses of incident atrial fibrillation and composite ventricular arrhythmias among hydroxychloroquine users versus nonusers*

Variable	Atrial fibrillation	*P*	Composite ventricular arrhythmias	*P*
Hydroxychloroquine				
Unadjusted	0.117 (0.034–0.39)	0.0005	2.39 (0.25–23)	0.45
Fully adjusted multivariable logistic regression†	0.137 (0.033–0.56)	0.006	4.38 (0.35–55)	0.25
Unadjusted conditional logistic regression with propensity score	0.154 (0.045–0.52)	0.003	3.01 (0.31–29)	0.34
Fully adjusted multivariable conditional logistic regression model with propensity score†	0.151 (0.04–0.63)	0.009	4.86 (0.41–58)	0.21

* Values are the odds ratio (95% confidence interval) unless indicated otherwise.
† Multivariable logistic regression included age, sex, race, alcohol use, any antiarrhythmic use, hypertension, diabetes mellitus, coronary artery disease, heart failure, cerebrovascular accident/transient ischemic attack, obstructive sleep apnea, peripheral vascular disease, and chronic kidney disease.

EKG confirmation to further enhance validity of incident atrial fibrillation, and availability of medication use data.

Limitations of our study include lack of data on SLE disease activity indices, incomplete data on duration and cumulative exposure of hydroxychloroquine, and lack of data on duration of SLE. Without SLE disease activity data, we were not able to evaluate hydroxychloroquine's antifibrillatory effect mediated specifically by better disease control. We also had incomplete prescription claims data on use of glucocorticoids and immunosuppressive medications, which could also impact cardiovascular risk. By nature of being observational data, there is potential for confounding by indication, which we attempted to minimize by using propensity scores for the probability of hydroxychloroquine use.

In conclusion, in our hypothesis-generating study, hydroxychloroquine was associated with a protective effect against incident atrial fibrillation in patients with lupus. Given the relative safety and low cost of hydroxychloroquine, and its favorable antithrombotic and cardiovascular risk benefit, a broader investigation of hydroxychloroquine in other cohorts or randomized studies to confirm its antifibrillatory effect would be warranted.

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. Sharma 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. Gupta, Manzi, Wasko, Sharma.
Acquisition of data. Gupta, Shields, Sharma.
Analysis and interpretation of data. Gupta, Shields, Sharma.

REFERENCES

1. Miner JJ, Kim AH. Cardiac manifestations of systemic lupus erythematosus. Rheum Dis Clin North Am 2014;40:51–60.
2. Teixeira RA, Borba EF, Pedrosa A, Nishioka S, Viana VS, Ramires JA, et al. Evidence for cardiac safety and antiarrhythmic potential of chloroquine in systemic lupus erythematosus. Europace 2014;16:887–92.
3. Barnado A, Carroll RJ, Casey C, Wheless L, Denny JC, Crofford LJ. Phenome-wide association studies uncover a novel association of increased atrial fibrillation in male patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70:1630–6.
4. Al Rayes H, Harvey PJ, Gladman DD, Su J, Sabapathy A, Urowitz MB, et al. Prevalence and associated factors of resting electrocardiogram abnormalities among systemic lupus erythematosus patients without cardiovascular disease. Arthritis Res Ther 2017;19:31.
5. Burrell ZL Jr, Martinez AC. Chloroquine and hydroxychloroquine in the treatment of cardiac arrhythmias. N Engl J Med 1958;258:798–800.
6. Allessie MA, Boyden PA, Camm AJ, Kléber AG, Lab MJ, Legato MJ, et al. Pathophysiology and prevention of atrial fibrillation. Circulation 2001;103:769–77.
7. Capel RA, Herring N, Kalla M, Yavari A, Mirams GR, Douglas G, et al. Hydroxychloroquine reduces heart rate by modulating the hyperpolarization-activated current I_f: novel electrophysiological insights and therapeutic potential. Heart Rhythm 2015;12:2186–94.
8. Sharma TS, Wasko MC, Tang X, Vedamurthy D, Yan X, Cote J, et al. Hydroxychloroquine use is associated with decreased incident cardiovascular events in rheumatoid arthritis patients. J Am Heart Assoc 2016;5:e002867.
9. Noujaim SF, Stuckey JA, Ponce-Balbuena D, Ferrer-Villada T, López-Izquierdo A, Pandit SV, et al. Structural bases for the different antifibrillatory effects of chloroquine and quinidine. Cardiovasc Res 2011;89:862–9.
10. Filgueiras-Rama D, Martins RP, Mironov S, Yamazaki M, Calvo CJ, Ennis SR, et al. Chloroquine terminates stretch-induced atrial fibrillation more effectively than flecainide in the sheep heart. Circ Arrhythm Electrophysiol 2012;5:561–70.
11. Jung H, Bobba R, Su J, Shariati-Sarabi Z, Gladman DD, Urowitz M, et al. The protective effect of antimalarial drugs on thrombovascular events in systemic lupus erythematosus. Arthritis Rheum 2010;62:863–8.
12. Petri M. Use of hydroxychloroquine to prevent thrombosis in systemic lupus erythematosus and in antiphospholipid antibody-positive patients. Curr Rheumatol Rep 2011;13:77–80.
13. Chen CY, Wang FL, Lin CC. Chronic hydroxychloroquine use associated with QT prolongation and refractory ventricular arrhythmia. Clin Toxicol (Phila) 2006;44:173–5.
14. O'Laughlin JP, Mehta PH, Wong BC. Life threatening severe QTc prolongation in patient with systemic lupus erythematosus due to hydroxychloroquine. Case Rep Cardiol 2016;2016:4626279.
15. Lazzerini PE, Capecchi PL, Acampa M, Morozzi G, Bellisai F, Bacarelli MR, et al. Anti-Ro/SSA-associated corrected QT interval prolongation in adults: the role of antibody level and specificity. Arthritis Care Res (Hoboken) 2011;63:1463–70.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 833–840
DOI 10.1002/acr.24188

American College of Rheumatology
Empowering Rheumatology Professionals

Association of Child Abuse and Systemic Lupus Erythematosus in Black Women During Adulthood

Yvette C. Cozier,[1] Medha Barbhaiya,[2] Nelsy Castro-Webb,[1] Carolyn Conte,[1] Sara Tedeschi,[3] Cianna Leatherwood,[3] Karen H. Costenbader,[3] and Lynn Rosenberg[1]

Objective. Exposure to psychosocial stressors may contribute to the onset of systemic lupus erythematosus (SLE) through dysregulation of the adaptive stress response. The present study was undertaken to assess the relationship of childhood physical and sexual abuse to risk of SLE among Black women.

Methods. Using data from the Black Women's Health Study, we followed 36,152 women from 1995 through 2015 with biennial questionnaires. Women reported on exposure to abuse during childhood (up to age 11) in 2005. Self-reported cases of incident SLE were confirmed as meeting the American College of Rheumatology SLE classification criteria by medical record review. Cox proportional hazards regression models were used to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs) for SLE among women exposed to physical or sexual abuse during childhood, controlling for potential confounders.

Results. We confirmed 101 cases of incident SLE and identified patients who had completed questions on child abuse during 670,822 person-years of follow-up. Both physical and sexual abuse during childhood were associated with statistically significant increases in SLE incidence. The HR for SLE associated with ≥2 episodes of severe sexual abuse compared to no abuse was 2.51 (95% CI 1.29–4.85) after adjustment for alcohol consumption, smoking, body mass index, oral contraceptive use, age at menarche, and parental education. The multivariable-adjusted HR for SLE with ≥5 episodes of severe physical abuse was 2.37 (95% CI 1.13–4.99).

Conclusion. Our results suggest that sexual and physical abuse during childhood increase SLE risk during adulthood among Black women. Research is necessary both to confirm this finding and to understand potential mediating mechanisms.

INTRODUCTION

Systemic lupus erythematosus (SLE) is an inflammatory autoimmune disease that affects Black individuals more frequently than White individuals, often with more severe manifestations and younger onset in Black patients (1). Environmental factors have been suggested as playing an important role in the pathogenesis of SLE in genetically predisposed individuals and are being actively investigated (2). Exposure to psychosocial stressors, such as depression and posttraumatic stress disorder (PTSD), have been shown to contribute to the onset of SLE and other autoimmune diseases (3,4), particularly among genetically predisposed individuals (5), through the dysregulation of the adaptive stress response (6). Studies linking trauma to psychiatric disorders in adulthood (e.g., depression, PTSD, psychosis, and anxiety), diseases such as obesity and cardiovascular disease, and autoimmune diseases such as rheumatoid arthritis suggest an important role for inflammation (7–9).

Childhood physical and sexual abuse occur with alarming frequency, affecting an estimated 1 in 4 children in their lifetimes and 1 in 7 in the past year (10). In nationally representative studies, Black women consistently report higher rates of childhood abuse than White women (11,12), but these differences diminish with adjustment for socioeconomic status (13). Children who experience victimization show elevated levels of inflammatory biomarkers several decades later (6,14). Maltreated children in a New Zealand cohort had an elevated risk of clinically relevant C-reactive protein (CRP) and other inflammation biomarkers at

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grants R01-AR-0573727 and K24-AR-066109 and National Cancer Institute grants R01-CA-058420 and U01-CA-164974). Dr. Barbhaiya's work was supported by the Rheumatology Research Foundation (Investigator award). Dr. Tedeschi's work was supported by the Lupus Foundation of America (Career Development award).

[1]Yvette C. Cozier, DSc, MPH, Nelsy Castro-Webb, ALM, Carolyn Conte, MPH, Lynn Rosenberg, ScD: Boston University, Boston, Massachusetts; [2]Medha Barbhaiya, MD, MPH: Hospital for Special Surgery, New York, New York; [3]Sara Tedeschi, MD, MPH, Cianna Leatherwood, MD, Karen H. Costenbader, MD, MPH: Brigham and Women's Hospital, Boston, Massachusetts.

Drs. Cozier and Barbhaiya contributed equally to this work. Drs. Costenbader and Rosenberg contributed equally to this work.

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

Address correspondence to Yvette C. Cozier, DSc, MPH, Slone Epidemiology Center at Boston University, Boston, MA 02118. Email: yvettec@bu.edu.

Submitted for publication May 8, 2019; accepted in revised form March 10, 2020.

SIGNIFICANCE & INNOVATIONS

- Exposure to psychosocial stressors have been shown to contribute to the onset of systemic lupus erythematosus (SLE) and other autoimmune diseases decades after exposure.
- To our knowledge, this is the largest study to explore the role of childhood abuse victimization in relation to adult-onset SLE in Black women in the US.
- In this study of Black women in the US, physical and sexual abuse during childhood, in particular severe and frequent abuse, was associated with increased incidence of SLE during adulthood.
- Our study contributes to a growing body of evidence demonstrating an association between psychosocial factors and SLE; yet it must be considered exploratory as it was based on a relatively small number of confirmed cases.

age 32 (6), independent of other key risk factors in childhood and adulthood. In an analysis of >17,000 adult members of the San Diego Kaiser Foundation Health Plan, childhood traumatic stress (including physical, emotional, or sexual abuse) was associated with increased likelihood of hospitalization with an autoimmune condition in adulthood (14). A recent analysis of 67,500 female nurses in the US found a significantly increased risk of SLE among those who experienced childhood physical and emotional abuse (15). Adverse childhood experiences have also been associated with an increased risk of overall poor health (16) and smoking (17). Among Black women, studies have found associations with obesity (18), asthma (19), uterine fibroids (20), and breast cancer (21). Both smoking and obesity have been independently associated with an increased risk of SLE in studies of both White (22,23) and Black women (24–26).

Using data derived from the Black Women's Health Study (BWHS), a prospective cohort study, we investigated the relationship between physical and sexual abuse as a child and risk of SLE in adulthood in Black women in the US, a population at high risk of SLE. We hypothesized that early life abuse, in particular increased severity of abuse, would be associated with increased risk of SLE in adulthood.

MATERIALS AND METHODS

BWHS. In 1995, 64,500 Black women ages 21–69 years (median age ±38 years) from the continental US enrolled in the BWHS by completing a 14-page health questionnaire; the 59,000 women whose addresses were considered to be valid 1 year later comprise the BWHS cohort that has been followed. The questionnaires were mailed largely to subscribers of Essence magazine, who comprise the vast majority of participants. A small percentage of participants were members of several professional organizations and friends and relatives of early responders. Only women who self-identified as Black or African American were included. More than 80% of participants lived in California, Georgia, Illinois, Indiana, Louisiana, Maryland, Massachusetts, Michigan, New Jersey, New York, South Carolina, Virginia, and the District of Columbia. The participants provided demographic, medical, and lifestyle information at baseline in 1995 and have been followed since then with biennial health questionnaires and yearly linkage with the National Death Index. All but 3% of respondents had completed high school, and 44% had completed college; 95% of participants had been born in the US. Follow-up of the cohort has been successful for >85% of potential person-years through 2015. The Institutional Review Board of Boston University Medical Center approved the study, and participants indicated their consent by filling out the questionnaires and signing consent forms for obtainment of medical records.

Data collection. At baseline, participants provided data on demographics, current weight and height, weight at age 18 years, medical and reproductive history, vigorous physical activity, cigarette smoking, alcohol use, and other variables. Self-completed biennial follow-up questionnaires have been used to update various data items.

Abuse victimization. On the 2005 BWHS follow-up questionnaire, participants provided information about abuse victimization as a child (up to age 11 years) and as an adolescent (ages 12–18 years). We used a 9-item abuse questionnaire adapted from the Conflict Tactics Scale and the Pregnancy Abuse Assessment Screen (27,28). Response categories were "never," "1–3 times," and "≥4 times." We defined childhood physical abuse as a report of a perpetrator having "pushed, grabbed, or shoved me," "threw something at me that could hurt me," "kicked, bit, or punched me," "hit me with something including hand or fist," "physically attacked me in some other way," "choked or burned me," or "seriously harmed someone I loved" at a frequency of ≥4 times during childhood. We defined childhood sexual abuse as a report of a perpetrator having "exposed genitals against my will" or "been sexual with me against my will" at a frequency of ≥4 times. To create a childhood physical abuse summary score variable, we assigned 1 point for each report of a physical abuse item that occurred ≥4 times (severe abuse); to create a childhood sexual abuse summary category, we assigned 1 point for each report of sexual abuse that occurred ≥4 times (18,21,29–31). We also employed an alternate method for both physical and sexual abuse by assigning 1 point for each report of 1–3 episodes and 2 points for each report of ≥4 times and then summed the results. Previous analyses in the BWHS have utilized these approaches and have found associations between childhood abuse victimization and obesity (18), asthma (19), uterine fibroids (20), and breast cancer (21).

Covariates. We selected variables related to early childhood and adolescent experiences that might be associated with SLE. Data on these variables, including age, smoking, body mass index (BMI), alcohol consumption, oral contraceptive use, and age

at menarche, were obtained in 1995 and updated on subsequent questionnaires. Information on education of parents was obtained in 2009.

Cases of SLE. The 1995 questionnaire asked about a list of diagnoses that included lupus. Every biennial questionnaire thereafter asked about SLE and the date of diagnosis. The doctors of women who gave consent were asked for copies of medical records concerning SLE or to fill out a checklist of American College of Rheumatology (ACR) criteria met for the SLE diagnosis (32,33). As previously described (24,25), medical record review by study rheumatologists confirmed cases as fulfilling at least 4 ACR classification criteria for SLE. An earlier validation in the BWHS found that for the 251 women reporting incident or prevalent SLE for whom a physician checklist or medical chart was obtained, 84% of patients fulfilled ACR criteria for definite or probable SLE or had clinical lupus (i.e., an SLE diagnosis recorded in a medical chart plus appropriate medication use) (34).

Analytic cohort. The current analysis utilizes data from the baseline questionnaire and 10 subsequent follow-up cycles (1995–2015). The 2005 questionnaire containing the abuse questions was completed by 43,179 participants. We excluded 483 women who reported SLE prior to 1995 and 6,544 women with missing information on abuse. The remaining 36,152 women comprised the analytic cohort. The women in the analytic cohort were similar to those excluded in terms of the proportion of SLE cases. There were 13 SLE patients among the 6,544 women with missing abuse data (0.20%), while there were 101 SLE patients among the 36,152 women with abuse data (0.28%).

Statistical analysis. We used Cox proportional hazards regression to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs) for the association of physical and sexual abuse, separately and combined, during childhood and adolescence with risk of SLE. Women contributed person-time from baseline (1995) until SLE diagnosis, death, loss to follow-up, or end of follow-up (2015), whichever occurred first. Women who reported SLE but for whom we were unable to obtain medical records to confirm the diagnosis were censored in the year of diagnosis. The Cox models were jointly stratified by age in 1-year intervals and questionnaire cycle (model 1). Multivariable models for child abuse included age at menarche (<11, 11, 12–14, and ≥15 years), BMI at age 18 (<25, 25–29, ≥30), age that patient began smoking (≤14 years or >14 years/never), age that patient began alcohol consumption (≤14 years or >14 years/never), age of first use of oral contraceptives (≤14 years or >14 years/never), and parental education (neither parent attended college, at least 1 parent attended college, at least 1 parent was a college graduate).

Table 1. Age-standardized baseline characteristics by child abuse type among 36,152 women who reported abuse during childhood in the Black Women's Health Study*

Characteristic	No childhood abuse (n = 8,070)	Physical abuse only (n = 19,274)	Sexual abuse only (n = 809)	Sexual and physical abuse (n = 7,999)	Any abuse (n = 28,082)
Baseline (1995) factors					
Age, mean ± SD years	41.5 ± 11.4	39.0 ± 10.5	37.9 ± 10.3	37.5 ± 9.7	38.6 ± 10.3
BMI, mean ± SD kg/m^2	27.6 ± 6.4	27.7 ± 6.5	28.2 ± 6.7	28.6 ± 7.1	28.0 ± 6.7
Region of residence					
Northeast	28	27	29	29	27
South	32	31	29	27	30
Education, years					
≤12	21	16	17	15	16
≥16	45	51	52	48	50
Neighborhood SES					
Quintile 1 (low)	19	18	18	19	18
Quintile 5 (high)	18	19	21	18	19
Oral contraceptive use, ever	84	85	84	85	85
Cigarette smoking, ever	32	33	33	39	34
Alcohol intake, ever	40	43	40	45	43
Childhood factors					
Parental education, HS or less	45	45	46	46	46
Age at menarche ≤11 years	26	28	33	33	29
Oral contraceptive use ≤14 years	2	2	3	3	2
BMI at age 18 years, mean ± SD kg/m^2	21.3 ± 3.9	21.4 ± 4.0	21.7 ± 4.2	21.7 ± 4.3	21.5 ± 4.1
Passive smoking ages 0–10 years	42	49	45	52	50
Started alcohol intake ≤14 years	1	2	1	3	2
Started smoking ≤14 years	2	3	3	6	4

* Values are the percentage unless indicated otherwise. Values are standardized to the 1995 age distribution of the study population. Percentages may not add to 100 because of rounding errors or missing values. BMI = body mass index; HS = high school; SES = socioeconomic status.

RESULTS

A total of 101 patients with incident SLE who completed the child abuse questions were confirmed among 670,822 person-years between 1995, the start of follow-up for this analysis, and 2015, the last completed follow-up cycle. Among these confirmed cases of incident SLE, the mean age at diagnosis was 43 years, the mean ± SD number of ACR criteria for SLE diagnosis was 5.1 ± 1.4, 66% had a hematologic disorder, and 33% had a renal disorder. The majority of all exposures to childhood and adolescent abuse were reported to have started during childhood (83% of all participants returning the questionnaire reporting adolescent abuse also reported childhood abuse). As only 1 patient with SLE reported new-onset abuse during adolescence, we were unable to analyze abuse beginning in adolescence separately.

As shown in Table 1, physical abuse was reported more frequently than sexual abuse. Women who reported physical or sexual abuse were younger and heavier during adulthood, had an earlier age at menarche, were more likely to smoke or drink alcohol and to start at an earlier age, to begin oral contraceptive use at an earlier age, and to have a higher level of education than women who reported no abuse. Physical and sexual abuse were unrelated to region of the country and neighborhood socioeconomic status.

Table 2 provides data on childhood physical and sexual abuse in relation to SLE. HRs of physical and sexual abuse associated with SLE that were adjusted for age and questionnaire cycle only were similar to those adjusted for age at menarche, BMI at age 18 years, age that patient began alcohol consumption, age that patient started smoking, age that patient began use of oral contraceptives, and parental education. The multivariable HRs for every report of sexual abuse only and every report of both physical and sexual abuse exceeded 2.0, and the estimate for physical and sexual abuse was statistically significant (HR 2.20 [95% CI 1.14–4.21]). The multivariable HR for 2 reports of ≥4 episodes of sexual abuse was HR 2.51 (95% CI 1.29–4.85); the HR for ≥5 reports of ≥4 episodes of physical abuse was HR 2.37 (95% CI 1.13–4.99). Analyses using an alternate scoring method for both physical and sexual abuse yielded similar estimates.

Table 3 shows HRs for the individual questions that contributed to childhood physical and sexual abuse according to the distribution of questionnaire responses. The HRs exceeded 2.0 and were statistically significant for "choked or burned me," "attacked me in some other way," "exposed genitals against my will," and "[had] been sexual with me against my will." The number of cases in the highest response category was 3 for "choked or burned me" and ≥8 for the other questions.

We sought to address the possibility of recall bias by conducting an analysis restricted to incident cases occurring after 2005 when the abuse questions were asked. Based on 21 cases overall, the HR for the highest category of sexual abuse score was 2.88 (95% CI 0.84–9.89); the HR for the highest category of physical abuse score was 1.68 (95% CI 0.85–3.35) (data not shown).

Table 2. Hazard ratios (HRs) for systemic lupus erythematosus in relation to type and frequency of childhood physical and sexual abuse among 36,152 participants in the Black Women's Health Study followed from 1995 to 2015*

	No. of cases	Person-years	Age- and questionnaire period–adjusted	Fully adjusted†
Child abuse type‡				
None	58	413,173	Ref.	Ref.
Physical only	28	212,301	0.91 (0.58–1.43)	0.90 (0.57–1.42)
Sexual only	4	12,877	2.11 (0.76–5.83)	2.04 (0.74–5.66)
Physical and sexual	11	32,469	2.24 (1.17–4.27)	2.20 (1.14–4.21)
Sexual abuse score§				
0	86	625,475	Ref.	Ref.
1	5	18,956	1.85 (0.75–4.56)	1.84 (0.74–4.54)
2	10	26,390	2.57 (1.33–4.96)	2.51 (1.29–4.85)
Physical abuse score§				
0	62	426,049	Ref.	Ref.
1–2	20	163,969	0.81 (0.49–1.35)	0.81 (0.48–1.34)
3–4	11	59,963	1.20 (0.63–2.27)	1.19 (0.63–2.27)
5+	8	20,840	2.44 (1.16–5.10)	2.37 (1.13–4.99)

* Values are the HR (95% confidence interval) unless indicated otherwise. Ref. = reference.
† Adjusted for age that participants began alcohol consumption ≤14 years, age that participants began smoking ≤14 years, body mass index at age 18 years, age of first oral contraceptive use ≤14 years, parental education level, and age at menarche.
‡ A report of ≥4 instances of each type of abuse.
§ To create a childhood physical abuse summary score variable, we assigned 1 point for each report of a physical abuse item that occurred ≥4 times. To create a childhood sexual abuse summary category, we assigned 1 point for each report of sexual abuse that occurred ≥4 times (e.g., 0 = 0 reports of abuse occurring ≥4 times, 5 = 5 reports of abuse occurring ≥4 times).

Table 3. Hazard ratios (HRs) for systemic lupus erythematosus in relation to individual components of childhood physical and sexual abuse among 36,153 participants in the Black Women's Health Study, 1995–2005*

	No. of cases	Person-years	Age- and questionnaire period–adjusted	Fully adjusted†
Pushed, grabbed, or shoved me				
Never	43	263,447	Ref.	Ref.
1–3 times	32	259,416	0.74 (0.47–1.18)	0.74 (0.47–1.17)
≥4 times	26	147,958	1.03 (0.632–1.68)	1.02 (0.62–1.66)
Threw something at me that could hurt me				
Never	65	445,463	Ref.	Ref.
1–3 times	23	169,717	0.90 (0.56–1.45)	0.89 (0.55–1.43)
≥4 times	13	55,640	1.52 (0.84–2.76)	1.49 (0.82–2.72)
Kicked, bit, or punched me				
Never	54	393,854	Ref.	Ref.
1–3 times	30	203,082	1.03 (0.66–1.62)	1.02 (0.65–1.59)
≥4 times	17	73,884	1.56 (0.90–2.70)	1.53 (0.88–2.65)
Hit me with something including hand and fist				
Never	41	265,398	Ref.	Ref.
1–3 times	27	228,708	0.74 (0.45–1.20)	0.73 (0.45–1.19)
≥4 times	33	176,714	1.13 (0.72–1.80)	1.13 (0.71–1.89)
Choked or burned me				
Never	94	642,837	Ref.	Ref.
1–3 times	4	23,042	1.13 (0.42–3.09)	1.10 (0.40–3.01)
≥4 times	3	4,942	3.74 (1.18–11.83)‡	3.77 (1.19–11.96)‡
Physically attacked me in some other way				
Never	76	544,213	Ref.	Ref.
1–3 times	12	88,082	0.96 (0.52–1.78)	0.95 (0.52–1.75)
≥4 times	13	38,526	2.33 (1.29–4.20)‡	2.27 (1.26–4.11)‡
Exposed their genitals against my will				
Never	73	548,835	Ref.	Ref.
1–3 times	17	89,646	1.36 (0.81–2.32)	1.34 (0.79–2.28)
≥4 times	11	32,339	2.38 (1.26–4.51)‡	2.33 (1.23–4.41)‡
Was sexual with me against my will				
Never	77	551,029	Ref.	Ref.
1–3 times	10	80,396	0.85 (0.44–1.65)	0.84 (0.43–1.62)
≥4 times	14	39,396	2.38 (1.34–4.23)‡	2.33 (1.31–4.15)‡
Seriously harmed someone I loved				
Never	81	569,728	Ref.	Ref.
1–3 times	12	65,883	1.25 (0.68–2.30)	1.23 (0.67–2.25)
≥4 times	8	35,209	1.53 (0.74–3.17)	1.50 (0.72–3.12)

* Values are the HR (95% confidence interval) unless indicated otherwise. HRs are presented according to the distribution of responses to the individual components of abuse. Ref. = reference.
† Adjusting for age, period, age that participant began alcohol consumption ≤14 years, age that participant began smoking ≤14 years, body mass index at age 18 years, age of first oral contraceptive use ≤14 years, parental education level, and age at menarche.
‡ Significant.

DISCUSSION

In the current study, abuse during childhood was associated with increased incidence of SLE during adulthood. The increase was ~2.5 fold for both physical and sexual abuse. Physical abuse was reported more frequently than sexual abuse, and the number of episodes associated with this increased risk was greater for physical abuse (at least 5 reports of physical abuse occurring at least 4 times) than for sexual abuse (2 reports of sexual abuse occurring at least 4 times). The numbers of exposed cases were insufficient to adequately study physical abuse in the absence of sexual abuse, or sexual abuse in the absence of physical abuse. The actions most strongly associated with increased SLE were "choked or burned me," "attacked me in some other way," "exposed genitals against my will," and "[had] been sexual with me against my will," but the numbers of exposed cases were small, and thus the estimates had wide confidence intervals.

To our knowledge, this is the largest study of abuse victimization in relation to SLE in Black women in the US. Among 269 prevalent cases of SLE identified in the San Francisco area of California, sexual abuse was reported more frequently by case than in a comparable sample of patients from the Behavioral Risk Factor Surveillance System, of which only 12% were Black (35). Perhaps the most relevant studies of other exposures are those of PTSD in relation to SLE. A study of Iraq and Afghanistan veterans found a higher absolute prevalence of PTSD (5.4% women, 1.7% men)

among those with autoimmune diseases. Veterans of both sexes diagnosed with PTSD were at significantly higher risk of diagnosis with any autoimmune condition, alone or combined, including SLE (36). In a study of predominantly White civilian female nurses, PTSD symptoms were associated with a >2-fold increased risk of incident SLE among women who experienced any traumatic event compared with those unexposed to trauma (4). The PTSD exposures studied in both the veterans and nurses were mainly adult exposures, whereas we looked at abuse during childhood. Another analysis of data from the Nurses' Health Study II by Feldman et al (15) assessed physical and emotional abuse during childhood and controlled for similar covariates (e.g., parental education and age at menarche) as in our analysis. Similar to us, they found a 2.57-times greater risk of SLE (95% CI 1.30–5.12) related to high levels of childhood exposure relative to low levels. They additionally found the association to be partially mediated by adult depression and PTSD.

A number of potential mechanisms may explain the association observed between childhood physical and sexual abuse and incident SLE. Animal models demonstrate an important link between PTSD and increased systemic inflammation via upregulation of microRNA in the brain, adrenal glands, and blood and higher circulating IgM levels (37,38). A meta-analysis of 25 studies demonstrated the association between childhood trauma and elevated levels of inflammatory biomarkers such as CRP, interleukin-6, and tumor necrosis factor at a mean age of ±42 years (39); in fact, subgroup analyses for specific types of trauma (physical, sexual, or emotional) revealed a differential impact on inflammatory markers by trauma type. While a number of studies have demonstrated that childhood trauma may be associated with high inflammation levels decades after exposure (6), a recent study demonstrated that childhood victimization predicted elevated CRP levels by age 18 years in young women independent of genetic and socioeconomic risk of inflammation (40). Additionally, dysregulation of the hypothalamic–pituitary–adrenal axis has been implicated as a modulator of inflammatory activity potentially leading to activation of the immune system (41).

The current study has several strengths. The data were collected using a prospective cohort study design with a lengthy period of follow-up. Validated data collection tools were used to assess childhood physical and sexual abuse (21,28,42); these instruments have been widely used and demonstrate high reproducibility within the BWHS (29) and in other studies (28,43). Furthermore, factor analysis of BWHS data indicate that the abuse questions identified the underlying constructs that they were intended to measure (29). Although self-report of child abuse as an adult may lead to underreporting of abuse, and thus potentially underestimating the association between child abuse and SLE in the current study, self-report measures of child abuse have strong discriminant validity for identifying those with a history of abuse (44). Potential cases of SLE were reviewed and confirmed by study rheumatologists as satisfying accepted classification criteria for SLE. Potential childhood and parental confounding factors were controlled in the analyses. The questions from the Conflict Tactics Scale have been associated with other outcomes in the BWHS, including age at menarche and obesity (18), factors also known to be associated with SLE. Additionally, the BWHS has similar prevalence estimates of childhood abuse compared to those found in nationally representative studies (10–12,16,45,46), suggesting that these findings may be generalizable to a broader population of Black women in the US.

Study limitations include the cross-sectional nature of data collection concerning abuse. Women were followed for SLE incidence from 1995 to 2015, but experiences of abuse during childhood were ascertained in 2005. Although experiences of physical and sexual abuse victimization in childhood would have preceded the occurrence of SLE in the BWHS, the temporal sequence of reporting abuse experiences did not precede the diagnosis for most cases in our analysis. Thus, recall bias could have occurred if women who had SLE overreported abuse or were more likely to remember childhood abuse compared to women who did not have SLE. Additionally, only those patients with SLE who survived until at least 2005 had the opportunity to report their experiences of abuse. Therefore, the cases analyzed from 1995 to 2005 may underrepresent the most aggressive cases of SLE in the cohort. In addition, since there was only 1 case of abuse that began when the patient was in adolescence, it was not possible to assess child and adolescent abuse separately. We conducted multiple testing (e.g., of the individual questions), which increased the possibility of false positives. The positive findings in our study need to be independently confirmed in other data. Although we did account for potentially important childhood and adult confounders, we did not perform mediation analyses to assess whether covariates associated with child abuse and SLE (such as cigarette smoking, alcohol consumption, obesity, reproductive factors, or depression) may actually lie on the causal pathway. Additionally, our study did not control for other childhood stressors that may be associated with childhood victimization (47). Finally, we were unable to assess early-onset SLE (in adolescence), which may be particularly related to childhood exposures.

In conclusion, this study suggests that childhood physical and sexual abuse, in particular severe and frequent abuse, are associated with increased risk of developing SLE among adult Black women. Our study contributes to a growing body of evidence demonstrating an association between psychosocial factors and SLE. However, the study must be considered exploratory, as it is the first to assess abuse in childhood and adolescence in relation to SLE, and it was based on small numbers of exposed cases. Confirmation in other data are required, and identification of biologic pathways could provide insight into disease etiology.

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. Cozier 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. Cozier, Barbhaiya, Costenbader, Rosenberg.

Acquisition of data. Cozier, Barbhaiya, Conte, Costenbader, Rosenberg.

Analysis and interpretation of data. Cozier, Barbhaiya, Castro-Webb, Tedeschi, Leatherwood, Costenbader, Rosenberg.

REFERENCES

1. 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.
2. Mak A, Tay SH. Environmental factors, toxicants and systemic lupus erythematosus. Int J Mol Sci 2014;15:16043–56.
3. De Brouwer SJ, Kraaimaat FW, Sweep FC, Creemers MC, Radstake TR, van Laarhoven AI, et al. Experimental stress in inflammatory rheumatic diseases: a review of psychophysiological stress responses. Arthritis Res Ther 2010;12:R89.
4. Roberts AL, Malspeis S, Kubzansky LD, Feldman CH, Chang SC, Koenen KC, et al. Association of trauma and posttraumatic stress disorder with incident systemic lupus erythematosus in a longitudinal cohort of women. Arthritis Rheumatol 2017;69:2162–9.
5. Luiz AP, Antico HA, Skare TL, Boldt AB, Nisihara R. Adverse childhood experience and rheumatic diseases. Clin Rheumatol 2018; 37:2863–7.
6. Danese A, Pariante CM, Caspi A, Taylor A, Poulton R. Childhood maltreatment predicts adult inflammation in a life-course study. Proc Natl Acad Sci U S A 2007;104:1319–24.
7. Miller G, Chen E, Cole SW. Health psychology: developing biologically plausible models linking the social world and physical health. Annu Rev Psychol 2009;60:501–24.
8. Danese A, Baldwin JR. Hidden wounds? Inflammatory links between childhood trauma and psychopathology. Annu Rev Psychol 2017;68:517–44.
9. Danese A, McEwen BS. Adverse childhood experiences, allostasis, allostatic load, and age-related disease. Physiol Behav 2012; 106:29–39.
10. Centers for Disease Control. National Center for Injury Prevention and Control. Child maltreatment: facts at a glance, 2014. 2016. URL: https://www.cdc.gov/violenceprevention/pdf/childmaltreatment-facts-at-a-glance.pdf.
11. Finkelhor D. Current information on the scope and nature of child sexual abuse. Future Child 1994;4:31–53.
12. Barnett OW, Miller-Perrin CL, Perrin RD. Physical child abuse. In: Barnett O, Miller-Perrin CL, Perrin RD, editors. Family violence across the lifespan. Thousand Oaks (CA): SAGE; 1996.
13. Rennison C, Planty M. Nonlethal intimate partner violence: examining race, gender, and income patterns. Violence Vict 2003;18:433–43.
14. Dube SR, Fairweather D, Pearson WS, Felitti VJ, Anda RF, Croft JB. Cumulative childhood stress and autoimmune diseases in adults. Psychosom Med 2009;71:243–50.
15. Feldman CH, Malspeis S, Leatherwood C, Kubzansky L, Costenbader KH, Roberts AL. Association of childhood abuse with incident systemic lupus erythematosus in adulthood in a longitudinal cohort of women. J Rheumatol 2019;46:1589–96.
16. Felitti VJ, Anda RF, Nordenberg D, Williamson DF, Spitz AM, Edwards V, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: the Adverse Childhood Experiences (ACE) Study. Am J Prev Med 1998; 14:245–58.
17. Nichols HB, Harlow BL. Childhood abuse and risk of smoking onset. J Epidemiol Community Health 2004;58:402–6.
18. Boynton-Jarrett R, Rosenberg L, Palmer JR, Boggs DA, Wise LA. Child and adolescent abuse in relation to obesity in adulthood: the Black Women's Health Study. Pediatrics 2012;130:245–53.
19. Coogan PF, Wise LA, O'Connor GT, Brown TA, Palmer JR, Rosenberg L. Abuse during childhood and adolescence and risk of adult-onset asthma in African American women. J Allergy Clin Immunol 2013;131:1058–63.
20. Wise LA, Palmer JR, Rosenberg L. Lifetime abuse victimization and risk of uterine leiomyomata in black women. Am J Obstet Gynecol 2013;208:272 e1–13.
21. Wise LA, Palmer JR, Boggs DA, Adams-Campbell LL, Rosenberg L. Abuse victimization and risk of breast cancer in the Black Women's Health Study [corrected]. Cancer Causes Control 2011; 22:659–69.
22. Barbhaiya M, Tedeschi SK, Lu B, Malspeis S, Kreps D, Sparks JA, et al. Cigarette smoking and the risk of systemic lupus erythematosus, overall and by anti-double stranded DNA antibody subtype, in the Nurses' Health Study cohorts. Ann Rheum Dis 2018;77: 196–202.
23. Costenbader KH, Kim DJ, Peerzada J, Lockman S, Nobles-Knight D, Petri M, et al. Cigarette smoking and the risk of systemic lupus erythematosus: a meta-analysis. Arthritis Rheum 2004;50:849–57.
24. Cozier YC, Barbhaiya M, Castro-Webb N, Conte C, Tedeschi SK, Leatherwood C, et al. Relationship of cigarette smoking and alcohol consumption to incidence of systemic lupus erythematosus in prospective cohort study of Black women. Arthritis Care Res (Hoboken) 2019;71:671–7.
25. Cozier YC, Barbhaiya M, Castro-Webb N, Conte C, Tedeschi S, Leatherwood C, et al. A prospective study of obesity and risk of systemic lupus erythematosus (SLE) among Black women. Semin Arthritis Rheum 2019;48:1030–4.
26. Tedeschi SK, Barbhaiya M, Malspeis S, Lu B, Sparks JA, Karlson EW, et al. Obesity and the risk of systemic lupus erythematosus among women in the Nurses' Health Studies. Semin Arthritis Rheum 2017;47:376–83.
27. Straus MA. Measuring intrafamily conflict and violence: the Conflict Tactics (CT) scales. J Marriage Fam 1979;41:75–88.
28. McFarlane J, Parker B, Soeken K, Bullock L. Assessing for abuse during pregnancy: severity and frequency of injuries and associated entry into prenatal care. JAMA 1992;267:3176–8.
29. Wise LA, Palmer JR, Rothman EF, Rosenberg L. Childhood abuse and early menarche: findings from the Black Women's Health Study. Am J Public Health 2009;99 Suppl S:460S–6S.
30. Wise LA, Zierler S, Krieger N, Harlow BL. Adult onset of major depressive disorder in relation to early life violent victimisation: a case-control study. Lancet 2001;358:881–7.
31. Rayworth BB, Wise LA, Harlow BL. Childhood abuse and risk of eating disorders in women. Epidemiology 2004;15:271–8.
32. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725–34.
33. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.
34. McAlindon TE, Formica M, Palmer JR, Lafyatis R, Rosenberg L. Assessment of strategies for identifying diagnosed cases of systemic lupus erythematosus through self-report. Lupus 2003;12:754–9.
35. DeQuattro K, Trupin L, Li J, Katz PP, Murphy LB, Yelin EH, et al. Relationships between adverse childhood experiences and health status in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2020;72:525–33.

36. O'Donovan A, Cohen BE, Seal KH, Bertenthal D, Margaretten M, Nishimi K, et al. Elevated risk for autoimmune disorders in Iraq and Afghanistan veterans with posttraumatic stress disorder. Biol Psychiatry 2015;77:365–74.
37. Wilson CB, McLaughlin LD, Nair A, Ebenezer PJ, Dange R, Francis J. Inflammation and oxidative stress are elevated in the brain, blood, and adrenal glands during the progression of post-traumatic stress disorder in a predator exposure animal model. PLoS One 2013;8:e76146.
38. Boscarino JA. Posttraumatic stress disorder and physical illness: results from clinical and epidemiologic studies. Ann N Y Acad Sci 2004;1032:141–53.
39. Baumeister D, Akhtar R, Ciufolini S, Pariante CM, Mondelli V. Childhood trauma and adulthood inflammation: a meta-analysis of peripheral C-reactive protein, interleukin-6 and tumour necrosis factor-alpha. Mol Psychiatry 2016;21:642–9.
40. Baldwin JR, Arseneault L, Caspi A, Fisher HL, Moffitt TE, Odgers CL, et al. Childhood victimization and inflammation in young adulthood: a genetically sensitive cohort study. Brain Behav Immun 2018; 67:211–7.
41. Santa Ana EJ, Saladin ME, Back SE, Waldrop AE, Spratt EG, McRae AL, et al. PTSD and the HPA axis: differences in response to the cold pressor task among individuals with child vs. adult trauma. Psychoneuroendocrinology 2006;31:501–9.
42. Straus MA. Measuring intrafamily conflict and violence: the Conflict Tactics (CT) scales. J Marriage Fam 1979;41:75–88.
43. Straus MA. The Conflict Tactics Scales and its critics: an evaluation and new data on validity and reliability. In: Straus MA, Gelles RJ, editors. Physical violence in American families: risk factors and adaptations to violence in 8,145 families. New Brunswick (NJ): Transaction Publishers; 1990.
44. Widom CS, Shepard RL. Accuracy of adult recollections of childhood victimization. Part I. Childhood physical abuse. Psychol Assess 1996;8:412–21.
45. Kilpatrick DG, Acierno R, Saunders B, Resnick HS, Best CL, Schnurr PP. Risk factors for adolescent substance abuse and dependence: data from a national sample. J Consult Clin Psychol 2000;68:19–30.
46. Finkelhor D, Ormrod RK, Turner HA. Re-victimization patterns in a national longitudinal sample of children and youth. Child Abuse Negl 2007;31:479–502.
47. Teegen F. Childhood sexual abuse and long term sequelae. Seattle (WA): Hogrefe and Huber Publishers; 1999.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 841–848
DOI 10.1002/acr.24206

American College of Rheumatology
Empowering Rheumatology Professionals

Variability in Interpretation of Magnetic Resonance Imaging of the Pediatric Sacroiliac Joint

Pamela F. Weiss,[1] Timothy G. Brandon,[2] John Bohnsack,[3] Merav Heshin-Bekenstein,[4] Michael L. Francavilla,[1] Jacob L. Jaremko,[5] Lester Liao,[5] Anne McHugh,[6] Edward J. Oberle,[7] Dax Rumsey,[5] Hemalatha Srinivasalu,[8] Matthew L. Stoll,[9] and Nancy A. Chauvin[10]

Objective. Magnetic resonance imaging (MRI) is pivotal in the assessment of early sacroiliitis in children. We aimed to evaluate the agreement between local radiology reports and central imaging reviewers for active inflammation and structural damage at the sacroiliac (SI) joints.

Methods. Eight hospitals each contributed up to 20 cases of consecutively imaged children and adolescents with juvenile idiopathic arthritis and suspected sacroiliitis. Studies were independently reviewed by 3 experienced musculoskeletal pediatric radiologists. Local assessments of global impression and lesions were coded from the local radiology reports by 2 study team members. Test properties of local reports were calculated using the central imaging team's majority as the reference standard.

Results. For 120 evaluable subjects, the median age was 14 years, half of the cases were male, and median disease duration at the time of imaging was 0.8 years (interquartile range 0–2). Sensitivity of local reports for inflammation was high, 93.5% (95% confidence interval [95% CI] 78.6–99.2), and specificity was moderate, 69.7% (95% CI 59.0–79.0), but positive predictive value (PPV) was low, 51.8% (95% CI 38.0–65.3). Twenty-seven cases (23%) had active inflammation reported locally but rated normal at the central reading, 19 (70%) with subsequent medication changes. The sensitivity of local reports detecting structural damage was low, 45.7% (95% CI 28.8–63.4), and specificity was high, 88.2% (95% CI 79.4–94.2); PPV was low, 61.5% (95% CI 40.6–79.8).

Conclusion. Substantial variation exists in the interpretation of inflammatory and structural lesions at the SI joints in children. To reliably identify pathology, additional training in the MRI appearance of the maturing SI joint is greatly needed.

INTRODUCTION

Imaging plays an increasingly important role in the assessment of suspected sacroiliitis in children, and magnetic resonance imaging (MRI) is the gold standard for evaluation of early disease. MRI findings consistent with sacroiliitis greatly influence treatment decisions for children, particularly the use of costly biologic therapy, and this fact underscores the importance of accurately diagnosing these patients to avoid both under- and overtreatment. In 2019, the American College of Rheumatology/Arthritis Foundation Guidelines advised that adding a tumor necrosis factor inhibitor is strongly recommended for children with continued active sacroiliitis despite nonsteroidal antiinflammatory drug monotherapy (1). MRI is used as the determining factor for defining sacroiliitis in those guidelines. Standard pelvic MRI sequences in children can be difficult to interpret, given the age-related changes that are primarily driven by marrow and cartilage maturation. To accurately identify pathology at the sacroiliac (SI) joints, the normal features of the maturing joint are critical to recognize.

[1]Pamela F. Weiss, MD, MSCE, Michael L. Francavilla, MD: Children's Hospital of Philadelphia and University of Pennsylvania Perelman School of Medicine, Philadelphia; [2]Timothy G. Brandon, MPH: Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; [3]John Bohnsack, MD: University of Utah, Salt Lake City; [4]Merav Heshin-Bekenstein, MD: University of California, San Francisco and Dana Children's Hospital of Tel Aviv Medical Center, Tel Aviv University, Tel Aviv, Israel; [5]Jacob L. Jaremko, MD, PhD, FRCPC, Lester Liao, MD, MTS, Dax Rumsey, MD, MSc: University of Alberta, Edmonton, Alberta, Canada; [6]Anne McHugh, MD: Stanford University, Stanford, California, and Wright State University Boonshoft School of Medicine, Dayton, Ohio; [7]Edward J. Oberle, MD, RhMSUS: Ohio State University College of Medicine and Nationwide Children's Hospital, Columbus, Ohio; [8]Hemalatha Srinivasalu, MD: Children's National Health System, Washington, DC; [9]Matthew L. Stoll, MD, PhD, MSCS: University of Alabama at Birmingham; [10]Nancy A. Chauvin, MD: Penn State Milton S. Hershey Children's Hospital, Hershey, Pennsylvania.

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

Address correspondence to Pamela F. Weiss, MD, MSCE, The Children's Hospital of Philadelphia, Roberts Center for Pediatric Research, 2716 South Street, 11th Floor, Room 11121, Philadelphia, PA 19146. Email: weisspa@chop.edu.

Submitted for publication December 6, 2019; accepted in revised form March 31, 2020.

SIGNIFICANCE & INNOVATIONS

- Imaging of the sacroiliac joint is challenging to interpret given the changes that occur with maturation.
- Significant variation and discordance exist in the interpretation of inflammatory and structural lesions at the sacroiliac joint in children across children's hospitals in North America.
- Findings consistent with active sacroiliitis on magnetic resonance imaging (MRI) greatly influence treatment decisions for children, particularly the use of costly biologic therapy, and underscore the importance of accurately diagnosing these patients to avoid both under- and overtreatment.
- Additional training for both radiologists and rheumatologists (adult and pediatric providers) in identifying active sacroiliitis on MRI in the maturing sacroiliac joint may improve the reliability of sacroiliac joint MRI interpretations.

Unfortunately, reference data for the appearance of the maturing skeleton are scarce. In 2019, the first article to document the physiologic changes in metaphyseal equivalent signal intensity in healthy children and adolescents was published by Chauvin et al (2). The key findings of that article were that prepubertal children frequently have homogeneous and symmetric increased metaphyseal-equivalent signal intensity along the sacral apophyses on fluid-sensitive sequences. During adolescence, this bright signal decreases with increased skeletal maturity. Chauvin et al also found that cortical irregularities are common, especially along the iliac surface of the joint, and that such irregularities were most prevalent in peripubertal children. These findings supported the results of an earlier autopsy-based study (3), which reported that the SI bony surfaces are smooth until puberty; after puberty, bony ridges and grooves can be seen, primarily on the ilium. Radiologists and rheumatologists need to recognize these features as variations in normal anatomy so that they are not mistaken for erosions, resulting in misdiagnosis and potentially unwarranted biologic therapy.

In clinical practice across institutions in the US and Canada, there is considerable variation in the musculoskeletal expertise and training of radiologists who interpret pelvic imaging for suspected sacroiliitis. This variability in training, along with unfamiliarity with the normal physiologic changes that occur at the maturing SI joint, raise concern that there are differences in interpretation of pathology both within and across institutions. In a single center study from the UK, interrater reliability was low to moderate for detection of bone marrow edema and erosion in children (4). We aimed to test the concordance of local interpretation of pelvic imaging from multiple sites across the US and Canada with that of a central imaging team with extensive experience evaluating the pediatric SI joint.

PATIENTS AND METHODS

The protocol for this multicenter study was reviewed and approved by the Children's Hospital of Philadelphia's Committee for the Protection of Human Subjects (IRB 17-014278). A waiver of consent/parental permission, a waiver of assent, and a waiver of Health Insurance Portability and Accountability Act authorization were granted for this retrospective study.

Subjects and study sites. Subjects were a retrospective convenience sample of children and adolescents who were consecutively imaged for suspicion of inflammatory sacroiliitis from 8 tertiary care children's hospitals across North America. Inclusion criteria were: 1) male or female patients, ages 6–17 years at the time of clinical care, 2) patients who underwent MRI of the pelvis for suspected sacroiliitis between January 1, 2015 and July 1, 2017, 3) MRI sequences that included a coronal T1-weighted and fluid-sensitive sequence (e.g., short tau inversion recovery or T2-weighted fat-saturated) with dedicated views of the SI joint, and 4) all subjects meeting the International League of Associations for Rheumatology (ILAR) juvenile idiopathic arthritis (JIA) criteria for enthesitis-related arthritis (ERA) or psoriatic arthritis. Children who met ERA criteria but were classified as having undifferentiated arthritis secondary to having a first-degree relative with psoriasis were also eligible. Cases were excluded if the front-line local radiologist was one of the musculoskeletal radiologists on the central imaging review team. Each site contributed up to 20 consecutive cases that met inclusion criteria.

Data. Data elements that were abstracted for each case included demographics (age, sex, race), ILAR JIA category, disease duration, indication for imaging, digital MRI scan, local radiologist report, radiologist experience (completed pediatric radiology fellowship, completed musculoskeletal fellowship, self-identification as a musculoskeletal radiologist, none of the above), the treating rheumatologist's interpretation of local radiology report (consistent with sacroiliitis yes/no), and documentation of changes to treatment regimen made based on local radiology report (yes/no). All local radiology reports were recorded at the point of care at each of the participating institutions and was done according to institutional standards. The central imaging team reviewed the MRI scans blinded to the local radiology report and all clinical details.

Prior to central imaging team review, all local radiologist findings were coded separately by 2 individuals (PFW and TGB) for the presence or absence of language in the local report indicating the following: bone marrow edema, erosion, sclerosis, and ankylosis. If a lesion was not noted in the report, we assumed it was absent. The global impression of sacroiliitis was coded based on the "Impression" section of the radiology report. If the impression did not differentiate between active inflammation and chronic sacroiliitis, cases were coded as "active inflammation" if bone marrow

edema in or along the SI joints was noted in the "Findings" section of the report. Discordant cases were reviewed and discussed until consensus was reached.

The central imaging team consisted of 3 radiologists (MLF, JLJ, and NAC) with extensive experience interpreting SI imaging in children. Two of the 3 raters (MLF and NAC) participated in a study that included review of SI imaging of 70 healthy children ages 7–18 years (2). All 3 central raters have been extensively calibrated to review inflammatory and structural lesions at the SI joint (5–8). Digital Imaging and Communications in Medicine–based anonymized cases were scored in randomized order, and study data were collected and managed using REDCap electronic data capture tools hosted at the Children's Hospital of Philadelphia (9). The Assessment of Spondyloarthritis international Society (ASAS) MRImagine consensus-based electronic case report form for recording MRI data was used to capture the following: global impression of acute/active inflammatory lesions compatible with sacroiliitis (yes/no), subchondral bone marrow edema, and global impression of structural lesions typical of axial spondyloarthritis: sclerosis, erosion, and ankylosis (10,11).

Lesions were rated according to previously published ASAS MRI definitions (12). The ASAS MRImagine case report form collects data on sidedness (left/right) and quadrant location (upper/lower and iliac/sacrum). All MRI scoring was dichotomized as either present (lesion noted in any quadrant on either the right or left SI joints) or absent for the analysis. All cases were reviewed by 2 central raters (MLF, JLJ, or NAC), who were blinded to clinical details and the local radiology report. When there was disagreement between the 2 raters regarding the presence or absence of inflammatory or structural lesions, the third radiologist, blinded to the type of lesion discrepancy, provided his/her interpretation of the case.

Statistical analysis. Subject demographic characteristics and local and central raters' assessments of lesions and global impression were summarized by frequencies and percentages or medians and interquartile ranges (IQRs). To compare local and central imaging assessment for lesions, all MRI scoring was dichotomized to either present or absent. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated to assess the performance of local reports in identifying active inflammation and lesions (present yes/no) using central impression of active inflammation and lesions (present yes/no) as the reference standard. The level of agreement between central reviewers and local reports, as well as among the central reviewers, was measured using percent agreement and raw concordance frequencies. Kappa coefficient was not included because it is highly dependent on the prevalence of the measured characteristic and assumes raters are the same for all subjects. All analyses were performed using Stata software, version14.2.

RESULTS

We enrolled 126 subjects to generate an evaluable population of 120 patients from 8 institutions across North America. Six cases were excluded due to failed transfer of files (n = 5) and age outside eligibility range (n = 1). The median age was 14 years (IQR 11.4–16) and half of all subjects were male. The majority of subjects were White/Caucasian (76.8%) and had a diagnosis of ERA (82.4%). The median disease duration at the time of imaging was 0.8 years (IQR 0–2), and the most frequent indication for imaging was back pain (65%). Local radiologist fellowships were primarily completed in pediatrics (96.7%), with 6.7% having completed a musculoskeletal-specific fellowship. Only 14.2% of local radiologists self-identified as a musculoskeletal radiologist. Local radiologists had a median of 12.8 years (range 0–43 years) experience as an attending physician.

Percent agreement on the presence/absence of active inflammation and chronic lesions was high among the central radiologists, 84% (95% confidence interval [95% CI] 79–89) and 86% (95% CI 81–91), respectively. The frequency of findings from the local radiology reports and the central imaging reviewers are

Table 1. Frequency of active inflammation and chronic lesions at the sacroiliac joints on MRI noted by local radiology reports and central radiologists across sites*

		Active inflammation†		Chronic lesions‡	
Site	No.	Local report	Central reviewers	Local report	Central reviewers
All	120	56 (46.7)	31 (25.8)	26 (21.7)	35 (29.2)
1	20	5 (25)	2 (10)	4 (20)	5 (25)
2	13	7 (53.8)	4 (30.8)	6 (46.2)	7 (53.8)
3	20	8 (40)	1 (5)	1 (5)	4 (20)
4	9	8 (88.9)	7 (77.8)	6 (66.7)	7 (77.8)
5	11	3 (27.3)	2 (18.2)	1 (9.1)	1 (9.1)
6	20	14 (70)	6 (30)	2 (10)	2 (10)
7	13	4 (30.8)	5 (38.5)	4 (30.8)	6 (46.2)
8	14	7 (50)	4 (28.6)	2 (14.3)	3 (21.4)

* Values are the frequency (%) unless indicated otherwise. MRI = magnetic resonance imaging.
† Defined as presence of acute/active inflammatory lesions (bone marrow edema) compatible with sacroiliitis.
‡ Presence of chronic lesions included erosion, sclerosis, fat metaplasia, and ankylosis.

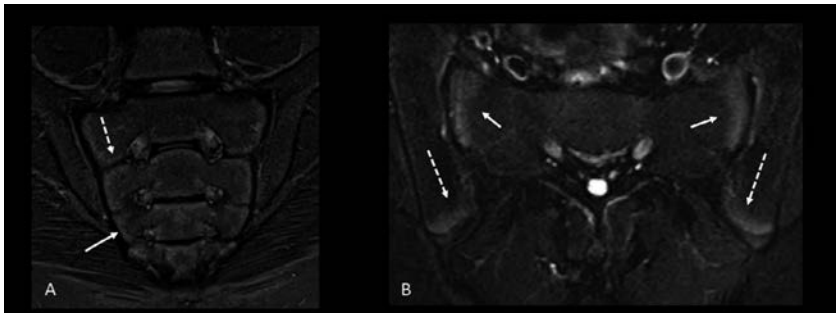

Figure 1. Representative short tau inversion recovery images of the sacrum in a female patient, age 11 years. **A**, Coronal oblique image of the sacrum shows homogeneous increased signal intensity within the periphery of the sacral ala, which appears uniform in thickness and symmetric. Homogenous increased signal intensity represents physiologic metaphyseal equivalent signal in the apophyses. The increased signal extends along the entire sacrum (solid arrow), not just the articular aspect, as well as within the cartilage between the sacral bodies (broken arrow). **B**, Axial T2-weighted fat-saturated image through the sacroiliac joints shows the hyperintense metaphyseal equivalent signal (solid arrows). Of note, the signal intensity is similar to the metaphyseal equivalent signal along the iliac crest apophyses (broken arrows). Local report interpretation: bone marrow edema present and indicative of sacroiliitis; central majority interpretation: normal.

shown in Table 1. Overall, subchondral bone marrow edema was noted more frequently by the local radiology reports (41.7%) than by the central reviewers (29.2%) across all sites except 2. Two examples of discordant local and central ratings for bone marrow edema are shown in Figure 1 (female patient, age 11 years) and Figure 2 (female patient, age 9 years). The frequency of chronic lesions indicative of structural damage (erosion, sclerosis, ankylosis) was noted more frequently by the central reviewers (29.2%) than by local reports (21.7%). An example of discordant local and central ratings for chronic lesions is shown in Figure 3 (male patient, age 17 years).

Table 2 shows the sensitivity, specificity, PPV, and NPV of local radiology reports in identifying active sacroiliitis and the presence of any chronic lesions using the central radiologists' findings (yes/no) as the reference standard. The sensitivity for active sacroiliitis on MRI, as depicted by subchondral marrow edema on fluid-sensitive sequences, was high and ranged from 80% to 100% across sites, with an overall sensitivity of 93.5% (95% CI 78.6–99.2). Specificity ranged from 0% to 100% across sites, with an overall specificity of 69.7% (95% CI 59.0–79.0). The PPV ranged widely from 12.5% to 100%, with an overall PPV of 51.8% (95% CI 38.0–65.3). The NPVs were high (95% CI 88.9–100) for all sites except 1.

Table 3 gives the rates of concordance and discordance for active inflammation and any chronic lesion. Across sites, the rate of discordance for active inflammation negative locally/positive centrally was very low (<2%). However, the rate of discordance for active inflammation positive locally/negative centrally ranged from 0% to 40%, with an overall discordance of 22.5% (27 of 120) across sites (Figures 1 and 2). Of those cases that were interpreted as active inflammation by the local radiologist but not by the central team, the median age was 13.4 years (IQR 11–16). Five cases were prepubertal (age <11 years), 8 were peripubertal (ages 11–13 years), and the remaining 14 were approaching skeletal maturity (ages 14–17 years).

In a sensitivity analysis that excluded all studies with any quality issues (n = 15) or missing coronal oblique sequences (n = 13), the rates of discordance were approximately the same

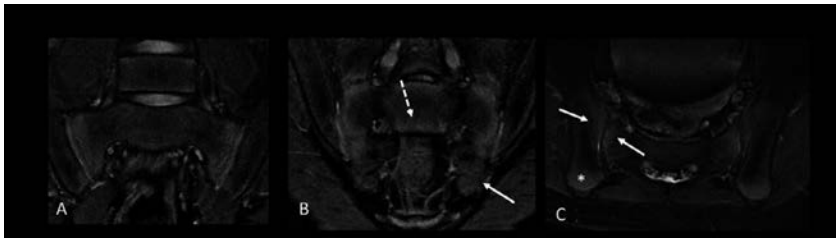

Figure 2. Representative short tau inversion recovery images of the sacrum in a female patient, age 9 years. **A** and **B**, Coronal oblique image of the sacrum shows homogeneous increased signal intensity within the periphery of the sacral ala and ilium, representing physiologic metaphyseal equivalent signal in the apophyses. Metaphyseal equivalent signal is more pronounced in this 9-year-old compared with the appearance of the older 11-year-old in Figure 1. The increased signal extends along the entire sacrum (solid arrow) as well as along the cartilage between the sacral bodies (broken arrow). **C**, Axial T2-weighted fat-saturated image through the sacroiliac joints shows the metaphyseal equivalent increased signal along both aspects of the joint (solid arrows) which is similar in signal along the iliac crest apophyses (*). Local report interpretation: bone marrow edema present and indicative of sacroiliitis; central majority interpretation: normal.

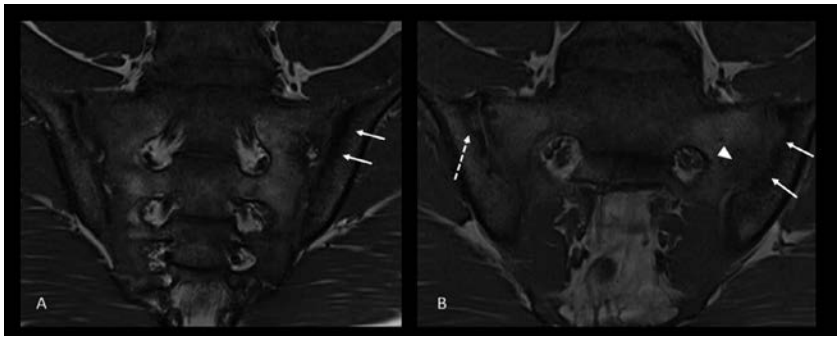

Figure 3. Representative T1-weighted coronal oblique image of the sacrum in a male patient, age 17 years. **A**, Dorsal aspect of the joint demonstrates hypointense signal along the left iliac wing, which extends >5 mm from the articular surface, indicative of sclerosis (solid arrows). **B**, Representative image along the ventral aspect of the joint shows extensive erosive change along the articular surface of the left iliac bone with loss of the normal cortex (solid arrows). There is early erosive change along the articular surface of the left sacrum with loss of the normal cortex (arrowhead). A small erosion is also seen within the right upper ilium (broken arrow). Local report interpretation: no structural changes noted; central majority interpretation: chronic inflammatory changes, including sclerosis and erosions.

(total unique excluded n = 24). The PPV and NPV for active inflammation were 50.0% (95% CI 34.6–65.4) and 98.1% (95% CI 89.9–100), respectively. The PPV and NPV for chronic lesions were 60.0% (95% CI 36.1–80.9) and 81.6% (95% CI 71.0–89.5), respectively.

In a second sensitivity analysis using only cases with total central radiologist agreement (no adjudication required), overall diagnostic test statistics were unchanged for active inflammation, with PPV and NPV of 52.4% (95% CI 36.4–68.0) and 95.9% (95% CI 86.0–99.5), respectively. PPV was slightly improved for the presence of chronic lesions at 70.0% (95% CI 45.7–88.1), and NPV was unchanged at 81.9% (95% CI 71.1–90.0).

The treating rheumatologist's interpretation of the study, as documented in the medical records (consistent with sacroiliitis yes/no), was in agreement with the local radiology report (sacroiliitis yes/no) for 88% of cases. Rheumatologists interpreted the case as negative when sacroiliitis was noted on the radiology report in 2 cases and in 12 cases interpreted the case as positive despite negative radiology reports. Of these discrepancies, the central radiologists agreed with the rheumatologist on both of the positive to negative calls and on only 1 of the 12 negative to positive calls. Of the 56 cases in which active inflammation was noted in the local radiology report, changes were made to the treatment regimen in 43 cases (76.8%). Nineteen of the 43 with medication

Table 2. Test properties of local radiology reports for detection of active inflammation and chronic lesions at the sacroiliac joints on MRI using majority central radiologists' interpretation as the reference standard*

Site	No.	Sensitivity	Specificity	PPV	NPV
Active inflammation†					
All	120	93.5 (78.6–99.2)	69.7 (59–79)	51.8 (38–65.3)	96.9 (89.2–99.6)
1	20	100 (15.8–100)	83.3 (58.6–96.4)	40 (5.3–85.3)	100 (78.2–100)
2	13	100 (39.8–100)	66.7 (29.9–92.5)	57.1 (18.4–90.1)	100 (54.1–100)
3	20	100 (2.5–100)	63.2 (38.4–83.7)	12.5 (0.3–52.7)	100 (73.5–100)
4	9	85.7 (42.1–99.6)	0 (0–84.2)	75 (34.9–96.8)	0 (0–97.5)
5	11	100 (15.8–100)	88.9 (51.8–99.7)	66.7 (9.4–99.2)	100 (63.1–100)
6	20	100 (54.1–100)	42.9 (17.7–71.1)	42.9 (17.7–71.1)	100 (54.1–100)
7	13	80 (28.4–99.5)	100 (63.1–100)	100 (39.8–100)	88.9 (51.8–99.7)
8	14	100 (39.8–100)	70 (34.8–93.3)	57.1 (18.4–90.1)	100 (59–100)
Chronic lesions‡					
All	120	45.7 (28.8–63.4)	88.2 (79.4–94.2)	61.5 (40.6–79.8)	79.8 (70.2–87.4)
1	20	40 (5.3–85.3)	86.7 (59.5–98.3)	50 (6.8–93.2)	81.3 (54.4–96)
2	13	57.1 (18.4–90.1)	66.7 (22.3–95.7)	66.7 (22.3–95.7)	57.1 (18.4–90.1)
3	20	25 (0.6–80.6)	100 (79.4–100)	100 (2.5–100)	84.2 (60.4–96.6)
4	9	71.4 (29–96.3)	50 (1.3–98.7)	83.3 (35.9–99.6)	33.3 (0.8–90.6)
5	11	0 (0–97.5)	90 (55.5–99.7)	0 (0–97.5)	90 (55.5–99.7)
6	20	50 (1.3–98.7)	94.4 (72.7–99.9)	50 (1.3–98.7)	94.4 (72.7–99.9)
7	13	33.3 (4.3–77.7)	71.4 (29–96.3)	50 (6.8–93.2)	55.6 (21.2–86.3)
8	14	33.3 (0.8–90.6)	90.9 (58.7–99.8)	50 (1.3–98.7)	83.3 (51.6–97.9)

* Values are the percentage (95% confidence interval) unless indicated otherwise. MRI = magnetic resonance imaging; NPV = negative predictive value; PPV = positive predictive value.
† Defined as presence of acute/active inflammatory lesions (bone marrow edema) compatible with sacroiliitis.
‡ Presence of chronic lesions included erosion, sclerosis, fat metaplasia, and ankylosis.

Table 3. Rates of concordance and discordance of active inflammation and chronic lesion detection at the sacroiliac joint on MRI between central radiologists and local radiology reports*

Site	No.	Central (+)/ local (+)	Central (+)/ local (–)	Central (–)/ local (+)	Central (–)/ local (–)
Active inflammation†					
All	120	29 (24.2)	2 (1.7)	27 (22.5)	62 (51.7)
1	20	2 (10)	0 (0)	3 (15)	15 (75)
2	13	4 (30.8)	0 (0)	3 (23.1)	6 (46.2)
3	20	1 (5)	0 (0)	7 (35)	12 (60)
4	9	6 (66.7)	1 (11.1)	2 (22.2)	0 (0)
5	11	2 (18.2)	0 (0)	1 (9.1)	8 (72.7)
6	20	6 (30)	0 (0)	8 (40)	6 (30)
7	13	4 (30.8)	1 (7.7)	0 (0)	8 (61.5)
8	14	4 (28.6)	0 (0)	3 (21.4)	7 (50)
Chronic lesions‡					
All	120	16 (13.3)	19 (15.8)	10 (8.3)	75 (62.5)
1	20	2 (10)	3 (15)	2 (10)	13 (65)
2	13	4 (30.8)	3 (23.1)	2 (15.4)	4 (30.8)
3	20	1 (5)	3 (15)	0 (0)	16 (80)
4	9	5 (55.6)	2 (22.2)	1 (11.1)	1 (11.1)
5	11	0 (0)	1 (9.1)	1 (9.1)	9 (81.8)
6	20	1 (5)	1 (5)	1 (5)	17 (85)
7	13	2 (15.4)	4 (30.8)	2 (15.4)	5 (38.5)
8	14	1 (7.1)	2 (14.3)	1 (7.1)	10 (71.4)

* Values are the number (%) unless indicated otherwise. Central = central radiologist reviewers; Local = local radiology report; MRI = magnetic resonance imaging; (+) = positive for active inflammation or presence of chronic lesions; (–) = negative for active inflammation or presence of chronic lesions.
† Defined as presence of acute/active inflammatory lesions (bone marrow edema) compatible with sacroiliitis.
‡ Presence of chronic lesions included erosion, sclerosis, fat metaplasia, and ankylosis.

changes (44.2%) were interpreted as not consistent with inflammation by the central reviewers.

DISCUSSION

We found a wide range of discrepancies between local and central radiologist interpretations of active and chronic sacroiliitis in pediatric SI joint MRI. These results highlight the fact that imaging of this joint is challenging to interpret. Key challenges include the age-related changes of the metaphyseal equivalent signal in the sacral apophyses and lack of materials available to familiarize radiologists with these changes. The high sensitivity for detection of active inflammation indicates that very few cases of active sacroiliitis were missed by local radiologist reports. However, the PPV was low, indicating frequent false positive results, presumably driven by age-related increased metaphyseal signal intensity being assessed as pathologic subchondral bone marrow edema. The sensitivity and PPV for chronic lesions were low. There was a high rate of agreement between the treating rheumatologists and local radiology reports. This agreement may be due to several factors, including collaboration and familiarity with team members at each hospital, the fact that many rheumatologists may not independently review imaging studies and instead rely on the radiology interpretation, and the fact that the issues that obscure interpretation of SI imaging may affect both radiologists and rheumatologists. Local radiology reports indicating active inflammation resulted in treatment changes by the rheumatologist for over 75% of cases, and almost half of these were interpreted as not consistent with active inflammation by the central reviewers. This inconsistency has important implications for children in regard to unnecessary exposure to biologic medications, potential side effects, and health care costs.

Our findings must be interpreted in the setting of several caveats. First, there was no standardized assessment for local interpretation, and imaging results were recorded according to institutional standards. Additionally, lesions were defined locally according to the clinical experience and training of the radiologist at the point of care. However, this variability was what this study aimed to capture. Second, the central imaging team was blinded to clinical details of the cases, whereas local radiologists had access to clinical information within the patient medical record, prior MRI studies (when available), and other imaging modality results. Local findings used for this analysis only represent abnormalities and impressions included in the radiology report and do not capture any consultation or dialogue that happened among the clinicians outside the MRI result report. If a second local radiologist had interpreted the study and adjudication was done where there was disagreement, the results may have been different. However, this study aimed to assess the accuracy of the clinical reports recorded in the medical chart and used to guide clinical decision-making at the point of care, not the accuracy of the individual radiologists.

Third, the imaging quality of the studies was highly variable, even within the same institution. Poor study quality can impact

the ability to make accurate and reliable assessments of inflammatory and structural lesions at the SI joint (13). For example, not all studies contained a coronal oblique view of the SI joint, making visualization of the synovial part of the joint more difficult. However, we found that even when studies of suboptimal quality or studies missing the coronal oblique view were excluded from analysis, the rates of discordance remained high.

Fourth, there is a possibility of sampling error because only a small number of studies (up to 20) were evaluated from each institution. Since site collaborators were instructed to choose consecutive studies, we anticipate the samples are representative of the case mix seen at varying time points. In addition, because all of the patients had juvenile arthritis, the prevalence of sacroiliitis was higher than would be expected in a study of otherwise healthy children with inflammatory back pain. Therefore, conclusions of this study cannot necessarily be extrapolated to that larger group of patients. Next, although the agreement among the central reviewers was high, it was not perfect, and several cases required adjudication. In a sensitivity analysis using only cases with total central radiologist agreement, overall diagnostic test statistics were largely unchanged for active inflammation and slightly improved for the presence of chronic lesions. Lastly, the imaging studies at each institution were read by multiple radiologists, and the interpretations at each site were considered collectively. Therefore, these results reflect the average SI imaging milieu across sites and are not reflective of anyone's individual performance. Perhaps most importantly, we note that a true external gold-standard pathologic diagnosis of sacroiliitis was not available in any case. This study focused on assessment of the reliability of MRI interpretations, not the accuracy of MRI findings in terms of diagnosis and clinical outcomes.

Recognizing inflammatory change from physiologic changes of the maturing SI joint is essential for radiologists interpreting pediatric pelvic MRI scans. Deficiencies in pediatric musculoskeletal imaging evaluation may be due to limitations of experience during fellowship training as well as to lack of available online educational tools focused on pediatric musculoskeletal radiology. Pediatric radiology fellowships are characterized by a diversity of educational experience, and perhaps a more formal needs assessment should be conducted to evaluate perceived deficiencies. In a study by Yablon et al (14), a needs assessment of musculoskeletal radiologists was performed to evaluate their musculoskeletal training experience. While the musculoskeletal radiologists believed that they were adequately trained for practice, pediatric musculoskeletal imaging was acknowledged to be a deficiency. To our knowledge, there is no MRI atlas that depicts the normal appearance of the maturing SI joint.

After review of our results, most of the discordance is perceived to be error differentiating normal physiologic metaphyseal equivalent signal from pathologic subchondral marrow edema. Five misclassified cases were prepubertal, 8 were peripubertal, and the remaining 14 were approaching skeletal maturity. In the study by Chauvin et al, increased signal on fluid-sensitive signal within the sacral metaphyseal equivalents was common along the healthy SI joint in all 3 of these groups, though less frequent as skeletal maturation was reached (2). One factor that aids in distinguishing normal physiologic metaphyseal signal from inflammatory edema on fluid-sensitive sequences is that the normal signal is uniform and symmetric, and that it follows the contours of the articular surface and extends along the length of the sacrum, inferior to the articulation.

In summary, we found that significant variation and discordance exist in the interpretation of inflammatory and structural lesions at the SI joint in children across children's hospitals in North America. Additional training for both radiologists and rheumatologists (adult and pediatric providers) in identifying active sacroiliitis on MRI in the maturing SI joint may improve the reliability of SI joint MRI interpretations. Diagnosis and treatment decisions for children never rely solely on imaging results and do take into account aspects of disease that are not reflected in the imaging results, such as patient/family preferences. However, improving the reliability of interpretations of pathology at the SI joint will increase the utility of MRI in the management of pediatric spondyloarthritis.

ACKNOWLEDGMENTS

The authors thank the following individuals who helped with various stages of identification, abstraction, and transfer of subject data: Eric Brown (Children's Hospital of Philadelphia), Jeffrey Prince and Suzy Jones (University of Utah), Tzielan Lee and Omar Hernandez (Stanford University), Houchun Hu (Nationwide Children's Hospital), Prem Madesh (Virginia Commonwealth University), and Joshua Pavlovec (University of Alabama at Birmingham).

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. Weiss 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. Weiss.

Acquisition of data. Weiss, Brandon, Bohnsack, Heshin-Bekenstein, Francavilla, Jaremko, Liao, McHugh, Oberle, Rumsey, Srinivasalu, Stoll, Chauvin.

Analysis and interpretation of data. Brandon.

REFERENCES

1. Ringold S, Angeles-Han ST, Beukelman T, Lovell D, Cuello CA, Becker ML, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the treatment of juvenile idiopathic arthritis: therapeutic approaches for non-systemic polyarthritis, sacroiliitis, and enthesitis. Arthritis Care Res (Hoboken) 2019;71:717–34.
2. Chauvin NA, Xiao R, Brandon TG, Biko DM, Francavilla M, Khrichenko D, et al. MRI of the sacroiliac joint in healthy children. AJR Am J Roentgenol 2019;212:1303–9.
3. Vleeming A, Schuenke MD, Masi AT, Carreiro JE, Danneels L, Willard FH. The sacroiliac joint: an overview of its anatomy, function and potential clinical implications. J Anat 2012;221:537–67.
4. Orr KE, Andronikou S, Bramham MJ, Holjar-Erlic I, Menegotto F, Ramanan AV. Magnetic resonance imaging of sacroiliitis in children:

frequency of findings and interobserver reliability. Pediatr Radiol 2018;48:1621–8.

5. Weiss PF, Xiao R, Brandon TG, Biko DM, Maksymowych WP, Lambert RG, et al. Radiographs in screening for sacroiliitis in children: what is the value? Arthritis Res Ther 2018;20:141.
6. Weiss PF, Maksymowych WP, Lambert RG, Jaremko JL, Biko DM, Paschke J, et al. Feasibility and reliability of the Spondyloarthritis Research Consortium of Canada sacroiliac joint inflammation score in children. Arthritis Res Ther 2018;20:56.
7. Weiss PF, Maksymowych WP, Lambert RG, Jaremko JL, Biko DM, Paschke J, et al. Feasibility and reliability of the Spondyloarthritis Research Consortium of Canada sacroiliac joint structural score in children. J Rheumatol 2018;45:1411–7.
8. Maksymowych WP, Krabbe S, Biko D, Weiss P, Maksymowych MP, Cheah J, et al. Validation of web-based calibration modules for imaging scoring systems based on principles of artificial intelligence: the SPARCC MRI sacroiliac joint inflammation score [abstract]. Ann Rheum Dis 2018;77 Suppl 2:822.
9. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81.
10. Maksymowych W, Lambert RG, Ostergaard M, de Hooge M, Pedersen SJ, Bennett AN, et al. MRI lesion definitions in axial spondyloarthritis: a consensus reappraisal from the Assessments in SpondyloArthritis International Society (ASAS) [abstract]. Ann Rheum Dis In press.
11. Maksymowych W, Ostergaard M, Lambert RG, Weber U, Pedersen SJ, Sieper J, et al. The contribution of structural MRI lesions to detection of sacroiliitis in patients in the assessments in SpondyloArthritis International Society (ASAS) classification cohort [abstract]. Ann Rheum Dis In press.
12. Lambert RG, Bakker PA, van der Heijde D, Weber U, Rudwaleit M, Hermann KG, et al. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis 2016;75:1958–63.
13. Orr KE, Andronikou S, Bramham MJ, Holjar-Erlic I, Menegotto F, Ramanan AV. Overcoming two technical pitfalls in MRI of paediatric and adolescent sacroiliitis. Clin Radiol 2019;74:235–41.
14. Yablon CM, Wu JS, Newman LR, Downie BK, Hochman MG, Eisenberg RL. A needs assessment of musculoskeletal fellowship training: a survey of practicing musculoskeletal radiologists. AJR Am J Roentgenol 2013;200:732–40.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 849–855
DOI 10.1002/acr.24203

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Ultrasound-Guided Biopsy of Suspected Salivary Gland Lymphoma in Sjögren's Syndrome

Alan N. Baer, Thomas Grader-Beck, Brendan Antiochos, Julius Birnbaum, and Joel M. Fradin

Objective. To evaluate the safety and utility of core needle biopsy (CNB) for diagnosis of salivary gland lymphoma in Sjögren's syndrome (SS).

Methods. We analyzed data from consecutive SS patients who underwent ultrasound-guided major salivary gland CNB for lymphoma diagnosis and determined whether CNB yielded an actionable diagnosis without need for further intervention.

Results. CNBs were performed in 24 patients to evaluate discrete parotid (n = 6) or submandibular (n = 2) gland masses or diffuse enlargement (n = 16; 15 parotid). One patient had 3 CNBs of the same mass. Of the 26 CNBs, 24 included flow cytometry, using CNB and/or fine needle aspirate material, and 14 targeted sonographically identified focal lesions. No patient reported complications. In the 23 patients with 1 CNB, final diagnoses were marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT; n = 6), atypical lymphoid infiltration (n = 3), benign lymphoepithelial sialadenitis (n = 9), normal gland tissue (n = 4), and lymphoepithelial cyst (n = 1). In the patient with serial CNBs, the initial one without flow cytometry was benign, but the next 2 showed atypical lymphoid infiltration. Monoclonal lymphoid infiltration was detected in 12 patients: 6 with MALT lymphoma, 3 were benign, and 3 with atypical lymphoid infiltration. Of the latter 3, 1 was treated with rituximab and 2 with expectant observation. The diagnosis changed from atypical lymphoid infiltration to MALT lymphoma in 1 patient following biopsy of inguinal adenopathy 6 months post-CNB. CNB provided actionable results and avoided open excisional biopsies in all cases.

Conclusion. CNB is safe and useful in the evaluation of suspected salivary gland lymphoma in SS.

INTRODUCTION

Ultrasound-guided core needle biopsy (CNB) is an important technique in the diagnosis of salivary gland masses (1). The procedure uses ultrasonography to guide the placement of a handheld spring-loaded device to obtain core samples of a salivary gland lesion. In contrast to fine needle aspiration (FNA) cytology, CNB provides biopsy samples with preserved architecture that are amenable to immunohistochemical staining and that can be used for staging and grading of neoplasms. This possibility with CNB is particularly relevant to the diagnosis of lymphoma, where analysis of tissue histopathology is essential. In a meta-analysis, sensitivity and specificity for the diagnosis of malignant salivary gland lesions were estimated to be 96% and 100%, respectively (1). A portion of the tissue can also be used for flow cytometry, thereby increasing the diagnostic yield in the evaluation of lymphoma (2). CNB has a superior diagnostic yield over FNA alone and is associated with few complications (1).

Patients with Sjögren's syndrome (SS) have a 6- to 18-fold increased risk of non-Hodgkin lymphoma, particularly marginal-zone lymphoma involving the parotid gland (3). The development of a salivary gland mass or persistent enlargement can be a sign of development of lymphoma. FNA can yield material for flow cytometry and the detection of a monoclonal B cell population, but flow cytometric analysis is not sufficient for the diagnosis of lymphoma. A larger sample is required for histopathology, including immunohistochemical staining. While excisional biopsies are preferred for lymphoma diagnosis, the trend has been to employ CNB whenever possible (4). Identification of salivary gland lymphoma with CNB avoids surgical resection, most often superficial

Supported by the Jerome L. Greene Foundation. Dr. Baer's work was supported by the NIH (RO1-DE-12354-15A1 and 75N92019P00427).

Alan N. Baer, MD, Thomas Grader-Beck, MD, Brendan Antiochos, MD, Julius Birnbaum, MD, MHS, Joel M. Fradin, MD: Johns Hopkins University School of Medicine, Baltimore, Maryland.

Dr. Baer has received consulting fees from Bristol Myers Squibb and Sanofi-Aventis (less than $10,000 each). Dr. Grader-Beck has received consulting fees from Novartis and Eli Lilly and Company (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Alan N. Baer, MD, Division of Rheumatology, 5200 Eastern Avenue, MFL Center Tower, Suite 4000, Baltimore, MD 21224. Email: alanbaer@jhmi.edu.

Submitted for publication August 2, 2019; accepted in revised form March 24, 2020.

SIGNIFICANCE & INNOVATIONS

- Salivary gland lymphoma has a cumulative lifetime prevalence of 5–10% in patients with Sjögren's syndrome (SS).
- Biopsy procedures that use an ultrasound-guided core needle have gained increasing use for the diagnosis of lymphoma and have the advantage of avoiding a surgical excision with its attendant risks.
- In the current study, ultrasound-guided core needle biopsy and pathologic analysis with both histopathology and flow cytometry in 24 SS patients was safe and avoided the need for excisional biopsy in all.
- This study supports the routine use of the biopsy procedure in the evaluation of suspected salivary gland lymphoma in SS.

parotidectomy, with its attendant risks. We sought to evaluate the safety and utility of CNB in the diagnostic evaluation of salivary gland lymphoma in patients with SS.

PATIENTS AND METHODS

Patients. We identified all SS patients who underwent ultrasound-guided CNB of either the parotid or submandibular gland as part of a diagnostic evaluation for salivary gland lymphoma in the Johns Hopkins Sjögren's Syndrome Center between July 2009 and August 2018. Clinical data on patients seen in the Center are maintained in a computer database, under a protocol approved by the Johns Hopkins Institutional Review Board. SS was defined by fulfillment of the American College of Rheumatology/European Alliance of Associations for Rheumatology classification criteria (5). None of the patients had a prior history of lymphoma.

Technique. CNBs were performed by radiologists with extensive experience in ultrasonography. Biopsy samples were typically obtained with an 18-gauge INRAD core biopsy needle (INRAD Inc). To accurately place the device either within the glandular parenchyma or within a focal intraglandular lesion, the CNB was performed under real-time ultrasound guidance, using an 18- or 12-mHz linear transducer (Figure 1). Typically, 2 passes were made in the same or very nearly the same location using the same core biopsy device. One core was placed in formalin for standard surgical pathology examination, and the other was placed in culture medium for flow cytometry. For parotid gland biopsy, the patient was placed in the decubitus position. The entry site was typically from the posterior glandular border, slightly caudal to the level of attachment of the ear lobe. Local infiltration anesthesia was employed in the skin and subcutaneous fat up to the surface of the gland but not within the gland. A small skin incision was made with a scalpel. The biopsy needle was directed from posterior to anterior, with care taken to keep the needle path located superficially within the gland parenchyma and the entire trough of the needle within the parenchyma.

For submandibular gland biopsy, the procedure was similar but with the patient in an oblique position, head mildly extended and rotated away from the operator. The approach was also from posterior to anterior. In the case of focal lesions, care was taken to minimize traversal of normal parotid parenchyma and to proceed from a posterior to anterior approach. FNA was routinely included in the same procedure as the CNB and targeted the same lesion or glandular location as the CNB. FNA and CNB material were routinely submitted together for flow cytometry. Patients were contacted by telephone by radiology department nurses on the day following the biopsy to determine whether any symptoms suggesting complications were present. Patients were queried as to whether swelling, pain, bleeding, or facial changes had taken place.

Pathologic interpretation and data analysis. All biopsy samples with suspicious lymphoid infiltrates were reviewed by hematopathologists, and a final diagnosis was achieved through integration of histopathology, immunohistochemistry, and flow cytometry and/or molecular studies. The utility of CNB was judged by whether it yielded an actionable diagnosis, allowing for either initiation of treatment or final determination of benignity requiring no further intervention (6).

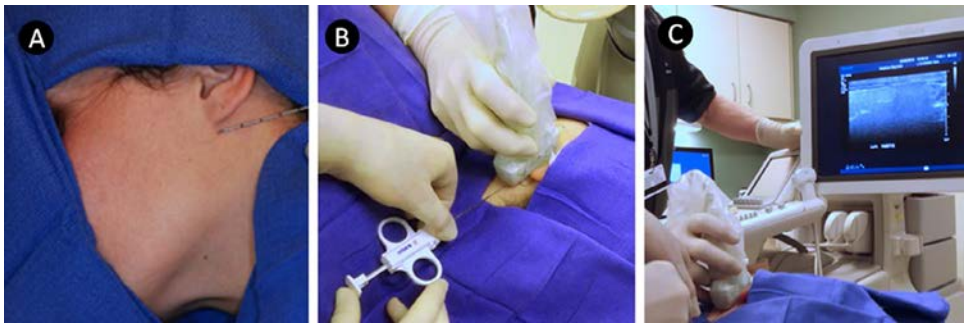

Figure 1. Parotid gland core needle biopsy with ultrasound guidance. **A**, The biopsy needle is directed from posterior to anterior, with care taken to keep the needle path located superficially within the gland parenchyma and the entire trough of the needle within parenchyma; **B**, Patient is in right side down decubitus position with head on the right side of the photograph as the biopsy needle approaches the left parotid gland from its posterior aspect. The biopsy is guided by real-time ultrasound, under aseptic conditions; **C**, Both the operator and the assisting sonographer monitor the path of the needle on the ultrasound screen.

RESULTS

Twenty-four SS patients underwent ultrasound-guided CNB to evaluate for possible salivary gland lymphoma (Table 1). The cohort included 22 women and 2 men with a median age of 53 years (range 18–74 years). Two patients had secondary SS. Additional phenotypic features are listed in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24203/abstract. The salivary gland abnormalities were bilateral parotid (n = 13), unilateral parotid (n = 2), or submandibular (n = 1) gland enlargement, and discrete masses in the parotid (n = 6) or submandibular (n = 2) glands. There were 26 procedures in total, with 1 patient (patient 16) undergoing 3 CNBs of the same mass over a 17-month period. The parotid was biopsied in 23 patients and the submandibular gland in 3 patients. The 26 procedures involved CNB alone (n = 5) or with FNA (n = 21). Sampling was restricted to 1 gland only, except for 2 patients (patients 4, 19) who underwent CNBs of both parotid glands. The average number of CNBs per gland per procedure was 2.29 ± 0.66 (range 1–4). Cyst aspiration was performed during 4 procedures. Biopsy material was sent for flow cytometry in 24 procedures; this material consisted of CNB specimens alone in 5 procedures or combined with FNA material in 21. None of the patients reported complications 1 day post-procedure or during longitudinal follow-up in our center (median [range] duration = 595 days [1–1,403 days]).

Small hypoechoic ovoid lesions, a characteristic ultrasonographic abnormality of SS, were present in the parotid gland parenchyma of 18 patients (7). In 16 of these 18, the hypoechoic lesions in composite occupied >50% of the glandular volume, corresponding to at least grade 2 severity on the Outcome Measures in Rheumatology scoring system (7).

Representative (i.e., random) sampling of glandular tissue was performed in 12 patients, opting for the salivary gland that was most enlarged. Targeted sampling of discrete sonographic lesions was performed in 12 patients. The targets included clinically palpable masses with corresponding sonographic hypoechoic solid or cystic mass lesions (n = 6; patients 11, 13, 15, 17, 21, 23) or partly solid/partly cystic masses (n = 2; patients 7, 16). Targets also included nonpalpable sonographically defined lesions, including a mass-like grouping of hypoechoic lesions (n = 1, patient 24) and a partly solid/partly cystic mass (n = 1; patient 22), as well as intraparotid lymph nodes that appeared sonographically abnormal (n = 2, patients 12, 18) (Figure 2). CNB was performed 3 times on the same mass lesion in patient 16, following an initial FNA alone with flow cytometry that showed 9% clonal B cells. The first CNB did not include repeat FNA/flow cytometry, while the second did. A third CNB was performed 16 months after the second because of progressive enlargement of the mass lesion.

A final pathologic diagnosis was established through the integration of CNB histopathology with immunohistochemistry (n = 12 procedures), cytology (n = 21), flow cytometry (n = 24), and molecular studies (n = 2). In 2 procedures (patients 2, 7), flow cytometry samples were largely comprised of debris and were thus without diagnostic utility (one was material from CNB only and the other from combined CNB and FNA) (see Supplementary Figures 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24203/abstract). Final pathologic diagnoses for the 23 patients with only 1 CNB were marginal-zone lymphoma of mucosa-associated lymphoid tissue (MALT; n = 6), atypical lymphoid infiltration suspicious for MALT lymphoma (n = 2), benign focal lymphoid infiltration (n = 8), benign lymphoepithelial sialadenitis (BLEL; n = 2), salivary gland tissue without inflammatory infiltrate (n = 4), and lymphoepithelial cyst/focal lymphoid infiltration (n = 1 with biopsies of right and left parotid glands). The diagnosis of MALT lymphoma was established a median of 6.1 years (range 1.3–8.0 years) after that of SS in the 6 affected patients. In the patient with serial CNBs, the first was interpreted as benign focal lymphoid infiltration, whereas the second (which included flow cytometry) confirmed an atypical lymphoid infiltrate seen on an earlier FNA. In the third CNB, a monoclonal B cell population was demonstrated, again by flow cytometry, but the lymphoid infiltrate in the CNB was too small for further characterization.

A clonally restricted cell population was detected by flow cytometry in 12 biopsy samples (B cell in 11, plasma cell in 1) performed on 11 patients and by molecular studies in additional procedure (patient 17). Among these 12 patients with monoclonal lymphoid infiltrates, 6 were diagnosed with MALT lymphoma based on histopathology and immunohistochemistry. Among the other 6 patients, 3 had interpretations of atypical lymphoid infiltration, but with insufficient morphologic findings to diagnose lymphoma. In 1 of these 3 patients (patient 18) with an atypical lymphoid infiltrate, biopsy of an enlarged inguinal node 6 months after the parotid biopsy showed MALT lymphoma. The other 3 (patients 5, 6, 15) were diagnosed with BLEL (patients 5, 6) or a benign T cell–predominant focal lymphocytic infiltrate (patient 15). The 2 patients described above with inadequate samples for flow cytometry (patients 2, 7) had no lymphoid infiltrate on CNB. Hypoechoic sonographic lesions were present in all patients with lymphoma, atypical lymphoid infiltration, and/or monoclonal B cell population detected in the CNB.

In 15 patients, the clinician concluded that the salivary gland process was benign and opted to make no change in treatment in 12 patients. In 3, the clinician opted to treat with rituximab due to persistent parotid gland enlargement (patients 8, 14) or the presence of interstitial nephritis (patient 15). In 3 patients, the clinician concluded that the patient had a possible lymphoproliferative disorder; 1 received rituximab treatment (patient 17), and 2 were observed expectantly (patients 16, 18). One of these latter patients (patient 18) developed inguinal adenopathy 6 months later, with a biopsy result showing MALT lymphoma. Of the 6 patients with MALT lymphoma, 5 were treated with rituximab monotherapy,

Table 1. Clinical abnormality, biopsy site, biopsy interpretation, and action taken based on biopsy results*

Case	Age (years), sex	Clinical concern	Site of biopsy	Random sampling or targeted focal lesion	Type of biopsy	Histology results	Monoclonal population on flow cytometry	Final clinical impression/action taken
1	56, M	Enlarged PG (R/L)	PG (R)	Random	Core only	No lymphoid infiltrate	–	Benign; no intervention
2	43, F	Enlarged PG (R/L)	PG (L)	Random	Core only	No lymphoid infiltrate	Mostly debris	Benign; no intervention
3	56, F	Enlarged PG (R/L)	PG (L)	Random	Core/FNA	No lymphoid infiltrate	None	Benign; no intervention
4	58, F	Enlarged PG (R/L)	PG (R/L)	Random	Core/FNA	LE cyst (R); FLS (L)	None	Benign; no intervention
5	58, F	Enlarged PG (R/L)	PG (L)	Random	Core/FNA	Benign LES	B cell (2%)	Benign; no follow-up available
6	33, F	Enlarged PG (R/L)	PG (R)	Random	Core/FNA	Benign LES	B cell (12%)	Benign; no intervention
7	39, F	PG mass (R)	PG (R)	2 solid/cystic masses	Core/FNA	No lymphoid infiltrate	Mostly debris	Benign; no intervention
8	36, F	Enlarged PG (R/L)	PG (R)	Random	Core/FNA	FLS	None	Benign; rituximab
9	53, F	Enlarged PG (R)	PG (R)	Random	Core only	FLS	–	Benign; no intervention
10	74, F	Enlarged SMG on US (L)	SMG (L)	Random	Core/FNA	FLS	None	Benign; no intervention
11	59, F	SMG mass (R)	SMG (R)	Hypoechoic mass	Core/FNA	FLS	None	Benign; no intervention
12	59, F	Enlarged PG (R)	PG (R)	Abnormal intra-PG LN	Core/FNA	FLS	None	Benign; no intervention
13	50, F	PG mass (R)	PG (R)	Grouped hypoechoic lesions	Core/FNA	FLS	None	Benign; no intervention
14	18, F	Enlarged PG (R/L)	PG (L)	Random	Core only	FLS	None	Benign; RTX
15	62, F	PG mass (R)	PG (R)	2 cystic masses	Core/FNA	FLS (mostly T cell)	Atypical B cell (4%), lacking surface light chains	Benign; RTX
16a	53, M	PG mass (R)	PG (R)	Solid/cystic mass	Core/FNA	FLS	–	LPD; repeat biopsy†
16b	53, M	PG mass (R)	PG (R)	Solid/cystic mass	Core only	Atypical lymphoid infiltrate	B cell	LPD; expectant observation
16c	54, M	PG mass (R)	PG (R)	Abnormal intra-PG LN	Core/FNA	FLS	B cell (11%)	LPD; expectant observation
17	66, F	PG mass (R)	PG (R)	Grouped hypoechoic lesions	Core/FNA	Atypical lymphoid infiltrate	None‡	LPD; RTX
18	50, F	Enlarged PG (R/L)	PG (L)	Abnormal intra-PG LN	Core/FNA	Atypical lymphoid infiltrate	B cell (17%)‡	LPD; expectant observation§
19	59, F	Enlarged PG (R/L)	PG (bilateral)	Random	Core/FNA	MALT lymphoma	Suspicious B cell	Lymphoma; RTX
20	47, F	Enlarged PG (R/L)	PG (R)	Random	Core/FNA	MALT lymphoma/ plasmacytic differentiation	Plasmacytic (26%)	Lymphoma; RTX

(Continued)

Table 1. *(Cont'd)*

Case	Age (years), sex	Clinical concern	Site of biopsy	Random sampling or targeted focal lesion	Type of biopsy	Histology results	Monoclonal population on flow cytometry	Final clinical impression/action taken
21	52, F	PG mass (L)	PG (L)	Grouped hypoechoic lesions	Core/FNA	MALT lymphoma	B cell (20%)	Lymphoma; RTX
22	31, F	Enlarged PG (R/L)	PG (L)	Solid/cystic mass	Core/FNA	MALT lymphoma	B cell	Lymphoma; RTX
23	58, F	SMG mass (R)	SMG (R)	Multilobulated solid mass	Core/FNA	MALT lymphoma	B cell (10%)	Lymphoma; RTX
24	46, F	Enlarged PG (R/L)	PG (R)	Grouped hypoechoic lesions	Core/FNA	MALT lymphoma	B cell (41%)	Lymphoma; RTX

* F = female; FLS = focal lymphocytic sialadenitis; FNA = fine needle aspirate; L = left; LE = lymphoepithelial; LES = lymphoepithelial sialadenitis; LN = lymph node; LPD = lymphoproliferative disorder, but insufficient evidence to classify as lymphoma; M = male; MALT = mucosa-associated lymphoid tissue; PG = parotid gland; R = right; RTX = rituximab; SMG = submandibular gland; US = ultrasound.
† Flow cytometry had shown a monoclonal B cell population on an FNA of the mass one month earlier, prompting continued concern for a possible lymphoproliferative disorder.
‡ IgH gene rearrangement studies documented a B cell clonal population.
§ Diagnosis of MALT lymphoma established 6 months later, via biopsy of new inguinal lymphadenopathy.

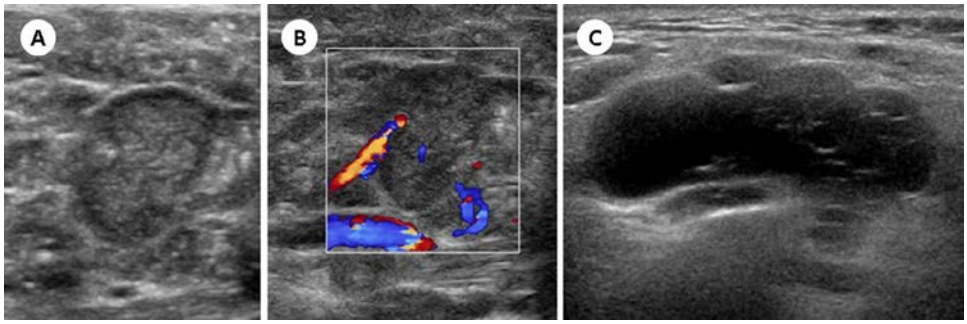

Figure 2. Ultrasound images of lymphomatous lesions of the parotid gland. **A**, An abnormal intraparotid lymph node is evident in this longitudinal view (patient 18). The node was considered abnormal because of its enlargement and rounded shape; **B**, A blood vessel is seen on color Doppler imaging penetrating the hilum of the intraparotid node seen in panel A; **C**, A heterogeneously hypoechoic, mixed solid and cystic mass is evident in this longitudinal view (patient 22).

while 1 was managed with expectant observation but then started on rituximab after an interval of 12 months.

DISCUSSION

We used ultrasound-guided CNB to evaluate salivary gland abnormalities suspicious for lymphoma in 24 patients with SS. The indications included persistent salivary gland enlargement or a discrete salivary gland mass or lesion identified clinically or by ultrasound. None of the patients had a rapidly enlarging mass, minimizing concern for a high-grade lymphoma (e.g., diffuse large B cell), which can involve the parotid gland (8). Importantly, flow cytometry of cellular material from the CNB and/or FNA was included to increase the diagnostic yield in 24 of the 26 procedures. None of our patients underwent open surgical resection of salivary gland tissue following CNB, so validation of the utility of this diagnostic protocol rests on whether the results were actionable, allowing a clinical decision to be made without the need for further testing. Our experience provides strong support for the safety and diagnostic utility of ultrasound-guided CNB in the evaluation of salivary gland enlargement or focal abnormalities in individuals with SS.

None of the patients reported complications from the procedure. Notably, CNB avoided an open surgical excision, with its attendant risk of facial nerve injury, fistula formation, sialocele formation, and unacceptable cosmetic deformity (9). A surgical excision of the major salivary gland is also time-consuming and often requires hospital admission and general anesthesia. Surgery may also exacerbate xerostomia in SS patients. In contrast to epithelial salivary gland tumors, the primary treatment of salivary gland lymphoma is not surgical excision, so diagnosis with a minimally invasive procedure can avoid the need for surgery (10,11).

Our CNB diagnostic protocol provided sufficient material to differentiate a range of salivary gland pathologic findings expected in a cohort of patients with SS, namely normal salivary gland tissue, fatty replacement, focal lymphocytic sialadenitis, and more diffuse lymphocytic infiltration, representing either BLEL or MALT lymphoma. The ability to differentiate a range of findings constituted an advantage of CNB (often with FNA) over FNA alone. None of the patients had alternative forms of benign salivary gland inflammation, such as IgG4-related or granulomatous sialadenitis. Ultrasonographic imaging during the biopsy procedure enabled targeting of lesions that were suspicious for lymphoma. These corresponded to palpable masses in 9 patients, but also clinically occult lesions in 5. On ultrasound, lesions that proved to be lymphomatous included mass-like conglomerates of hypoechoic lesions and abnormal intraparotid lymph nodes. Notably, the parenchyma of the parotid and submandibular gland showed numerous hypoechoic lesions in all patients with lymphoma or atypical lymphoid infiltrates. This observation suggests that clinical enlargement of salivary glands in the absence of sonographic abnormality is not suspicious for lymphoma. A larger sample size would be necessary to generalize our experience with this cohort. Similar concerns have been raised by Jousse-Joulin et al (12).

In 3 of our patients with diffuse glandular lymphoid infiltration, a definitive diagnosis of indolent low-grade lymphoma could not be established. In each, flow cytometry demonstrated a monoclonal B cell population, but the histologic findings were not sufficient to allow a definitive diagnosis of lymphoma. Differentiation of BLEL from low-grade lymphoma, most often MALT, can be difficult (13,14). Both are characterized by polymorphous lymphoid infiltrates with an admixture of B and T cells and lymphoepithelial lesions. A monoclonal cell population is characteristic of MALT lymphoma (15), but can also be found in BLEL from SS as well as reactive nodes. The monoclonal cell population may also represent reactive follicular center cells (16). While BLEL is a histologic precursor of MALT lymphoma, evolution to MALT lymphoma is slow and infrequent.

With the understanding that the distinction between a benign and malignant lymphoproliferative process of the salivary gland may not be definitive in patients with SS, the decision about whether to treat is based on a number of factors, including systemic disease activity, symptoms, or cosmetic concerns related to the salivary gland enlargement or mass, and overall health (10,11). Treatment can be effective in reducing salivary gland enlargement and controlling systemic disease activity but is not known to impact the

recurrence rate of MALT lymphoma. Similarly, whether treatment of BLEL prevents or delays the development of frank lymphoma is not known.

The information provided by ultrasound-guided CNB, inclusive of flow cytometry, was actionable in all cases, allowing the practitioner to differentiate a benign salivary gland process from the presence of MALT lymphoma or a possible lymphoproliferative disorder. With respect to the latter 2 diagnoses, the treatment decision is between expectant observation and the institution of B cell–depleting therapy, either alone or with another agent. This decision is made on the basis of the clinical context, and in our patients could be made in conjunction with CNB without the need for excisional biopsy. Note is made of 2 patients who underwent additional biopsies, 1 due to enlarging inguinal nodes (showing MALT lymphoma) 6 months after the parotid gland biopsy, and a second who underwent repeat core biopsies of the same parotid gland mass over the ensuing 17 months, each time showing an atypical lymphoid infiltrate but no frank lymphoma.

We advocate for CNB only in certain clinical situations, namely persistent salivary gland enlargement, a palpable mass, or a sonographic abnormality such as an abnormal intraparotid lymph node or mass-like conglomeration of the hypoechoic lesions characteristic of SS. Given the low grade and indolent behavior of salivary gland lymphoid neoplasms in SS, there must be circumspection in the use of CNB for routine disease monitoring. Early detection of lymphoma through repeated biopsies is unlikely to affect outcome and may engender undue anxiety.

The limitations of our study include our relatively small cohort, which might not have been fully representative of the range of salivary gland lesions that may be encountered in this population. On the other hand, our study is the largest experience as yet reported for CNB in suspected lymphoma in SS and is concordant with what has been reported by others.

In summary, CNB of the salivary glands in patients with SS is a safe and useful technique for evaluation of suspected salivary gland lymphoma, providing sufficient pathologic material to allow for definitive pathologic evaluation and appropriate clinical decision-making, while avoiding the risks of an excisional biopsy.

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. Baer 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. Baer, Fradin.
Acquisition of data. Baer, Grader-Beck, Antiochos, Birnbaum, Fradin.
Analysis and interpretation of data. Baer, Grader-Beck, Fradin.

REFERENCES

1. Witt BL, Schmidt RL. Ultrasound-guided core needle biopsy of salivary gland lesions: a systematic review and meta-analysis. Laryngoscope 2014;124:695–700.
2. Huang YC, Wu CT, Lin G, Chuang WY, Yeow KM, Wan YL. Comparison of ultrasonographically guided fine-needle aspiration and core needle biopsy in the diagnosis of parotid masses. J Clin Ultrasound 2012;40:189–94.
3. Nocturne G, Mariette X. Sjogren syndrome-associated lymphomas: an update on pathogenesis and management. Br J Haematol 2015;168:317–27.
4. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol 2014;32:3059–68.
5. Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome: a consensus and data–driven methodology involving three international patient cohorts. Arthritis Rheumatol 2017;69:35–45.
6. Drylewicz MR, Watkins MP, Shetty AS, Lin MF, Salter A, Bartlett NL, et al. Formulating a treatment plan in suspected lymphoma: ultrasound-guided core needle biopsy versus core needle biopsy and fine-needle aspiration of peripheral lymph nodes. J Ultrasound Med 2019;38:581–6.
7. Jousse-Joulin S, D'Agostino MA, Nicolas C, Naredo E, Ohrndorf S, Backhaus M, et al. Video clip assessment of a salivary gland ultrasound scoring system in Sjogren's syndrome using consensual definitions: an OMERACT ultrasound working group reliability exercise. Ann Rheum Dis 2019;78:967–73.
8. Zenone T. Parotid gland non-Hodgkin lymphoma in primary Sjogren syndrome. Rheumatol Int 2012;32:1387–90.
9. Wyss E, Mueller-Garamvolgyi E, Ghadjar P, Rauch D, Zbaren P, Arnold A. Diagnosis and treatment outcomes for patients with lymphoma of the parotid gland. Laryngoscope 2013; 123:662–9.
10. Voulgarelis M, Ziakas PD, Papageorgiou A, Baimpa E, Tzioufas AG, Moutsopoulos HM. Prognosis and outcome of non-Hodgkin lymphoma in primary Sjogren syndrome. Medicine (Baltimore) 2012; 91:1–9.
11. Pollard RP, Pijpe J, Bootsma H, Spijkervet FK, Kluin PM, Roodenburg JL, et al. Treatment of mucosa-associated lymphoid tissue lymphoma in Sjogren's syndrome: a retrospective clinical study. J Rheumatol 2011;38:2198–208.
12. Jousse-Joulin S, D'Agostino MA, Hocevar A, Naredo E, Terslev L, Ohrndorf S, et al. Could we use salivary gland ultrasonography as a prognostic marker in Sjogren's syndrome? Response to: 'Ultrasonographic damages of major salivary glands are associated with cryoglobulinemic vasculitis and lymphoma in primary Sjogren's syndrome: are the ultrasonographic features of the salivary glands new prognostic markers in Sjogren's syndrome?' by Coiffier et al [letter]. Ann Rheum Dis 2019. Epub ahead of print.
13. Carbone A, Gloghini A, Ferlito A. Pathological features of lymphoid proliferations of the salivary glands: lymphoepithelial sialadenitis versus low-grade B-cell lymphoma of the malt type. Ann Otol Rhinol Laryngol 2000;109 Pt 1:1170–5.
14. Hsi ED, Zukerberg LR, Schnitzer B, Harris NL. Development of extrasalivary gland lymphoma in myoepithelial sialadenitis. Mod Pathol 1995;8:817–24.
15. Stacchini A, Aliberti S, Pacchioni D, Demurtas A, Isolato G, Gazzera C, et al. Flow cytometry significantly improves the diagnostic value of fine needle aspiration cytology of lymphoproliferative lesions of salivary glands. Cytopathology 2014;25:231–40.
16. Kussick SJ, Kalnoski M, Braziel RM, Wood BL. Prominent clonal B-cell populations identified by flow cytometry in histologically reactive lymphoid proliferations. Am J Clin Pathol 2004;121:464–72.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 856–860
DOI 10.1002/acr.24174

The Phenotype of Axial Spondyloarthritis: Is It Dependent on HLA–B27 Status?

Laura C. Coates,[1] Xenofon Baraliakos,[2] Francisco J. Blanco,[3] Elena Alonso Blanco-Morales,[3] Jurgen Braun,[2] Vinod Chandran,[4] Jose Luis Fernandez-Sueiro,[†] Oliver FitzGerald,[5] Phil Gallagher,[5] Dafna D. Gladman,[4] Elena Gubar,[6] Tatiana Korotaeva,[6] Elena Loginova,[6] Ennio Lubrano,[7] Juan Mulero,[8] Jose Pinto-Tasende,[3] Ruben Queiro,[9] Jesús Sanz Sanz,[8] Agnes Szentpetery,[10] and Philip S. Helliwell[11]

Objective. To describe the radiographic phenotype of axial spondyloarthritis (SpA) according to the presence of HLA–B27.

Methods. An international collaboration compared the radiographic phenotype of axial SpA according to HLA–B27 status. Patients with ankylosing spondylitis (AS) and axial psoriatic arthritis (PsA) were collected. Radiographs were read centrally, blinded to clinical details. The symmetry of the sacroiliac joints and lumbar syndesmophytes and the morphology of syndesmophytes (typical marginal versus atypical chunky), together with the modified Stoke Ankylosing Spondylitis Spine Score and the Psoriatic Arthritis Spondylitis Radiographic Index, were recorded.

Results. A total of 244 patients with PsA and 198 patients with AS were included. In PsA, 60 patients (25%) were HLA–B27 positive while in AS, 148 patients (75%) were HLA–B27 positive. Patients with HLA–B27 were younger and more often male and had a longer duration of disease. In multivariable logistic regression, HLA–B27 was significantly associated with syndesmophyte symmetry (odds ratio [OR] 3.02 [95% confidence interval (95% CI) 1.38, 6.61]) and marginal syndesmophytes (OR 1.97 [95% CI 1.16, 3.36]) but not with sacroiliac symmetry. Mean radiographic scores were higher for patients with HLA–B27.

Conclusion. Patients with axial SpA who are positive for HLA–B27 have more severe radiographic damage, more marginal syndesmophytes, and more frequent syndesmophyte symmetry compared to patients who are negative for HLA–B27.

INTRODUCTION

Axial spondyloarthritis (axial SpA) is an inflammatory disease of the spine and sacroiliac joints that leads to new bone formation and has the potential to cause total ankylosis of the spine. Ankylosing spondylitis (AS) represents the classical manifestation of axial SpA and was the hallmark clinical manifestation of SpA, first described in detail by Moll et al (1). Psoriatic arthritis (PsA) is a common form of inflammatory arthritis affecting between 15% and 30% of people with psoriasis and is a member of the SpA group of conditions. The most common phenotype of PsA is predominant peripheral arthritis (2,3), but up to 50% of patients with PsA develop inflammation in their axial skeleton (axial PsA), and a few (approximately 5%) have isolated axial inflammation (4).

Although axial involvement in PsA can be indistinguishable from axial disease in AS, it can also differ in several respects, raising the question of whether axial PsA and AS, with or without psoriasis, are different clinical presentations of the same disease,

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Supported by a starter grant for clinical lecturers from the Academy of Medical Sciences and by the NIHR Leeds and Oxford Biomedical Research Centre. Dr. Coates' work was supported by an NIHR Clinician Scientist award.

[1]Laura C. Coates, MBChB, PhD: Leeds Institute of Rheumatic and Musculoskeletal Medicine, Leeds, and University of Oxford, Oxford, UK; [2]Xenofon Baraliakos, MD, PhD, Jurgen Braun, MD: Rheumazentrum Ruhrgebiet, Ruhr-Universität, Bochum, Germany; [3]Francisco J. Blanco, MD, PhD, Elena Alonso Blanco-Morales, MD, Jose Pinto-Tasende, MD, PhD: Complexo Hospitilario, Universitario de A Coruña, A Coruña, Spain; [4]Vinod Chandran, MBBS, MD, DM, PhD, Dafna D. Gladman, MD, FRCPC: University of Toronto, Toronto, Ontario, Canada; [5]Oliver FitzGerald, MD, Phil Gallagher, RGN: Saint Vincent's University Hospital, Dublin, Ireland; [6]Elena Gubar, PhD, VA, Tatiana Korotaeva, MD, PhD, VA, Elena Loginova, PhD, VA: Nasonova Research Institute of Rheumatology, Moscow, Russia; [7]Ennio Lubrano, MD, PhD: Università degli Studi del Molise, Campobasso, Italy; [8]Juan Mulero, MD, Jesús Sanz Sanz, MD: Hospital Puerta de Hierro, Majadahonda, Madrid, Spain; [9]Ruben Queiro, MD, PhD: Hospital Universitario Central de Asturias, Oviedo, Spain; [10]Agnes Szentpetery, MD, PhD: Saint Vincent's University Hospital, Dublin, Ireland, and Uppsala University Hospital, Uppsala, Sweden; [11]Philip S. Helliwell, MA, MD: Leeds Institute of Rheumatic and Musculoskeletal Medicine and University of Leeds, Leeds, UK.

[†]Dr. Fernandez-Sueiro is deceased.

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

Address correspondence to Philip S. Helliwell, MA, MD, Chapel Allerton Hospital, Chapeltown Road, Leeds LS7 4SA, UK. Email: p.helliwell@leeds.ac.uk.

Submitted for publication June 10, 2019; accepted in revised form February 18, 2020.

SIGNIFICANCE & INNOVATIONS

- Patients with axial spondyloarthritis (SpA) who are positive for HLA–B27 have higher levels of radiographic damage, more symmetry, and more marginal syndesmophytes.
- Most patients with axial psoriatic arthritis are HLA–B27 negative and less frequently have sacroiliac joint involvement.
- Existing classification criteria for axial SpA may not be applicable to axial psoriatic arthritis.

axial SpA, or whether they are separate diseases that have overlapping features (5). Recent clinical (6) and genetic (7,8) studies have shown that axial PsA is nonhomogeneous, based on the presence of HLA–B27, a result confirmed for early axial SpA in the Devenir des Spondylarthropathies Indifférenciées Récentes cohort (9).

Our study hypothesis was that the radiographic phenotype of patients with axial SpA depends on the presence of HLA–B27. We hypothesized that HLA–B27 positivity is associated with a more severe, classical AS phenotype: these patients have a more symmetrical appearance on the radiographs of both the spine and the sacroiliac joints, and manifest classical syndesmophyte morphology with smooth, contiguous calcification between adjacent vertebrae. In contrast, patients who are negative for HLA–B27 will represent an alternative phenotype, with less radiographic severity, less involvement of the sacroiliac joints and spine, less symmetry, and different morphology of syndesmophytes, with unusual-shaped, bulkier, nonmarginal syndesmophytes (10–12). To achieve phenotypic diversity, we studied patients with PsA and axial involvement (a group of patients recognized to have less frequent carriage of HLA–B27), and AS, with patients drawn from a number of geographically diverse populations.

MATERIALS AND METHODS

Patients. Cross-sectional clinical, radiographic, and laboratory data from several cohorts in Ireland, Canada, Italy, Germany, Russia, and Spain were included. All sites have clinics dedicated to axial SpA, with both PsA and AS. The data were extracted from existing databases and digital film archives. Patient consent was not sought specifically for this study although consent was collected within each existing cohort to study both clinical and radiographic data. Formal ethics review was not obtained. The inclusion criteria were as follows: 1) age ≥18 years; 2) either a clinical diagnosis of PsA and fulfillment of the Classification of Psoriatic Arthritis criteria with a physician diagnosis of axial involvement, or a clinical diagnosis of AS and fulfillment of the modified New York criteria; 3) HLA–B27 status available; and 4) plain radiographs of sacroiliac joints and lumbar and cervical spines within the last 5 years.

Minimal clinical data were collected, including basic demographic data, a recent patient-completed disease activity measure (the Bath Ankylosing Spondylitis Disease Activity Index [BASDAI] [13]), and a recent measure of C-reactive protein (CRP) level. In addition, the presence/absence of diabetes mellitus was recorded, because there is an association between diabetes mellitus and diffuse idiopathic skeletal hyperostosis (DISH), the radiographic appearance of which may make interpretation of syndesmophyte morphology more difficult (14).

Radiographs. The majority of the images (>90%) obtained were in the DICOM format; the rest were JPEG images. The images were read by consensus by 2 observers (LCC and PSH), blind to diagnosis. The lateral spinal images were scored using the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) (15) and Psoriatic Arthritis Spondylitis Radiology Index (PASRI) (16) scoring systems. The mSASSS scores the corners of the vertebral bodies from the lower border of C2 to the upper border of T1, and from the lower border of T12 to the upper border of S1. At each vertebral corner, scores range from 0 to 3, thus giving a total score range of 0–72. The PASRI scores the vertebral bodies in a similar way to the mSASSS, but in addition, scores the zygo-apophyseal joints at C2/C3, C3/C4, and C4/C5 for fusion and scores the sacroiliac joints using the modified New York criteria (17), with a score range for each joint of 0–4. These methods have been proven reliable in both AS and PsA (18).

Using further review of the lateral alongside the anteroposterior images, other features were recorded at each vertebral level: syndesmophyte morphology (marginal or nonmarginal), Andersson lesions, zygo-apophyseal joint fusion in the cervical spine, and paravertebral ossification (ossification adjacent to, but separate from, a vertebral body usually contiguous with a syndesmophyte) at each level. The symmetry of sacroiliac joints was defined as a 2-point difference in scores between each side. The symmetry of syndesmophytes was assessed on anteroposterior views and was determined from the ratio of the number of matched pairs of syndesmophytes to the total number of syndesmophytes; a ratio of 0.5 or above was deemed to indicate symmetry, as previously defined for peripheral joint involvement (19).

Statistical analysis. Summary statistics, according to data, are presented with appropriate univariate statistical tests. Logistic regression models were used to investigate the predictors of symmetry and syndesmophyte morphology using binary multivariable logistic regression models, entering all independent variables together, assessing goodness of fit by the Hosmer-Lemeshow method and the percentage of accurate prediction. Independent variables were age, sex, HLA–B27 status, duration of disease, and diabetes mellitus status. Using the available data (radiographic scoring of sacroiliac joints, HLA–B27 status, diagnosis of psoriasis, the presence of inflammatory back pain, and

Table 1. Demographic details of the cohort, fulfillment of ASAS criteria, and radiographic damage scores*

	HLA–B27 positive (n = 208)	HLA–B27 negative (n = 234)	Difference between B27+ and B27– (continuous data) and odds ratios (categorical data)	*P*
Age, mean ± SD years	49.1 ± 14.2	53.8 ± 13.8	–4.7 (–7.4, –2.1)†	< 0.0001
Male	152 (73)	138 (59)	1.9 (1.3, 2.8)‡	0.002
Duration of disease, mean ± SD years	13.6 ± 11.9	11.0 ± 10.2	2.6 (0.5, 4.7)†	0.02
Fulfills clinical arm, ASAS criteria	68 (33)	0	NA	< 0.0001
Fulfills radiographic arm, ASAS criteria	177 (85)	149 (64)	3.3 (2.1, 5.2)‡	< 0.0001
mSASSS score, median (range)	6 (0–72)	2 (0–72)	0.5 (0–3)§	0.04
PASRI score, median (range)	12 (0–71)	6 (0–71)	5 (3–7)§	< 0.0001
BASDAI, mean ± SD	4.1 ± 2.0	3.5 ± 2.4	0.6 (0.2, 1.1)†	0.009

* Values are the number (%) unless indicated otherwise. ASAS = Assessment of SpondyloArthritis international Society; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; mSASSS = Stoke Ankylosing Spondylitis Spine Score; NA = not applicable; PASRI = Psoriatic Arthritis Spondylitis Radiology Index.
† Mean (95% confidence interval [95% CI]).
‡ Odds ratio (95% CI).
§ Median difference (range).

elevated CRP level), the Assessment of SpondyloArthritis international Society (ASAS) criteria for axial SpA (20) were applied.

RESULTS

A total of 8 sites contributed data on 244 patients with PsA and 198 patients with AS. In patients with PsA, 60 (25%) were HLA–B27 positive, while in AS, 148 (75%) were HLA–B27 positive. Patients with HLA–B27 were younger, were more often male, and had a longer duration of disease. Patients with HLA–B27 had higher BASDAI, mSASSS, and PASRI scores (Table 1). A total of 54 patients in the PsA group did not have radiographic sacroiliitis. Of these, 33 also did not have syndesmophytes. In the AS group, 13 patients did not have any syndesmophytes. In total, 338 patients met either the clinical or radiographic arm of the axial SpA criteria (n = 270 met the radiographic arm, n = 12 met the clinical arm, and n = 56 met both). Fulfillment of these criteria was significantly higher for those with HLA–B27.

Sacroiliac joint involvement. Patients who were negative for HLA–B27 were more likely to have bilateral normal sacroiliac joints (grade 0 or 1), and less likely to have bilateral grade 4 sacroiliac joints (Table 2). However, there was no difference in symmetry of the sacroiliac joints according to HLA–B27 status, after exclusion of cases with bilateral normal sacroiliac joints. Multivariable logistic regression assessing predictors of sacroiliac symmetry found age to be the only significant predictor (odds ratio [OR] 1.04 [95% confidence interval (95% CI) 1.01, 1.06]) (Table 3).

Spinal involvement. Not all cases had syndesmophytes on the anteroposterior view of the lumbar spine, but nevertheless, there was a clear difference between groups in terms of syndesmophyte symmetry, and in the presence of marginal syndesmophytes, particularly in the lumbar spine, where these were more frequently seen in those who were HLA–B27 positive (Table 2). There were no differences in nonmarginal syndesmophytes according to HLA–B27 status.

The only predictor of syndesmophyte symmetry was HLA–B27 positivity (OR 3.02 [95% CI 1.38, 6.61]). The presence of marginal syndesmophytes showed a significant relationship with age (OR 1.08 [95% CI 1.05, 1.10]), HLA–B27 status (OR 1.97 [95% CI 1.16, 3.36]), and male sex (OR 1.66 [95% CI 1.04, 2.66]). For nonmarginal syndesmophytes, only age (OR 1.05 [95% CI 1.03, 1.07]) and male sex (OR 2.55 [95% CI 1.46, 4.64]) were significant predictors (Table 3).

Table 2. Radiographic phenotype according to HLA–B27 status*

	HLA–B27 positive (n = 208)	HLA–B27 negative (n = 234)	OR (95% CI)	*P*, two-way
Bilateral normal sacroiliac joints: grade 0 or 1	11 (5)	39 (17)	0.3 (0.1, 0.7)	< 0.0001
Bilateral grade 4 sacroiliac joints	82 (39)	37 (16)	3.5 (2.2, 5.4)	< 0.0001
Symmetry at sacroiliac joint (excludes bilateral zero), no./total no. (%)	175/193 (91)	169/194 (87)	1.4 (0.8, 2.7)	NS
Symmetry syndesmophytes (lumbar spine), no./total no. (%)	88/113 (78)	50/86 (58)	2.5 (1.4, 4.7)	0.003
Marginal syndesmophytes	128 (62)	119 (51)	1.5 (1.1, 2.3)	0.02
Cervical	103 (50)	97 (42)	1.4 (0.9, 2.0)	NS
Lumbar	95 (46)	73 (31)	1.9 (1.3, 2.7)	0.002
Nonmarginal syndesmophytes	46 (22)	53 (23)	1.0 (0.6, 1.5)	NS
Cervical	28 (14)	37 (16)	0.8 (0.5, 1.4)	NS
Lumbar	28 (14)	31 (13)	1.0 (0.6, 1.7)	NS

* Values are the number (%) unless indicated otherwise. 95% CI = 95% confidence interval; NS = not significant; OR = odds ratio.

Table 3. Predictors of radiographic phenotype (symmetry and syndesmophyte morphology)*

Multivariable logistic regression analysis	OR (95% CI)	*P*
Sacroiliac joint symmetry†		
Male	0.81 (0.55, 2.14)	NS
HLA–B27 positive	1.54 (0.72, 3.28)	NS
Age	1.04 (1.01, 1.06)	0.02‡
Years of diagnosis	1.00 (0.96, 1.04)	NS
Diagnosis of PsA	0.75 (0.36, 1.59)	NS
Syndesmophyte symmetry§		
Male	0.89 (0.41, 1.95)	NS
HLA–B27 positive	3.02 (1.38, 6.61)	0.006‡
Age	1.03 (0.99, 1.06)	NS
Years of diagnosis	1.00 (0.97, 1.03)	NS
Diagnosis of PsA	0.80 (0.39, 1.67)	NS
Marginal syndesmophytes¶		
Male	1.66 (1.04, 2.66)	0.035‡
HLA–B27 positive	1.97 (1.16, 3.36)	0.013‡
Age	1.08 (1.05, 1.10)	<0.0001‡
Years of diagnosis	1.02 (0.99, 1.051.64)	NS
Diagnosis of PsA	0.80 (0.47, 1.36)	NS
Nonmarginal syndesmophytes#		
Male	2.55 (1.46, 4.64)	0.001‡
Presence of diabetes mellitus	1.65 (0.73, 3.76)	NS
HLA–B27 positive	1.20 (0.67, 2.17)	NS
Age	1.05 (1.03, 1.07)	<0.0001‡
Years of diagnosis	1.00 (0.98, 1.03)	NS
Diagnosis of PsA	1.17 (0.66, 2.10)	NS

* 95% CI = 95% confidence interval; NS = not significant; OR = odds ratio; PsA = psoriatic arthritis.
† Chi-square = 7.9, *P* = 0.45, overall correct prediction = 88.8%.
‡ Statistically significant.
§ Chi-square = 6.6, *P* = 0.59; overall correct prediction = 69.5%.
¶ Chi-square = 8.77, *P* = 0.36; overall correct prediction = 70.2%.
# Chi-square = 6.6, *P* = 0.58; overall correct prediction = 78.0%.

DISCUSSION

In this observational cross-sectional study, differences in radiographic phenotype according to HLA–B27 status were largely as we had hypothesized. Thus, the patients who were HLA–B27 positive had more severe radiographic damage, as measured by mSASSS and PASRI, more bilateral fused sacroiliac joints, more typical marginal syndesmophytes, and more symmetry in the spine. However, this study has shown no difference in sacroiliac symmetry, and no difference in nonmarginal syndesmophytes, according to HLA–B27 status.

The strengths of this study are the large, international sample size, with a mixed population of axial PsA and AS, and the blinded reading of the radiographs. The readers were therefore not subject to bias due to knowledge of diagnosis that may have influenced the results, particularly with regard to subjective interpretations, such as syndesmophyte morphology. This is the first study, to our knowledge, to describe the phenotype of established axial SpA according to HLA–B27 carriage, using a mixed population in which HLA–B27 carriage varied markedly. Unlike studies by Jadon et al (6) and Haroon et al (8), we did not focus specifically on disease status (axial PsA compared to AS), hypothesizing that HLA–B27 status was the main influence of radiographic phenotype, a result largely confirmed by this study.

This study has some limitations. We collected a large number of patients with axial PsA and AS from a number of cohorts in Europe and North America. Disease groups were not matched for age, sex, and duration of disease, all of which may influence the phenotype (Table 3). Further, central reading was done by consensus, not independently, and the recognition of syndesmophyte morphology was subjective, because no standard definitions are available. Because these participants were collected from existing cohorts, the study did not attempt to standardize case definition where currently no accepted criteria for axial PsA are available. Some differences may be due to case selection, because contributors possibly handpicked the cases for inclusion.

The limitations noted above may reflect the discordance of results between this study and previous studies, particularly with respect to sacroiliac symmetry. In a similar, but single-center study published in 1998, the proportion of symmetrical sacroiliitis in cases of AS and PsA were 0.85 and 0.74, respectively (11), compared to 0.88 for both conditions in the current study. Evaluation of the PsA cohort in Dublin found a proportion of symmetrical sacroiliitis at 0.27, but this cohort evaluated all patients with PsA, rather than only those with physician-diagnosed axial disease. In the latter study, asymmetry was associated with HLA–B*08 and symmetry with HLA–B27 (8).

The patients included in this study had already been diagnosed with axial SpA by the investigators, so presumably cases of DISH had been excluded prior to referral. However, differentiating between nonmarginal syndesmophytes and the appearances of DISH can be difficult, especially where the sacroiliac joints appear normal. DISH may coexist with axial SpA, and DISH may be found in approximately 8% of patients with PsA, according to 1 study (21). We acknowledge that some cases of DISH may have been included inadvertently but were unlikely to influence the major findings in relation to HLA–B27 status and phenotype.

This study and others have implications for the diagnosis and classification of spondylitis in people with psoriasis. In AS, the prevalence of the major histocompatibility complex class I allele HLA–B27 is 85–90%, but in PsA the prevalence is much lower at 20–50% (5,22,23), so the axial phenotype would be expected to differ between AS and PsA. The ASAS classification criteria included patients with concomitant psoriasis, and thus, by definition, psoriatic spondylitis, although the majority of patients presumably had nonpsoriatic axial SpA. The ASAS criteria include a clinical arm, which is dependent on HLA–B27 status, and a radiographic arm, which includes imaging evidence of sacroiliitis (20). Given the lower frequency of HLA–B27 in the spondylitis associated with psoriasis, and the lower frequency of sacroiliac involvement, patients with PsA are less likely to fulfill both the clinical and the imaging arms of the classification criteria. Developing an alternative clinical and radiologic definition of axial PsA may be

necessary for classification. If this development were to be done, an entirely new classification study would be required, including cases of psoriatic spondylitis and classical AS, selecting consecutive cases attending outpatient clinics.

In summary, this analysis suggests less difference in radiographic phenotype between AS and axial PsA than previously found but emphasizes the importance of HLA–B27 status in severity and the phenotypic expression of disease radiographically. Future studies, including those assessing classification criteria, should allow for the disparity in HLA–B27 frequency between AS and axial PsA.

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. Helliwell 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. Coates, Helliwell.

Acquisition of data. Coates, Baraliakos, Blanco, Blanco-Morales, Braun, Chandran, FitzGerald, Gallagher, Gladman, Gubar, Korotaeva, Loginova, Lubrano, Mulero, Pinto-Tasende, Queiro, Sanz Sanz, Szentpetery, Helliwell.

Analysis and interpretation of data. Coates, Helliwell.

REFERENCES

1. Moll JM, Haslock I, Macrae I, Wright V. Associations between ankylosing spondylitis, psoriatic arthritis, Reiters' disease, the intestinal arthropathies, and Behçet's syndrome. Medicine (Baltimore) 1974;53:343–64.
2. 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.
3. Taylor WJ, Zmierczak HG, Helliwell PS. Problems with the definition of axial and peripheral disease patterns in psoriatic arthritis. J Rheumatol 2005;32:974–7.
4. Chandran V, Barrett J, Schentag CT, Farewell VT, Gladman DD. Axial psoriatic arthritis: update on a longterm prospective study. J Rheumatol 2009;36:2744–50.
5. Feld J, Chandran V, Haroon N, Inman R, Gladman D. Axial disease in psoriatic arthritis and ankylosing spondylitis: a critical comparison. Nat Rev Rheumatol 2018;14:363–71.
6. Jadon DR, Sengupta R, Nightingale A, Lindsay M, Korendowych E, Robinson G, et al. Axial disease in psoriatic arthritis study: defining the clinical and radiographic phenotype of psoriatic spondyloarthritis. Ann Rheum Dis 2017;76:701–7.
7. Haroon M, Winchester R, Giles JT, Heffernan E, FitzGerald O. Certain class I HLA alleles and haplotypes implicated in susceptibility play a role in determining specific features of the psoriatic arthritis phenotype. Ann Rheum Dis 2016;75:155–62.
8. Haroon M, Winchester R, Giles JT, Heffernan E, Fitzgerald O. Clinical and genetic associations of radiographic sacroiliitis and its different patterns in psoriatic arthritis. Clin Exp Rheumatol 2017;35:270–6.
9. Chung HY, Machado P, van der Heijde D, D'Agostino MA, Dougados M. HLA–B27 positive patients differ from HLA–B27 negative patients in clinical presentation and imaging: results from the DESIR cohort of patients with recent onset axial spondyloarthritis. Ann Rheum Dis 2011;70:1930–6.
10. Bywaters EG, Dixon AS. Paravertebral ossification in psoriatic arthritis. Ann Rheum Dis 1965;24:313–31.
11. Helliwell PS, Hickling P, Wright V. Do the radiological changes of classic ankylosing spondylitis differ from the changes found in the spondylitis associated with inflammatory bowel disease, psoriasis, and reactive arthritis? Ann Rheum Dis 1998;57:135–40.
12. McEwen C, DiTata D, Lingg C, Porini A, Good A, Rankin T. Ankylosing spondylitis and spondylitis accompanying ulcerative colitis, regional enteritis, psoriasis and Reiter's disease: a comparative study. Arthritis Rheum 1971;14:291–318.
13. 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.
14. Olivieri I, D'Angelo S, Palazzi C, Padula A. Spondyloarthritis and diffuse idiopathic skeletal hyperostosis: two different diseases that continue to intersect. J Rheumatol 2013;40:1251–3.
15. Creemers MC, Franssen MJ, van't Hof MA, Gribnau FW, van de Putte LB, van Riel PL. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis 2005;64:127–9.
16. Lubrano E, Marchesoni A, Olivieri I, D'Angelo S, Spadaro A, Parsons WJ, et al. Psoriatic arthritis spondylitis radiology index: a modified index for radiologic assessment of axial involvement in psoriatic arthritis. J Rheumatol 2009;36:1006–11.
17. 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.
18. Biagioni BJ, Gladman DD, Cook RJ, Eder L, Wakhlu A, Shen H, et al. Reliability of radiographic scoring methods in axial psoriatic arthritis. Arthritis Care Res (Hoboken) 2014;66:1417–22.
19. Helliwell PS, Hetthen J, Sokoll K, Green M, Marchesoni A, Lubrano E, et al. Joint symmetry in early and late rheumatoid and psoriatic arthritis: comparison with a mathematical model. Arthritis Rheum 2000;43:865–71.
20. Rudwaleit M, Landewe R, van der Heijde D, Listing J, Brandt J, Braun J, et al. The development of Assessment of SpondyloArthritis International Society classification criteria for axial spondyloarthritis (part I): classification of paper patients by expert opinion including uncertainty appraisal. Ann Rheum Dis 2009;68:770–6.
21. Haddad A, Thavaneswaran A, Toloza S, Chandran V, Gladman DD. Diffuse idiopathic skeletal hyperostosis in psoriatic arthritis. J Rheumatol 2013;40:1367–73.
22. Chandran V, Tolusso DC, Cook RJ, Gladman DD. Risk factors for axial inflammatory arthritis in patients with psoriatic arthritis. J Rheumatol 2010;37:809–15.
23. Helliwell PS, Wright V. Psoriatic arthritis: clinical features. In: Klippel JH, Dieppe PA, editors. Rheumatology. 2nd ed. London: Mosby; 1998. p. 6.21.1–6.21.8.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 861–872
DOI 10.1002/acr.24184

American College of Rheumatology
Empowering Rheumatology Professionals

Tumor Necrosis Factor Inhibitor Dose Reduction for Axial Spondyloarthritis: A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Daeria O. Lawson,[1] Maria Eraso,[2] Lawrence Mbuagbaw,[3] Marianinha Joanes,[4] Theresa Aves,[5] Alvin Leenus,[6] Ahmed Omar,[2] and Robert D. Inman[7]

Objective. The present study was undertaken to investigate the effectiveness and safety of dose reduction of tumor necrosis factor inhibitor (TNFi) therapy in the treatment of axial spondyloarthritis (SpA) compared to usual care.

Methods. We searched the Cochrane Central Register of Controlled Trials, Embase, Medline, and trial registries. We screened, extracted data, and assessed risk of bias in duplicate. Data were pooled using random-effects models; subgroup analyses were performed for type of TNFi, prior TNFi exposure, and follow-up duration. Outcomes of interest were Assessment of SpondyloArthritis international Society (ASAS) response and remission criteria, disease activity, relapse, and safety.

Results. We included 6 randomized trials with 747 participants (442 with ankylosing spondylitis and 305 with nonradiographic axial SpA). Compared to the standard dose, there were fewer events with the reduced dose for the ASAS criteria for 40% improvement (risk ratio [RR] 0.62 [95% confidence interval (95% CI) 0.49, 0.78]) and for ASAS partial remission (RR 0.17 [95% CI 0.06, 0.46]). There was a mean increase in the Bath Ankylosing Spondylitis Disease Activity Index score (mean difference [MD] 0.35 [95% CI 0.10, 0.60]) and no difference in C-reactive protein levels (MD 0.16 [95% CI –0.76, 1.07]) with the reduced dose. There were more disease flares/relapses (RR 1.73 [95% CI 1.32, 2.27]) with the reduced dose. There were no differences in infection rates (incidence rate ratio [IRR] 0.98 [95% CI 0.76, 1.25]) or injection/infusion reactions (IRR 0.71 [95% CI 0.42, 1.19]).

Conclusion. Patients with axial SpA may experience little to no clinical benefit from reduction of TNFi therapy. Maintaining the standard dose probably improves the sustained effect on disease activity and helps to prevent disease flare.

INTRODUCTION

Axial spondyloarthritis (SpA) is a chronic, inflammatory arthritis characterized by spinal stiffness, joint inflammation, pain, and decreased spinal mobility and physical function (1,2). For inclusion into research trials (i.e., not for diagnostic purposes), axial SpA can be identified through several standardized criteria, including the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for clinical abnormalities (e.g., inflammatory back pain) (1). Axial SpA can be further categorized into 2 disease forms: nonradiographic axial SpA and ankylosing spondylitis (AS). There is currently no cure for axial SpA; however, it is commonly treated with biologic therapy to manage symptoms, to modulate disease progression, and to minimize associated damaging inflammatory side effects (3,4). Tumor necrosis factor inhibitors (TNFi) are biologic agents that target TNF, a proinflammatory molecule implicated in spondyloarthritis pathogenesis. These drugs have shown significant sustained clinical improvement in axial SpA and are introduced in patients with axial disease or as the next line of treatment after inadequate response to nonsteroidal antiinflammatory drugs (NSAIDs) (1,2,5,6). TNFi modulate immune response, and treatment entails some degree of immunosuppression. This

Ms Lawson's work was supported in part through the Ontario Drug Policy Research Network Student Training Program.

[1]Daeria O. Lawson, HBSc: McMaster University, Hamilton, and Toronto Western Hospital, Toronto, Ontario, Canada; [2]Maria Eraso, MD, Ahmed Omar, MBBCh, MSc: Toronto Western Hospital, Toronto, Ontario, Canada; [3]Lawrence Mbuagbaw, MD, MPH, PhD: McMaster University and St. Joseph's Healthcare Hamilton, Hamilton, Ontario, Canada; [4]Marianinha Joanes, MD: Boston University School of Medicine, Boston, Massachusetts; [5]Theresa Aves, MSc: St. Michael's Hospital, Toronto, Ontario, Canada; [6]Alvin Leenus, BHSc: McMaster University, Hamilton, Ontario, Canada; [7]Robert D. Inman, MD: Toronto Western Hospital and University of Toronto, Toronto, Ontario, Canada.

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

Address correspondence to Robert D. Inman, MD, Division of Rheumatology, Toronto Western Hospital, 399 Bathurst Street, Toronto, Ontario M5T 2S8, Canada. Email: robert.inman@uhn.ca.

Submitted for publication July 29, 2019; accepted in revised form March 3, 2020.

SIGNIFICANCE & INNOVATIONS

- Individuals living with axial spondyloarthritis may consider the reduction of tumor necrosis factor inhibitor (TNFi) therapy at some point during their course of treatment.
- A meta-analysis of randomized controlled trials (RCTs) was completed to bring together high-quality evidence and determine the safety and efficacy of reducing the standard TNFi dose in regard to disease activity and symptom severity.
- Data from 6 RCTs suggest that, in comparison to dose reduction strategies or complete withdrawal of treatment, standard TNFi doses promote sustained improvement in disease activity and help to prevent disease relapse.
- There should be further study of different TNFi dose reduction regimens to evaluate the potential for a dose response or differences in treatment effect depending on the type of TNFi that is reduced.

can result in higher rates of infections and other potential long-term side effects (5,7,8).

However, with sustained efficacy in patients achieving remission in disease, it is possible to wean patients off TNFi (9). Dose reduction of TNFi has been applied using various approaches, including step-down, extending regular intervals of drug administration, and full withdrawal; however, there is currently no standardized strategy for this (6,9,10). The understanding of the underlying axial SpA disease activity such as inflammatory changes and radiographic progression when patients are tapered off of TNFi is limited, and these strategies require close monitoring for indicators of disease relapse (6,10). Two systematic reviews broadly examined the literature up to 2014 addressing the effects of TNFi dose reduction or discontinuation in rheumatic diseases including axial SpA. In both reviews, based largely on evidence from nonrandomized studies, it was suggested that discontinuation of the standard of care TNFi regimens tends to result in disease flares and increased disease activity (10,11). However, some randomized controlled trials (RCTs) and cohort studies incorporating dose reduction regimens in patients with both active and stable disease have concluded that with cautious monitoring patients can maintain remission and low disease activity (11–17). It still remains unclear in which patients it is appropriate to implement dose reduction strategies (if at all) and how to incorporate safety checkpoints for disease flare (10).

Meta-analyses of RCTs can pave the way for a better understanding and interpretation of study findings on the safety and efficacy of TNFi dose reduction as well as encourage the development of more succinct recommendations for safely withdrawing TNFi. The aim of this systematic review and meta-analysis was to provide a summary of the current evidence and to determine the effects of reduced or discontinued TNFi therapy as compared to standard doses in adult patients with axial SpA.

MATERIALS AND METHODS

Our results are reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (18), and the protocol has been registered in Prospero (registration number CRD42018091146) (19). Research ethics approval was not required since we incorporated only secondary participant data.

Eligible participants and interventions. Adults with a diagnosis of axial SpA (i.e., AS or nonradiographic axial SpA), based on the ASAS classification criteria, were eligible. RCTs evaluating any of the 5 reference TNFi currently approved for treatment in axial SpA (i.e., adalimumab, etanercept, golimumab, certolizumab, and infliximab) were included. Studies investigating TNFi biosimilars (i.e., non-originator drug versions) were excluded, as these are still an emerging class of drugs. We planned to evaluate different strategies for treatment reduction. Dose reduction, dose tapering, extending administration intervals, and complete discontinuation of the drug were all eligible interventions for inclusion. Studies that compared the reduced dose (intervention) to the maintenance of the standard dose (control) for the same TNFi were included. Intervention arms with cointerventions (e.g., TNFi with methotrexate) or placebo as the control were excluded. Studies retrospectively investigating TNFi discontinuation that was not due to a predetermined treatment strategy or where TNFi withdrawal was mandated by adverse events or intercurrent events such as infection, pregnancy, or surgery were excluded.

Outcomes. The prespecified outcomes of interest were as follows: 1) efficacy of dose reduction as measured by a) symptom severity (using the Ankylosing Spondylitis Disease Activity Score [ASDAS] and ASAS criteria for 40% improvement [ASAS40]) and b) disease activity (using Bath Ankylosing Spondylitis Disease Activity Index [BASDAI] score and C-reactive protein [CRP] level in milligrams per liter) (2,20–25); 2) ASAS partial remission (26,27); 3) relapse; 4) safety (e.g., rate of admissions, adverse events, or infections); 5) quality of life (QoL).

Search methods. We conducted initial hand searches in PubMed and checked reference lists of relevant articles to identify potential studies for inclusion. We searched Medline and Embase as a combination search with deduplication via the Ovid platform and the Cochrane Central Register of Controlled Trials from inception to May 2019. There were no restrictions based on language, study design, or publication status. Clinical experts were contacted for additional unpublished data. Ongoing and unpublished trials were identified through the ClinicalTrials.gov and World Health Organization International Clinical Trials Registry Platform registries. The full search strategy is presented in Supplementary Tables 1–4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract.

Study selection. Title and abstract screening was completed independently and in duplicate (DOL and ME, MJ, AL, or TA) through the online review application Rayyan (28). The screening form was piloted (50 articles in duplicate) and revised accordingly to update any items lacking clarity. Authors independently examined the retrieved articles for relevance and excluded articles not meeting the inclusion criteria. Full texts of all relevant articles were assessed independently and in duplicate (DOL and ME) using an Excel spreadsheet (Microsoft) after piloting the screening form. Conflicts were resolved through discussion.

Assessment of risk of bias. Two review authors (DOL and ME or TA) independently assessed the risk of bias in studies and for each outcome using the Cochrane Collaboration's tool for assessing risk of bias in RCTs (29). Additionally, we had planned to use funnel plots to assess reporting bias. Disagreements between reviewers were resolved through discussion, and a third reviewer (LM or RDI) was consulted if no consensus could be reached.

Data collection and synthesis. Two review authors (DOL and ME) independently extracted data into a spreadsheet. Only published, intent-to-treat data were extracted from each study where available. The weighted mean difference (MD) was computed for continuous outcomes (e.g., BASDAI score) and risk ratios (RRs) for dichotomous outcomes (e.g., relapse). The natural log of the incidence rate ratio (IRR) and its SE were computed for safety outcomes. We obtained raw data from graphs and converted data from medians, proportions, and SE to SD as needed when not directly reported (30,31). Potentially skewed data (e.g., SD > mean) were not pooled, and no value imputations or assumptions were made for missing data. Pooled estimates were reported with 95% confidence intervals (95% CIs). All analyses, based on the DerSimonian-Laird random-effects model, were completed in Stata, version 15.1, using the Metan package (32). We assessed statistical heterogeneity using the chi-square test, and the I^2 statistic was used to determine if pooling studies was appropriate (33,34). Outcomes that could not be pooled were summarized narratively. When data were available, considerable heterogeneity ($I^2 \geq 75\%$) was evaluated through prespecified subgroup analyses by prior TNFi exposure, TNFi drug, type of reduction strategy, follow-up duration, disease classification, and sex (35,36). A priori sensitivity analyses were planned to further investigate heterogeneity by removing studies with high risk of bias, follow-up visits <6 months, and studies investigating multiple TNFi.

RESULTS

Details of the screening process, study selection, and reasons for exclusion are summarized in Figure 1. Our search yielded a total of 2,717 records for screening. We reviewed 299 full texts

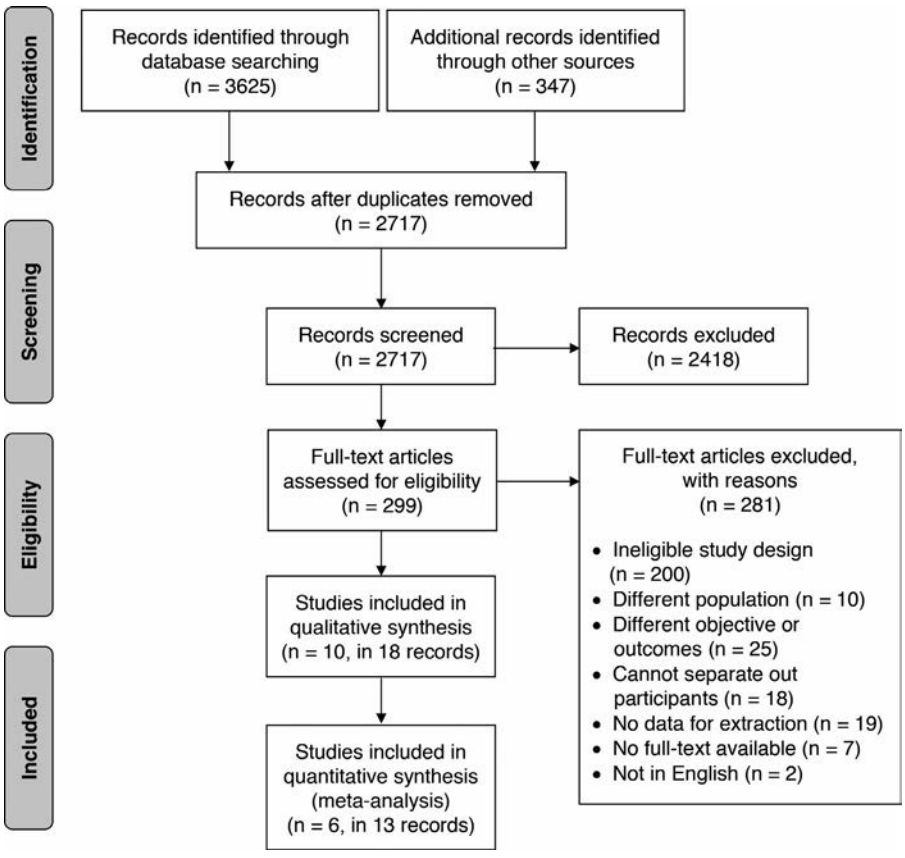

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of search and study selection.

for eligibility and included 13 articles and abstracts reporting on 6 RCTs with 747 participants in the meta-analysis. The study characteristics are summarized in Table 1. Of the included trials, there were 442 participants classified as having AS, and 305 classified as having nonradiographic axial SpA. The median age of participants across studies was 39 years, and the majority were male, with a median proportion of 76.9%. Disease duration ranged, where reported, with a median of 12 years. Two studies (37,38) included patients who were naive to TNFi therapy. One study investigated full discontinuation of TNFi (39) as a method of reduction, and the remaining 5 studies incorporated either a reduced dose (40,41) or an extended dosing interval (37,38,40,42) of the TNFi (see Table 2 for TNFi regimens). Three studies used etanercept (38,41,42), one used infliximab (37), one used adalimumab (39), and one used any of adalimumab, etanercept, golimumab, and infliximab (40), which could not be separated out by outcome. The studies were conducted primarily in Europe (France, Italy, Spain, and the UK), one in China, and there was one large multinational trial. Four additional studies were identified from trial registries, for which details are presented in Table 3.

Risk of bias. Our assessment of the overall risk of bias for each included study is summarized in Figure 2. We were unable to assess reporting bias as there were <10 eligible studies retrieved. The selective reporting and incomplete outcome data domains were judged to have the lowest risk of bias across studies, although

Table 1. Description and characteristics of included studies*

	Breban et al, 2008 (37)	Cantini et al, 2013 (42)	Yates et al, 2015 (41)	Li et al, 2016 (38)	Landewé et al, 2018 (39)	Gratacós et al, 2019 (40)
Design	Parallel	Parallel	Parallel (NI)	Parallel	Parallel	Parallel (NI)
Country	France	Italy	UK	China	Multiple†	Spain
No. of sites	32	1	2	2	107	22
TNFi	Infliximab	Etanercept	Etanercept	Etanercept	Adalimumab	Adalimumab, etanercept, golimumab, infliximab
Type of reduction	Extended interval	Extended interval	Reduced dose	Extended interval	Discontinuation	Extended interval and reduced dose (infliximab)
Naive to TNFi	Yes	No	No	Yes	No	No
Active disease at enrollment	Yes	No	NR	Yes	No	No
Form of disease	AS	AS	AS	AS	Nonradiographic axial SpA	AS
Classification criteria	Modified NY	Modified NY	Modified NY	Modified NY	ASAS	ASAS
Total no. of participants						
Standard dose	124	21	24	17	152	62
Reduced dose	62	22	23	26	153	61
Age, mean ± SD years						
Standard dose	41.4 ± 12.3	38‡	46.7 ± 14.1	22.0 ± 4.0§	34.7 ± 10.3	46.2 ± 13.7
Reduced dose	40.0 ± 9.6	37‡	46.7 ± 14.1	22.0 ± 4.0§	35.3 ± 10.2	43.7 ± 12.4
Sex, male						
Standard dose	93 (75.0)	16 (76.1)	41 (87.2)	14 (82.4)	96 (63.2)	53 (88.3)
Reduced dose	45 (72.6)	18 (81.8)	41 (87.2)	20 (76.9)	93 (60.8)	49 (81.7)
HLA–B27 positive						
Standard dose	92 (80.0)¶	NR	NR	NR	132 (86.8)	NR
Reduced dose	48 (82.8)#	NR	NR	NR	134 (87.6)	NR
Disease duration, mean ± SD years						
Standard dose	14.6 ± 10.5	12‡	NR	0.58 ± 0.23§	1.9 ± 2.9**	12.8 ± 10.4**
Reduced dose	13.8 ± 7.0	13‡	NR	0.58 ± 0.23§	1.8 ± 2.9**	10.7 ± 9.4**
Outcome measurement, no. of weeks	58	93††	26‡‡	12	40	52§§
Relapsed¶¶						
Standard dose	–	2 (9.5)	4 (16.7)	–	45 (29.6)	4 (6.4)
Reduced dose	–	3 (13.6)	11 (47.8)	–	81 (52.9)	6 (10.1)
Time to relapse, months##						
Standard dose	–	10.0 ± 1.1	6.0	–	9.2	12.0
Reduced dose	–	8.0 ± 3.2	6.0	–	9.2	12.0
Analysis***	ITT	ITT	ITT	ITT	ITT	ITT and per protocol
Trial number	NCT00439283 (ClinicalTrials)	NR	2010-029013-10 (EudraCT)	NR	NCT01808118 (ClinicalTrials)	2011-005871-18 (EudraCT); NCT01604629 (ClinicalTrials)

(Continued)

Table 1. (*Cont'd*)

	Breban et al, 2008 (37)	Cantini et al, 2013 (42)	Yates et al, 2015 (41)	Li et al, 2016 (38)	Landewé et al, 2018 (39)	Gratacós et al, 2019 (40)
Funding sources	Schering-Plough	NR	Pfizer	NR	AbbVie	Spanish Ministry of Health; ERDF; MINECO–ISCIII

* Values are the number (%) unless indicated otherwise. Standard dose = control arm; reduced dose = intervention arm. AS = ankylosing spondylitis; ASAS40 = Assessment of SpondyloArthritis international Society criteria for 40% improvement; ERDF = European Fund for Regional Development; ITT = intent-to-treat; MINECO–ISCIII = Ministerio de Economia y Consumo–Instituto de Salud Carlos III, Subdirección General de Evaluación; NI = noninferiority; NR = not reported; SpA = spondyloarthritis; TNFi = tumor necrosis factor inhibitor.
† Australia, Belgium, Brazil, Canada, Czechia, Denmark, Finland, France, Germany, Ireland, Italy, Republic of Korea, Mexico, Netherlands, New Zealand, Norway, Poland, Russian Federation, Slovakia, Spain, Sweden, Switzerland, UK, and US (based on trial record).
‡ Only median reported.
§ Only mean and range reported; SD was estimated using the range rule (i.e., SD = maximum–minimum/4).
¶ Only 115 of 124 with data.
# Only 58 of 62 with data.
** Reported as years from diagnosis.
†† Pooled mean ± SD 21.5 ± 1.4 months was derived from reported mean follow-up of 21 ± 1.6 months and 22 ± 1.1 months for standard- and reduced-dose arms, respectively.
‡‡ Reported as 6 months.
§§ Reported as 12 months.
¶¶ Per author definition (i.e., Breban et al: relapse; Yates et al: loss of clinical response; Landewé et al: any flare; Gratacós et al: relapse based on an adjusted percentage estimate from binomial regression); reported only for studies that required patients to be in remission or have stable disease at start of study.
Reported only for studies that required patients to be in remission or have stable disease at start of study.
*** Data extracted for review. Extracted ITT data for all outcomes except as follows: Yates et al (ASAS40 data based on 1 participant missing from standard- and reduced-dose arms); Landewé et al (ASAS40 and relapse rate data based on modified ITT, which incorporated a nonresponder imputation where only data prior to rescue therapy were included); Gratacós et al (baseline data were reported for a full analysis set, which included 120 participants who started the allocated treatment [2 and 1 participants lost to follow-up in standard- and reduced-dose arms, respectively]). Infections (any) and upper respiratory tract infection rates were based on ITT (standard dose: 62; reduced dose: 61). Relapse rates were based on per-protocol analysis (standard dose: 55, reduced dose: 58).

one study was judged at high risk of bias for both (39). Except for one double-blind study (39), the remaining studies were open-label designs, and blinding of outcome assessment, participants, and personnel was judged to have the highest risk of bias. Methods for random sequence generation and allocation concealment were not reported in 2 studies (38,42). Additionally, many outcomes (e.g., adverse events, ASAS40, and BASDAI score, some of which can be included in the criteria for remission/relapse) were subjectively measured. For most, all outcomes were reported as planned in the methods section. A summary of all pooled outcomes is presented in Table 4 (for forest plots, see Supplementary Figures 1–7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

Efficacy. *ASDAS.* There were insufficient data to pool the ASDAS. According to one study with 47 participants, the mean ± SD ASDAS increased in the dose reduction group from 2.15 ± 1.02 to 2.18 ± 0.94 and decreased in the standard-dose group from 1.70 ± 0.70 to 1.60 ± 0.79, with the latter differences being reported as statistically significant at $P < 0.05$ (41). One study reported the baseline median ASDAS with CRP for 120 of 123 participants and an adjusted score from a regression model at follow-up for a per-protocol set of participants (40). Additionally, some means were reported, but for different subsets of patients and not the originally allocated groups. Another study reported the mean ASDAS at baseline for all 305 participants but reported proportions at predefined cutoffs for the follow-up (e.g., inactive disease, major improvement, and clinically important improvement) (39).

ASAS40. Three studies with 535 participants reported on ASAS improvement criteria (37,39,41). Counts represent a maximum of 1 event (i.e., an event or state of improvement at the end of follow-up) per participant where more events are desirable. Fewer patients achieved ASAS40 (RR 0.62 [95% CI 0.49, 0.78]) with the reduced dose compared to the standard dose. The heterogeneity was low (I^2 = 21%) and appeared consistent visually (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

BASDAI. Four studies reported on the BASDAI, but one study was omitted from the analysis due to the presence of skewed data (e.g., large SDs) (41). Three studies with 272 participants were included in the analysis (37,38,42). Compared to the standard dose, there was a mean increase in BASDAI score with the reduced dose (MD 0.35 [95% CI 0.10, 0.60]) and low heterogeneity (I^2 = 0%) (see Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

CRP. Four studies reported on CRP level, but 2 studies were omitted from the analysis due to skewed data (37,41). Due to high heterogeneity (I^2 = 77.5%), the remaining 2 studies with 86 participants were not pooled (38,42). Upon visual inspection, studies

appeared inconsistent in their results, with a mean increase (MD 0.63 [95% CI 0.01, 1.24]) with the reduced dose in one study, and no difference (MD –0.31 [95% CI –0.92, 0.31]) between groups in the other study. Both studies evaluated extended intervals of etanercept; however, the study with no reported difference in CRP levels also enrolled TNFi-naive participants with newly diagnosed active disease over a short follow-up duration (38).

Remission. We distinguished between achieving and maintaining remission, only pooling those studies in which the eligibility criteria did not require patients to be in remission for enrollment. Among the studies that required patients to have stable disease or to be in remission at the start of the study, because the criteria varied between studies (e.g., BASDAI score ≤2, BASDAI score <4, ASDAS inactive disease, and other criteria), we could

Table 2. Summary of reduced (intervention arm) and standard (control arm) dose regimens of TNFi in included studies*

	Breban et al, 2008 (37)	Cantini et al, 2013 (42)	Yates et al, 2015 (41)	Li et al, 2016 (38)	Landewé et al, 2018 (39)	Gratacós et al, 2019 (40)
TNFi	Infliximab	Etanercept	Etanercept	Etanercept	Adalimumab	Adalimumab, etanercept, golimumab, infliximab
Route	Intravenous	Subcutaneous	Subcutaneous	Subcutaneous	Subcutaneous	Adalimumab, etanercept, and golimumab were subcutaneous; infliximab was intravenous
Standard of care for TNFi	5 mg/kg every 6–8 weeks	50 mg every week	50 mg every week	25 mg twice every week	40 mg every 2 weeks	Adalimumab, 40 mg every 2 weeks; etanercept 25 mg twice every week or 50 mg every week; golimumab 50 mg every 4 weeks; infliximab 5 mg/kg every 6–8 weeks
Standard dose administered	5 mg/kg at weeks 4, 6, and 10 (loading), then 5 mg/kg every 6 weeks; 7.5 mg/kg not earlier than at week 40 if relapse during 2 consecutive visits (n = 6)†	50 mg every week	50 mg every week	25 mg twice every week	40 mg every 2 weeks	Adalimumab 40 mg every 2 weeks; etanercept 25 mg every 3 days or 50 mg every week; golimumab 50 mg every 4 weeks; infliximab 5 mg/kg every 6–8 weeks
Reduced dose administered	5 mg/kg at weeks 4, 6, and 10 (loading), then 5 mg/kg on demand only if relapse, with a minimum interval of 4 weeks between infusions; 7.5 mg/kg not earlier than at 4th on-demand dose, if relapse (n = 4)†	50 mg every 2 weeks; 50 mg every week if relapse (n = 3)‡	25 mg every week; 50 mg every week if lost clinical response (n = 4)§	25 mg twice every week for 4 weeks, then 25 mg every week for 8 weeks	Placebo; 40 mg every 2 weeks for at least 12 weeks if flare (n = 68)¶	Adalimumab 40 mg every 3 weeks; etanercept 50 mg every 10 days; golimumab 50 mg every 6 weeks; infliximab 3 mg/kg every 8 weeks
Time point reduced dose administered	4 weeks after enrollment	Exact time point not specified	Exact time point not specified	0 weeks after randomization/allocation	0 weeks after randomization/allocation	Exact time point not specified
Duration of treatment (i.e., last follow-up visit)	Up to week 52 (standard), up to week 54 (reduced); study entry: April to October 2003	Up to weeks ~84–100; study entry: January 2010	Up to week 26 (6 months); study entry: November 2010 to September 2012	Up to week 12; study entry: March 2009	Up to week 40; study entry: April 2013 to September 2015	Up to week 52 (12 months); study entry: July 2012 to May 2014

(Continued)

Table 2. (*Cont'd*)

	Breban et al, 2008 (37)	Cantini et al, 2013 (42)	Yates et al, 2015 (41)	Li et al, 2016 (38)	Landewé et al, 2018 (39)	Gratacós et al, 2019 (40)
Additional notes	DMARDs (e.g., CSA, HCQ, IM gold, IV bisphosphonate MTX, SSZ, thiol compound) had to be withdrawn ≥4 weeks prior to enrollment; NSAID and corticosteroid doses had to be stable for ≥4 weeks prior to enrollment; 3-arm study with 2 intervention arms including 1 reduced TNFi + cointervention arm, which was excluded from this review	IM and systemic NSAIDs and corticosteroids not permitted during study; for pain, analgesics (acetaminophen or tramadol) were permitted; no patients received physical therapy or psychotherapy	Receipt of adjuvant treatments (e.g., glucocorticoids or NSAIDs) was permitted	Oral SSZ for 3 weeks (0.75 gm/day, 3 times every day in 1st week; 1.5 gm/day, 3 times every day in 2nd week; 2.0 gm/day, twice every day in 3rd week) and oral LEF for 12 weeks (20 mg/day) permitted for peripheral arthritis	Doses of DMARDs had to be stable for 28 days, and NSAIDs, analgesics, or corticosteroids for 14 days prior to baseline; start or changes in concomitant medications not permitted until either the end of the study, or at least 12 weeks of rescue therapy; after rescue, start or changes in concomitant medications (e.g., analgesics, corticosteroids at maximum dose ≤10 mg prednisone equivalent every day, DMARDs, NSAIDs) permitted	NSAIDs and dose modification permitted throughout the study (details about any dose modifications that may have occurred not provided); DMARD and NSAID use was not standardized and could be adjusted by investigators as needed to control symptoms

* ASDAS = Ankylosing Spondylitis Disease Activity Score; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; CRP = C-reactive protein; CSA = cyclosporine; DMARDs = disease-modifying antirheumatic drugs; HCQ = hydroxychloroquine; IM = intramuscular; IV = intravenous; LEF = leflunomide; MTX = methotrexate; NSAIDs = nonsteroidal antiinflammatory drugs; SSZ = sulfasalazine; TNFi = tumor necrosis factor inhibitor.
† Relapse: as indicated by a negative answer to "Since the last connection, do you think that your disease has remained under control?", a positive answer to "Since the last connection, do you think that your disease has been worsening?", and either an increase in BASDAI score of ≥1 on a 10-point scale, or an increase in the patient's assessment of pain of ≥2 on a 10-point scale compared with the lowest score reached by that patient since the first infusion.
‡ Relapse: no definition provided.
§ Clinical response: as indicated by a 50% reduction in BASDAI score, or a fall ≥2 units and ≥2-unit reduction in spinal pain as measured on a 10-point scale at 6 months postrandomization.
¶ Flare: as indicated by ASDAS ≥2.1 (high disease activity) on 2 consecutive visits; for the second consecutive visit to qualify as a flare, the ASDAS was calculated using the last available high-sensitivity CRP level prior to the visit.

not assess the maintenance of remission. Two studies with 233 participants reported on remission according to the ASAS criteria for partial remission (i.e., the proportion of participants achieving remission among those who did not have stable disease at baseline) (37,41). Compared to the standard dose, there were fewer remission events achieved with the reduced dose (RR 0.17 [95% CI 0.06, 0.46]) and low heterogeneity (I^2 = 0%) (see Supplementary Figure 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

Relapse. Although the basis for measuring relapse was not always clear, we sought to take a more conservative approach for assessing relapse and therefore used the study authors' definitions including: BASDAI score ≥4 (40,42), disease flare (39), and increased infusion dose (37). Four studies with 647 participants were pooled and compared to the standard dose; there were more relapse and/or disease flare events with the reduced dose (RR 1.73 [95% CI 1.32, 2.27]). Heterogeneity was low (I^2 = 0%), suggesting that it was reasonable to pool these studies (see Supplementary Figure 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

Safety. The reporting of safety data varied widely across studies. Rate of infections was the most commonly reported, and among these, upper respiratory tract infections (URTI) were reported separately most often.

Any infections. Six studies with 747 participants reported on the total number of infections during the study period (37–42). Infections could not be differentiated on the basis of seriousness due to differences in reporting and were instead pooled as "any infection." There was no difference in infection rates between groups (IRR 0.98 [95% CI 0.76, 1.25]), and the heterogeneity was low (I^2 = 0%) (see Supplementary Figure 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

Table 3. Summary of ongoing or recently completed unpublished trials identified from registries*

Trial Number	NCT01610947	NCT02505542	NCT03253796	NTR7640
Design	Parallel	Parallel	Parallel	Parallel
Country	France	Multiple†	Multiple‡	Netherlands
No. of sites	1	108	71	NR
TNFi	Adalimumab, etanercept, golimumab, infliximab	Certolizumab	Golimumab	Adalimumab, certolizumab, etanercept, golimumab, infliximab
Type of reduction	Extended interval	Extended interval and discontinuation	Extended interval and discontinuation	Extended interval, discontinuation, and reduced dose (infliximab)
Naive to TNFi	No	No	No	No
Form of disease	SpA	Axial SpA	Nonradiographic axial SpA	Axial SpA
Enrolled participants	398	736	300 (estimated)	190 (estimated)
Age, years	18+	18–45	18–45	18+
Funding sources or sponsor	University Hospital, Montpellier	UCB Biosciences, Parexel	Merck Sharp & Dohme	Sint Maartenskliniek
Status	Completed, results unavailable	Active, recruitment closed	Active, recruitment closed	Not started

* NR = not reported; SpA = spondyloarthritis; TNFi = tumor necrosis factor inhibitor.
† Belgium, Bulgaria, Czechia, France, Germany, Hungary, Netherlands, Poland, Romania, Spain, Taiwan, Turkey, UK, and US.
‡ Czechia, Germany, Netherlands, Poland, Romania, Russian Federation, Spain, Turkey, and Ukraine.

URTI. Three studies with 471 participants reported on rates of URTIs (39,40,42). There was no difference between groups (IRR 0.71 [95% CI 0.41, 1.19]), and the heterogeneity was low (I^2 = 0%) (see Supplementary Figure 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

Injection or infusion reactions. Four studies with 319 participants reported rates of infusion reactions (i.e., for infliximab, which is administered intravenously) (37) and injection reactions (i.e., for etanercept, which is administered subcutaneously) (38,41,42). There was no difference between groups (IRR 1.14 [95% CI 0.58, 2.25]), and the heterogeneity was low (I^2 = 22.6%) (see

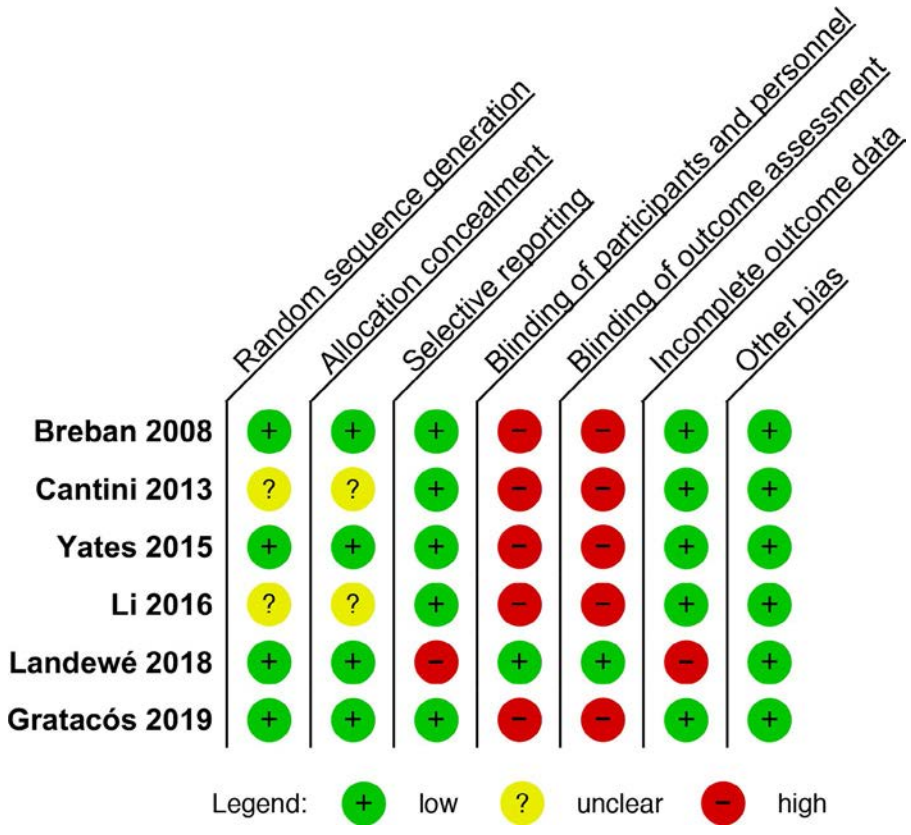

Figure 2. Summary of risk of bias in the included studies with review authors' judgments. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract.

Table 4. Summary of meta-analyses*

Outcome	No. of studies (no. of participants)	References	Effect estimate (95% CI)	I^2, %	Favors
Efficacy					
ASAS40	3 (535)	37,39,41	RR 0.62 (0.49, 0.78)	21	Standard
BASDAI score	3 (272)	37,38,42	MD 0.35 (0.10, 0.60)	0	Standard
Remission					
ASAS partial remission	2 (233)	37,41	RR 0.17 (0.06, 0.46)	0	Standard
Relapse					
Relapse	4 (647)	37,39,40,42	RR 1.73 (1.32, 2.27)	0	Standard
Safety					
Any infections	6 (747)	37–42	IRR 0.98 (0.76, 1.25)	0	ND
URTI	3 (471)	39,40,42	IRR 0.71 (0.42, 1.19)	0	ND
Injection/infusion reactions	4 (319)	38,41,42	IRR 1.14 (0.58, 2.25)	22.6	ND

* 95% CI = 95% confidence interval; ASAS40 = Assessment of SpondyloArthritis international Society criteria for 40% improvement; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; IRR = incidence rate ratio; MD = mean difference; ND = no difference; RR = risk ratio; URTI = upper respiratory tract infections.

Supplementary Figure 7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

QoL. There were insufficient data to pool QoL. Three trials reported on generic measures, including the EuroQoL 5-domain (EQ-5D) questionnaire, Short Form 36 (SF-36) health survey, and the Health Assessment Questionnaire for the Spondyloarthropathies (HAQ-S). One study reported a statistically significant ($P < 0.02$) difference between the standard dose and reduced dose on the mental component of the SF-36 at follow-up. However, the reduced dose included data for both reduction arms, including the reduced dose plus methotrexate cointervention arm (37). There were no differences on the SF-36 physical component. Another study reported no differences ($P = 0.092$) between groups in the mean change from baseline for HAQ-S, although not all randomized participants were assessed (39). A third study used the EQ-5D and reported no differences for this outcome measure for all randomized participants (41). Two trials (40,41) reported data for disease-specific measures: the Ankylosing Spondylitis Quality of Life (ASQoL) questionnaire and the Evaluating Ankylosing Spondylitis Quality of Life (Easi-QoL) questionnaire. One study reported increased limitation (e.g., higher scores) in physical function at follow-up, according to the EasiQoL, in the standard-dose group ($P < 0.05$) and no differences for the other domains or in the reduced-dose group (41). Similarly, there were no reported differences in ASQoL, although scores were increased (e.g., worse QoL) at follow-up for both groups, and data were likely skewed in the standard-dose arm (41). The other trial did not provide baseline data for ASQoL for comparison, and follow-up values were based on an adjusted regression model. Additional data for this measure were reported for a smaller subset of participants and not the randomized groups (40).

Subgroup and sensitivity analyses. We planned to investigate considerable heterogeneity (i.e., overall $I^2 \geq 75\%$) for possible sources, which was only evident in the meta-analysis for CRP level. However, since only 2 studies were pooled for this outcome, the a priori subgroup analyses and exclusion of studies based on predetermined criteria were not possible. Additionally, it was not possible to assess any sex differences because no studies reported data by sex. We completed post hoc subgroup analyses for all outcomes when >2 studies were pooled and found no significant subgroups.

Our results were robust to sensitivity analyses, with the exclusion of studies that had short follow-up periods (38), investigated several TNFi therapies (40), and enrolled patients with active disease or where this was unreported (37,38,41). The observed effect for relapse vanished with the exclusion of a study that investigated TNFi discontinuation (RR 1.39 [95% CI 0.64, 2.99; $I^2 = 0\%$]); however, this was also the largest trial contributing 305 participants (39) (see Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24184/abstract).

DISCUSSION

To our knowledge, this is the first systematic review with a meta-analysis of RCTs investigating the safety and efficacy of TNFi dose reduction for the treatment of axial SpA. To strengthen the quality of this review and provide a clearer assessment of the effect of TNFi dose reduction in axial SpA, we restricted our analysis of the current evidence to RCTs only. Our findings confirm previous studies in which standard TNFi doses suggest some benefit for disease activity and are more efficacious for achieving and maintaining stable disease. There were no discernable group differences for general measures of safety, including rates of infection. From a clinical standpoint, it is unclear whether there is a strong enough clinical basis to withdraw therapy altogether. Although treatment recommendations for the best dose reduction strategies cannot be made at this time given the heterogeneity in tapering strategies reported in the literature, this decision should be an individualized one between the patient and their physician. Additionally, since most of the studies only included patients with

AS, our study findings may not be generalizable to nonradiographic axial SpA, which is considered to be an earlier or milder form of disease than AS. It is possible that TNFi dose reduction strategies may be acceptable in this group, but further research is needed.

We have faced several limitations during the conduct of this review. Due to the small sample of available RCTs, it was impractical to compare subgroups based on dose regimens because all studies implemented various strategies. Across all studies, we found that for most outcomes there was some degree of high-risk bias (e.g., blinding and selection bias). This can lead to concerns with detection and performance biases if participants are aware of the TNFi regimen that they receive, impacting the perception of their condition and the measurement of outcomes relating to their disease state. For example, participants might be more cognizant of specific symptoms and overreport if they are aware that their dose has been reduced or discontinued. This can also result in methodologic challenges in study conduct and introduces difficulties in the critical appraisal of study designs. Researchers should employ better practices in addressing important biases (i.e., selection, performance, and detection biases), as well as practicing better reporting by following the Consolidated Standards of Reporting Trials (CONSORT) statement guidelines for RCTs (43). Although concerns with blinding and self-reporting cannot be neglected, 2 double-blind RCTs, one recently completed and another ongoing at the time of this study (NCT02505542 and NCT03253796) can provide valuable data by addressing these issues.

Due to insufficient data, we were unable to assess some of the most patient-important outcomes (ASDAS, maintenance of remission, and quality of life) in this meta-analysis. There are additional outcomes that we did not assess that can further support the evaluation of efficacy and safety of TNFi dose reduction. However, the availability of certain data, such as the modified Stoke Ankylosing Spondylitis Spinal Score for AS, is typically limited in the context and duration of RCTs. Long-term follow-up data from cohort studies may be of greater value for assessing such outcomes. As such, incorporation of evidence from only RCTs may detract from the more pragmatic approaches to dose tapering that are sought, especially when application of tapering strategies in a real-world setting is the primary goal. For example, there is value in assessing the economic impacts of reducing TNFi doses. In settings where drugs are not available through a subscribed health plan, the option for a dose reduction or extended administration interval can be an important consideration. Future research should also aim to study both tapering and withdrawal approaches in more detail to better characterize the mechanisms and outcomes involved for each pathway of dose reduction. Our analysis is also limited in its completeness without the incorporation of newer, TNFi biosimilar therapies that are rapidly becoming available.

Although there are large variations in the available observational studies on this topic, some designs, including large prospective cohorts, can offer insights into the types of patients who may be most responsive to dose reduction strategies. Covariates that may be informative and should be explored with meta-regression analyses include optimal criteria for determining remission (e.g., ASAS criteria, ASDAS, or BASDAI score), form of disease, smoking history, symptom and disease duration, positivity for HLA–B27, sex, and previous drug exposure. For example, since symptoms of axial SpA can wax and wane, the severity of disease may be an effect modifier. Therefore, various definitions of disease stability may impact our understanding of the effects of different dose reduction strategies on patient outcomes. Interpretation of efficacy can also be misconstrued with the ASAS change score because ASAS40 can represent a change from 9 to 5 on the BASDAI, where 5 is still a high BASDAI score. In this case, the ASDAS may be the ideal measure because it represents a state. When considering change that is clinically meaningful, the minimum clinically important difference for BASDAI score has been explored in the literature, and 2 studies have shown changes of ~1.0–1.1 units on a 0–10 visual analog scale to be considered clinically important for improvement (44,45). In our meta-analysis based on 3 RCTs, there was a mean increase of 0.35 units in the reduced-dose group, which may suggest a difference that is not clinically meaningful. Examining the marginally opposite effects on CRP level between studies, since variability in results is multifactorial and extends beyond different patient characteristics, there may be alternative reasons for the observed differences, including permitted concurrent medications (e.g., NSAIDs), unreported factors, and chance.

Dose response and individual patient data meta-analyses can provide insight into some of the observed heterogeneity across trials and should be explored in future studies (46). Given the growing body of evidence on this topic, future evidence syntheses should also incorporate network meta-analyses, which allow for the comparison of multiple interventions and combinations (47). As there are currently no standard TNFi tapering or withdrawal regimens for axial SpA, network meta-analyses can aid our understanding of the differences in clinical and patient-important outcomes between different TNFi regimens including dose reduction strategies. This can also help determine optimal strategies with regard to better patient outcomes.

ACKNOWLEDGMENTS

We thank the Health Research Methodology Graduate Program systematic review course coordinators at the Department of Health Research Methods, Evidence, and Impact at McMaster University for their invaluable support, critiques, and insightful comments to improve the first draft of this review.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Ms Lawson 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. Lawson, Mbuagbaw, Omar, Inman.

Acquisition of data. Lawson, Eraso, Joanes, Aves, Leenus.
Analysis and interpretation of data. Lawson, Eraso, Mbuagbaw, Inman.

REFERENCES

1. Inman R, Sieper J. Oxford textbook of axial spondyloarthritis. Oxford: Oxford University Press; 2016.
2. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis 2009;68 Suppl 2:ii1–44.
3. Gaston H. Mechanisms of disease: the immunopathogenesis of spondyloarthropathies. Nat Clin Pract Rheumatol 2006;2:383–92.
4. Inman RD. Mechanisms of disease: infection and spondyloarthritis. Nat Clin Pract Rheumatol 2006;2:163–9.
5. Maxwell LJ, Zochling J, Boonen A, Singh JA, Veras MM, Tanjong Ghogomu E, et al. TNF-alpha inhibitors for ankylosing spondylitis. Cochrane Database Syst Rev 2015:CD005468.
6. Van der Heijde D, Ramiro S, Landewe R, Baraliakos X, van den Bosch F, Sepriano A, et al. 2016 update of the ASAS–EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis 2017;76:978–91.
7. Silva LC, Ortigosa LC, Benard G. Anti–TNF-α agents in the treatment of immune–mediated inflammatory diseases: mechanisms of action and pitfalls. Immunotherapy 2010;2:817–33.
8. Gensler LS. Axial spondyloarthritis: the heart of the matter. Clin Rheumatol 2015;34:995–8.
9. Baraliakos X, Kiltz U, Heldmann F, Sieper J, Braun J. Withdrawal of biologic therapy in axial spondyloarthritis: the experience in established disease. Clin Exp Rheumatol 2013;31 Suppl 78:S43–6.
10. Edwards CJ, Fautrel B, Schulze-Koops H, Huizinga TW, Kruger K. Dosing down with biologic therapies: a systematic review and clinicians' perspective. Rheumatology (Oxford) 2017;56:1847–56.
11. Navarro-Compan V, Plasencia-Rodriguez C, de Miguel E, Balsa A, Martin-Mola E, Seoane-Mato D, et al. Anti-TNF discontinuation and tapering strategies in patients with axial spondyloarthritis: a systematic literature review. Rheumatology (Oxford) 2016;55:1188–94.
12. Almirall M, Salman-Monte TC, Lisbona MP, Maymo J. Dose reduction of biological treatment in patients with axial spondyloarthritis in clinical remission: are there any differences between patients who relapsed and to those who remained in low disease activity? Rheumatol Int 2015;35:1565–8.
13. De Stefano R, Frati E, De Quattro D, Menza L, Manganelli S. Low doses of etanercept can be effective to maintain remission in ankylosing spondylitis patients. Clin Rheumatol 2014;33:707–11.
14. Jois RN, Leeder J, Gibb A, Gaffney K, Macgregor A, Somerville M, et al. Low-dose infliximab treatment for ankylosing spondylitis: clinically- and cost-effective. Rheumatology (Oxford) 2006;45:1566–9.
15. Lee SH, Lee YA, Hong SJ, Yang HI. Etanercept 25 mg/week is effective enough to maintain remission for ankylosing spondylitis among Korean patients. Clin Rheumatol 2008;27:179–81.
16. Plasencia C, Kneepkens EL, Wolbink G, Krieckaert CL, Turk S, Navarro-Compan V, et al. Comparing tapering strategy to standard dosing regimen of tumor necrosis factor inhibitors in patients with spondyloarthritis in low disease activity. J Rheumatol 2015;42:1638–46.
17. Steel L, Gaffney K. Maintenance of low disease activity following tumor necrosis factor inhibitor dose tapering in ankylosing spondylitis. J Rheumatol 2017;44:1292.
18. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.
19. Lawson DO, Eraso M, Mbuagbaw L, Joanes M, Aves T, Leenus A, et al. TNF inhibitor dose tapering in axial spondyloarthritis: a systematic review and meta-analysis. 2018. URL: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018091146.
20. Benhamou M, Gossec L, Dougados M. Clinical relevance of C-reactive protein in ankylosing spondylitis and evaluation of the NSAIDs/cox-ibs' treatment effect on C-reactive protein. Rheumatology (Oxford) 2010;49:536–41.
21. Brandt J, Listing J, Sieper J, Rudwaleit M, van der Heijde D, Braun J. Development and preselection of criteria for short term improvement after anti-TNF alpha treatment in ankylosing spondylitis. Ann Rheum Dis 2004;63:1438–44.
22. Landewe R, van Tubergen A. Clinical tools to assess and monitor spondyloarthritis. Curr Rheumatol Rep 2015;17:47.
23. Machado P, Landewe R, Lie E, Kvien TK, Braun J, Baker D, et al. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 2011;70:47–53.
24. Spoorenberg A, van Tubergen A, Landewe R, Dougados M, van der Linden S, Mielants H, et al. Measuring disease activity in ankylosing spondylitis: patient and physician have different perspectives. Rheumatology (Oxford) 2005;44:789–95.
25. Van der Heijde D, Lie E, Kvien TK, Sieper J, van den Bosch F, Listing J, et al. ASDAS, a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:1811–8.
26. Zochling J, Braun J. Remission in ankylosing spondylitis. Clin Exp Rheumatol 2006;24 Suppl 43:S-88–92.
27. Sieper J. How to define remission in ankylosing spondylitis? Ann Rheum Dis 2012;71 Suppl 2:i93–5.
28. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016;5:210.
29. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.
30. Rohatgi A. WebPlotDigitizer version 4.1. 2018. URL: https://automeris.io/WebPlotDigitizer.
31. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol 2014;14:135.
32. Harris R, Bradburn M, Deeks J, Harbord R, Altman D, Sterne J. Metan: fixed- and random-effects meta-analysis. Stata J 2008;8:3–28.
33. Deeks JJ, Higgins JP, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins P, Green S, editors. Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration; 2008. p. 649.
34. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.
35. Wallis D, Haroon N, Ayearst R, Carty A, Inman RD. Ankylosing spondylitis and nonradiographic axial spondyloarthritis: part of a common spectrum or distinct diseases? J Rheumatol 2013;40:2038–41.
36. Ward MM, Deodhar A, Akl EA, Lui A, Ermann J, Gensler LS, et al. American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and Treatment Network 2015 recommendations for the treatment of ankylosing spondylitis and nonradiographic axial spondyloarthritis. Arthritis Care Res (Hoboken) 2016;68:151–66.
37. Breban M, Ravaud P, Claudepierre P, Baron G, Henry YD, Hudry C, et al. Maintenance of infliximab treatment in ankylosing spondylitis: results of a one-year randomized controlled trial comparing systematic versus on-demand treatment. Arthritis Rheum 2008;58:88–97.
38. Li J, Wang X, Han Z, Zhang Y, Wang Y, Zhang Y, et al. Dose reduction of recombinant human tumor necrosis factor inhibitors (etanercept) can be effective in ankylosing spondylitis patients with synovitis

of the hip in a Chinese population. Int J Immunopathol Pharmacol 2016;29:510–5.

39. Landewe R, Sieper J, Mease P, Inman RD, Lambert RG, Deodhar A, et al. Efficacy and safety of continuing versus withdrawing adalimumab therapy in maintaining remission in patients with non-radiographic axial spondyloarthritis (ABILITY-3): a multicentre, randomised, double-blind study. Lancet 2018;392:134–44.
40. Gratacos J, Pontes C, Juanola X, Sanz J, Torres F, Avendano C, et al. Non-inferiority of dose reduction versus standard dosing of TNF-inhibitors in axial spondyloarthritis. Arthritis Res Ther 2019;21:11.
41. Yates M, Hamilton LE, Elender F, Dean L, Doll H, MacGregor AJ, et al. Is etanercept 25 mg once weekly as effective as 50 mg at maintaining response in patients with ankylosing spondylitis? a randomized control trial. J Rheumatol 2015;42:1177–85.
42. Cantini F, Niccoli L, Cassara E, Kaloudi O, Nannini C. Duration of remission after halving of the etanercept dose in patients with ankylosing spondylitis: a randomized, prospective, long-term, follow-up study. Biologics 2013;7:1–6.
43. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869.
44. Pavy S, Brophy S, Calin A. Establishment of the minimum clinically important difference for the bath ankylosing spondylitis indices: a prospective study. J Rheumatol 2005;32:80–5.
45. Van der Heijde D, van der Linden S, Dougados M, Bellamy N, Russell AS, Edmonds J. Ankylosing spondylitis: plenary discussion and results of voting on selection of domains and some specific instruments. J Rheumatol 1999;26:1003–5.
46. Higgins JP, Green S, editors. Cochrane handbook for systematic reviews of interventions version 5.1.0. The Cochrane Collaboration; 2011. URL: www.cochrane-handbook.org.
47. Mbuagbaw L, Rochwerg B, Jaeschke R, Heels-Andsell D, Alhazzani W, Thabane L, et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. Syst Rev 2017;6:79.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 873–884
DOI 10.1002/acr.24195

American College of Rheumatology
Empowering Rheumatology Professionals

Most Appropriate Conventional Disease-Modifying Antirheumatic Drug to Combine With Different Advanced Therapies in Rheumatoid Arthritis: A Systematic Literature Review With Meta-Analysis

Guillaume Decarriere,[1] Thomas Barnetche,[2] Bernard Combe,[1] Cécile Gaujoux-Viala,[3] Cédric Lukas,[1] Jacques Morel,[1] and Claire Daien[1]

Objective. In rheumatoid arthritis, the association between advanced therapies (including biologic disease-modifying antirheumatic drugs [DMARDs] and targeted synthetic DMARDs) and methotrexate (MTX) is recommended by international societies. When MTX cannot be used, other conventional synthetic DMARDs (csDMARDs) may be proposed. We aimed to compare the safety and efficacy of MTX and non-MTX csDMARDs in combination with advanced therapies.

Methods. We systematically searched the literature for studies comparing the effectiveness, retention rate, and safety of MTX versus non-MTX csDMARDs (leflunomide or others) in combination with tumor necrosis factor inhibitors (TNFi), abatacept, rituximab, tocilizumab, and JAK inhibitors. Meta-analysis was performed with RevMan, using an inverse variance approach with fixed or random-effects models. Risk ratios (RRs) and 95% confidence intervals (95% CIs) were estimated.

Results. The literature search revealed 3,842 articles; 41 studies were included for the systematic literature review and 21 for the meta-analysis: 13 with TNFi, 3 with abatacept, and 5 with rituximab. For TNFi, the European Alliance of Associations for Rheumatology (EULAR) response at 6 months was lower for patients receiving non-MTX csDMARDs than for those using MTX (RR 0.93 [95% CI 0.87, 1.0], $P = 0.04$; n = 3,843; $I^2 = 28\%$), with a lower retention rate at 12 months. For abatacept, effectiveness and safety were similar between the 2 groups. For rituximab, a good EULAR response was higher with leflunomide than MTX (RR 1.38 [95% CI 1.13, 1.68], $P = 0.001$; n = 2,078; $I^2 = 0\%$), with similar adverse event rates. Meta-analysis for tocilizumab or JAK inhibitors could not be performed.

Conclusion. The different csDMARDs seem safe and efficient to combine with advanced therapies in RA patients. Although MTX seems slightly superior to other csDMARDs in combination with TNFi, leflunomide might be superior to MTX in combination with rituximab.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterized by synovial inflammation leading to both cartilage destruction and bone erosions. Advanced therapies (biologic disease-modifying antirheumatic drugs [DMARDs] and targeted synthetics DMARDs) have changed RA management and the outcomes for RA patients. Methotrexate (MTX) is the first-line therapy

Presented in part at the 2018 Annual Scientific Meeting of the American College of Rheumatology and as a poster session at the 2018 Annual Meeting of the French Society of Rheumatology.

Supported by AbbVie France.

[1]Guillaume Decarriere, MD, Bernard Combe, MD, PhD, Cédric Lukas, MD, PhD, Jacques Morel, MD, PhD, Claire Daien, MD, PhD: CHU Montpellier and Montpellier University, Montpellier, France; [2]Thomas Barnetche, PhD: University Hospital Pellegrin, Bordeaux, France; [3]Cécile Gaujoux-Viala, MD, PhD: University Hospital Caremeau, Nîmes, France.

Dr. Decarriere has received an honorarium from AbbVie (less than $10,000). Dr. Barnetche has received honoraria from Eli Lilly and Company (less than $10,000). Dr. Combe has received honoraria from AbbVie, Bristol Myers Squibb, Gilead, Janssen, Eli Lilly and Company, MSD, Novartis, Pfizer, Roche-Chugai, Sanofi, and UCB (less than $10,000 each). Dr. Gaujoux-Viala has received honoraria from AbbVie, Bristol Myers Squibb, Celgene, Janssen, Eli Lilly and Company, Medac, MSD, Nordic Pharma, Novartis, Pfizer, Roche-Chugai, Sanofi, Schering, and UCB (less than $10,000 each). Dr. Lukas has received honoraria from AbbVie, Bristol Myers Squibb, Celgene, Janssen, Medac, MSD, Novartis, Pfizer, Sanofi, Schering, Roche-Chugai, and UCB (less than $10,000 each). Dr. Morel has received honoraria from AbbVie, Bristol Myers Squibb, Celgene, Janssen, MSD, Novartis, Pfizer, Sanofi, Schering, Roche-Chugai, and UCB (less than $10,000 each). Dr. Daien has received honoraria from AbbVie, Bristol Myers Squibb, Eli Lilly and Company, Novartis, Pfizer, Roche-Chugai, Sanofi, and UCB (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Claire Daien, MD, PhD, Department of Rheumatology, CHU Montpellier and Montpellier University, 191 Avenue du Doyen Gaston Giraud, 34295 Montpellier cedex 5, France. Email: c-daien@chu-montpellier.fr.

Submitted for publication June 4, 2019; accepted in revised form March 17, 2020.

SIGNIFICANCE & INNOVATIONS

- The different conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) are safe and efficient in combination with advanced therapies in the management of rheumatoid arthritis.
- Methotrexate (MTX) might be superior to other csDMARDs in combination with a tumor necrosis factor inhibitor.
- Leflunomide might be superior to MTX in combination with rituximab.

for RA and represents an anchor drug (1,2). With an inadequate response to MTX, advanced therapies, including tumor necrosis factor inhibitors (TNFi: golimumab, etanercept [ETA], adalimumab, certolizumab, and infliximab [INF]), abatacept (ABA), rituximab (RTX), tocilizumab (TCZ), and JAK inhibitors (tofacitinib and baricitinib) can be used. The association between these advanced therapies and MTX is recommended by international societies (3). In case of contraindication or intolerance to MTX, other conventional synthetic DMARDs (csDMARDs) such as leflunomide (LEF) or sulfasalazine (SSZ) can be used as first line-therapy or combined with advanced therapies (2). LEF and SSZ have demonstrated efficacy and safety as monotherapy (4). However, their effectiveness and safety versus MTX in combination with advanced therapies remain unclear. The aim of this study was to compare the efficacy and safety of non-MTX csDMARDs and MTX combined with advanced therapies for RA based on a systematic literature review and meta-analysis.

MATERIALS AND METHODS

Literature search. This systematic literature review with meta-analysis followed the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement. Articles in English, French, or Spanish published up to February 2018 were searched in MEDLINE via PubMed, Embase, and databases from the American College of Rheumatology (ACR) and European Alliance of Associations for Rheumatology (EULAR) 2016 and 2017 annual meetings. The search keywords used are summarized in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract. Two reviewers (GD and CD) independently screened titles and abstracts using the website COVIDENCE (https://www.covidence.org/home). Disagreements were discussed to achieve consensus. Initially, the reference lists of relevant articles were manually searched to identify articles. The first search strategy used did not capture all articles, so we modified the strategy in order not to miss any relevant article. This new search strategy captured all articles identified by the manual search. The trials were selected on the basis of titles and abstracts, then on full text; duplicates were removed.

Study selection. The research questions were structured using the PICOT format (Patients, Intervention, Comparators, Outcome, Type of study). Studies had to investigate adults (age ≥18 years) with RA regardless of disease activity, and duration and had to involve an advanced therapy including TNFi (golimumab, certolizumab, etanercept, adalimumab, INF), ABA, TCZ, RTX, or JAK inhibitors (tofacitinib, baricitinib, upadacitinib, filgotinib, solcitinib, peficitinib, or itacitinib) in patients receiving MTX. The comparator had to involve introduction of the same advanced therapy in patients receiving non-MTX csDMARDs (all csDMARDs or LEF alone). Outcomes had to include effectiveness by using the Disease Activity Score in 28 joints (DAS28) at follow-up, change in DAS28 score, achievement of DAS28 remission and low-disease activity, EULAR response, Health Assessment Questionnaire disability index (HAQ DI) and change in HAQ DI, radiographic progression, retention rate (percentage), safety assessing all adverse events (AEs) and serious AEs (SAEs), and the number of drug interruptions due to AEs. We included observational cohort studies and randomized controlled trials (RCTs) but not case reports or reviews.

Data extraction and quality assessment. Data on the study design, sample size, patient and control characteristics (age, sex, disease duration, advanced therapy use, csDMARD use and dose, prednisone use and dose, clinical baseline characteristics, previous csDMARDs, previous advanced therapy), and efficacy and safety outcomes were collected using a predetermined Excel form. The quality and risk of bias of studies suitable for meta-analysis were assessed using the Cochrane Risk of Bias tool (ROBINS-I) (5), which evaluates participation, attrition, prognostic factor measurement, outcome measurement, confounding, and analysis.

Data synthesis and analysis. Meta-analysis was performed when at least 2 studies evaluated the same outcome, in the non-MTX csDMARDs group versus MTX group, with all data available (number or percentage of patients and mean ± SD). Statistical analysis of data involved a comprehensive meta-analysis performed with RevMan software and an inverse variance approach with fixed or random-effects models, depending on the presence of heterogeneity (cutoff: I^2 = 20%). We estimated risk ratios (RRs) and 95% confidence intervals (95% CIs). Publication bias was evaluated by the Egger test and funnel plots.

RESULTS

Literature search results and study characteristics. From 3,842 articles revealed by the search, 144 articles and abstracts were of potential interest; further examination resulted in 21 studies included for the meta-analysis (Figure 1): 13 TNFi, 3 ABA, and 5 RTX. No meta-analysis could be performed for TCZ or JAK inhibitors. For a systematic literature review, 20 studies were included: 5 TCZ, 3 JAK inhibitors, 1 ABA, 4 RTX, and 7 TNFi.

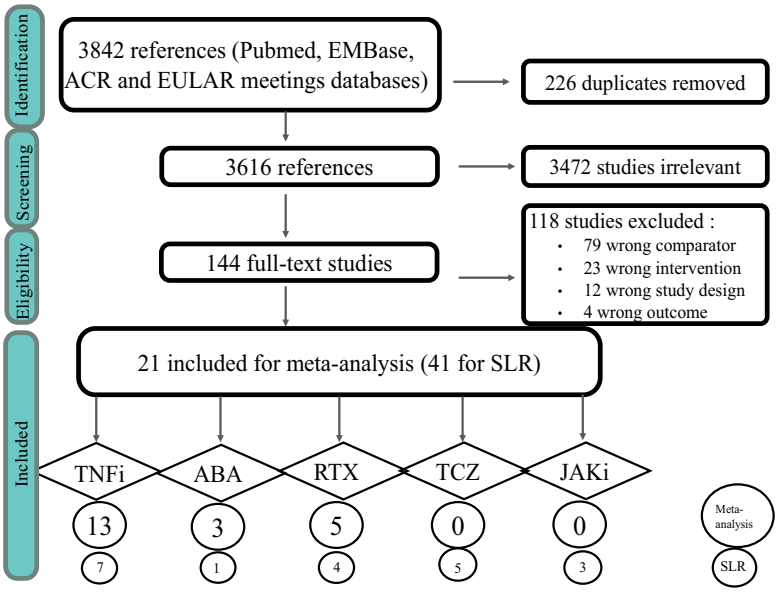

Figure 1. Flow chart of the systematic literature review with meta-analysis. ABA = abatacept; ACR = American College of Rheumatology; EULAR = European Alliance of Associations for Rheumatology; JAKi = JAK inhibitors; RTX = rituximab; SLR = systematic literature review; TCZ = tocilizumab; TNFi = tumor necrosis factor inhibitors.

For studies that were not included in the meta-analysis for RTX and TNF (systematic literature review), no difference in term of effectiveness or safety was identified (not included due to insufficient data, lack of mean ± SD, or other outcome evaluated). If studies had 2 non-MTX csDMARDs groups, we considered each group (comparator and intervention) as 2 independent studies.

Study characteristics. Table 1 shows the characteristics of studies included in the meta-analysis of TNFi, ABA, and RTX. Most of the studies were cohort studies. Other studies were post hoc analyses of RCTs. Baseline characteristics of patients in each study are in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract. Baseline disease activity (DAS28 score) and severity (age, sex, rheumatoid factor, anti–citrullinated protein antibodies [ACPAs]) as well as disease duration were similar among patients receiving LEF or MTX in each biologic DMARD group, when available. The only difference was the use of glucocorticoids, with a higher percentage of use in the MTX group with RTX. For MTX, the mean dose was 15 mg/week, except for Koyama et al (6 mg/week) (6). For LEF, most patients received 20 mg/day. The methodologic quality of studies was evaluated and is shown in Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract. Studies mostly had moderate risk of bias. No study was excluded because of low quality assessment.

TNFi. *Effectiveness.* The relative risk of good and moderate EULAR response at 6 months (3 studies, n = 3,843) was significantly lower with non-MTX csDMARDs than with MTX combined with TNFi (RR 0.93 [95% CI 0.87, 1.0], $P = 0.04$; n = 3,843; $I^2 = 28\%$) (Figure 2A) (6–8). Most patients with non-MTX csDMARDs received LEF (n = 464 of 681). The mean difference of changed DAS28 score (delta DAS28) decreased less with non-MTX csDMARDs than with MTX (mean difference = 0.29 [95% CI 0.10, 0.49], $P = 0.003$; 4 studies; $I^2 = 0\%$, n = 2,075) (Figure 2B) (6,7,9,10). The relative risk of DAS28 remission at 6 months was lower with non-MTX csDMARDs than with MTX (RR 0.65 [95% CI 0.44, 0.96], $P = 0.03$; 2 studies; $I^2 = 0\%$, n = 1,927) (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract) (7,11).

For LEF in association with TNFi, the relative risk of good and moderate EULAR response was lower with LEF than with MTX (RR 0.93 [95% CI 0.90, 0.96], $P < 0.0001$; n = 5,620) (Figure 2C), with no heterogeneity (2 studies; $I^2 = 0\%$) (8,12). The mean difference of the delta DAS28 score was lower, but not significantly, with LEF than with MTX (mean difference = 0.32 [95% CI –0.05, 0.70], $P = 0.09$; n = 226; 2 studies; $I^2 = 0\%$) (Figure 2D) (9,10). No publication bias was found (Egger test: Figure 2A: $P = 0.642$; Figure 2B: $P = 0.094$; Figures 2C and 2D: not significant) (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract; $P = 0.05$).

Persistence rate. The relative risk of persistence at 12 months was lower with non-MTX csDMARDs than with MTX (RR 0.91 [95% CI 0.88, 0.95], $P < 0.0001$; 4 studies; $I^2 = 0\%$, n = 8,764) (see Supplementary Table 4A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract) (13–16). Most of the patients with non-MTX csDMARDs received LEF (1,264 of 1,448). As expected, for LEF, the relative risk of persistence was similarly lower with LEF–TNFi as compared with

Table 1. Characteristics of studies included in the meta-analysis*

Author, date (ref.)	Study type	Country	Period	Non-MTX csDMARDs, no.	csDMARDs	MTX, no.	bDMARDs	Evaluation date, months
Abatacept								
Alten et al, 2018 (19), post hoc analysis of Schiff et al, 2009	Post hoc analysis ARRIVE trial	International	2005–2007	176	LEF, AZA, SSZ, HCQ	490	ABA	6
Alten et al, 2018 (19), post hoc analysis of Weinblatt et al, 2006	Post hoc analysis ASSURE trial	International	2002–2004	143	LEF, AZA, SSZ, HCQ	535	ABA	6, 12
Alten et al, 2018 (19), post hoc analysis of Genovese et al, 2005	Post hoc analysis ATTAIN trial	International	2002–2004	42	LEF, AZA, SSZ, HCQ	175	ABA	6
Rituximab								
Narváez et al, 2011 (25)	Cohort	Spain	2007–2009	32	LEF	45	RTX	6
Chatzidionysiou et al, 2012 (24)	Cohort	European	NA	177	LEF	1,195	RTX	6, 12
Wendler et al, 2009 (27), abstract EULAR	Cohort	Germany	2006–2009	90	LEF	442	RTX	4
Soliman et al, 2012 (26)	Cohort	England	2006–2010	160	LEF, AZA, SSZ, HCQ, CsA, D-P, Gold, Min	343	RTX	6
Lukina et al, 2010 (28), abstract EULAR	Cohort	Russia	2006–2009	18	LEF	79	RTX	6
Tumor necrosis factor inhibitors								
Hyrich et al, 2006 (7): 1	Cohort	England	2001–2005	245	LEF, AZA, SSZ, HCQ, CsA, D-P, Gold, Min	250	ETA	6
Hyrich et al, 2006 (7): 2	Cohort	England	2001–2005	121	LEF, AZA, SSZ, HCQ, CsA, D-P, Gold	1,204	INF	6
Koyama et al, 2012 (6)	Cohort	Japan	2006–2010	12	SSZ	27	ETA	6
Iwamoto et al, 2009 (11)	Cohort	Japan	2009	47	LEF, TAC, SSZ, CsA, MZB, TP, D-P, BUC, Actarit	60	ETA	6
De Stefano et al, 2010 (9)	Cohort	Italy	2005–2007	60	LEF	60	TNFi (ETA, ADA, INF)	6
Benucci et al, 2011 (10)	Cohort	Italy	2010	42	LEF	54	TNFi (ADA, ETA)	6
Combe et al, 2014 (8)	Post hoc analysis	International	2009–2011	303	LEF	1,681	GOL	4, 6
Burmester et al, 2007 (12)	Post hoc analysis	International	2001–2007	842	LEF	2,794	ADA	3
Keystone et al, 2017 (18), abstract EULAR 2017	Post hoc analysis	International	2000–2002	60	SSZ, HCQ, LEF	114	ADA	6
Strangfeld et al, 2009 (16): 1	Cohort	Germany	2001–2006	76	LEF	361	INF	6
Strangfeld et al, 2009 (16): 2	Cohort	Germany	2001–2006	144	LEF	448	ETA	6
Strangfeld et al, 2009 (16): 3	Cohort	Germany	2001–2006	174	LEF	566	ADA	6
Soliman et al, 2011 (15)	Cohort	England	2001–2009	610	LEF	4,418	TNFi (ETA, ADA, INF)	12
Finckh et al, 2008 (13)	Cohort	Switzerland	1996–2006	116	LEF	843	TNFi (ETA, ADA, INF)	6, 12
Nordstrom et al, 2006 (17)	Cohort	Finland	1999–2004	44	LEF, AZA, HCQ, SSZ, ATM	44	INF	3, 6
Kristensen et al, 2006 (14): 1	Cohort	Sweden	1999–2004	116	LEF, AZA, SSZ, HCQ	501	INF	12
Kristensen et al, 2006 (14): 2	Cohort	Sweden	1999–2004	68	LEF, AZA, SSZ, HCQ	179	ETA	12

* ABA = abatacept; ADA = adalimumab; ATM = aurothiomalate; AZA = azathioprine; bDMARD = biologic disease-modifying antirheumatic drug; BUC = bucillamine; CsA = cyclosporin A; csDMARD = conventional synthetic DMARD; D-P = D-penicillamine; ETA = etanercept; EULAR = European Alliance of Associations for Rheumatology; GOL = golimumab; HCQ = hydroxychloroquine; INF = infliximab; LEF = leflunomide; Min = minocycline; MTX = methotrexate; MZB = mizoribine; NA = not applicable; RTX = rituximab; SSZ = sulfasalazine; TAC = tacrolimus; TNFi = tumor necrosis factor inhibitor; TP = tiopronin.

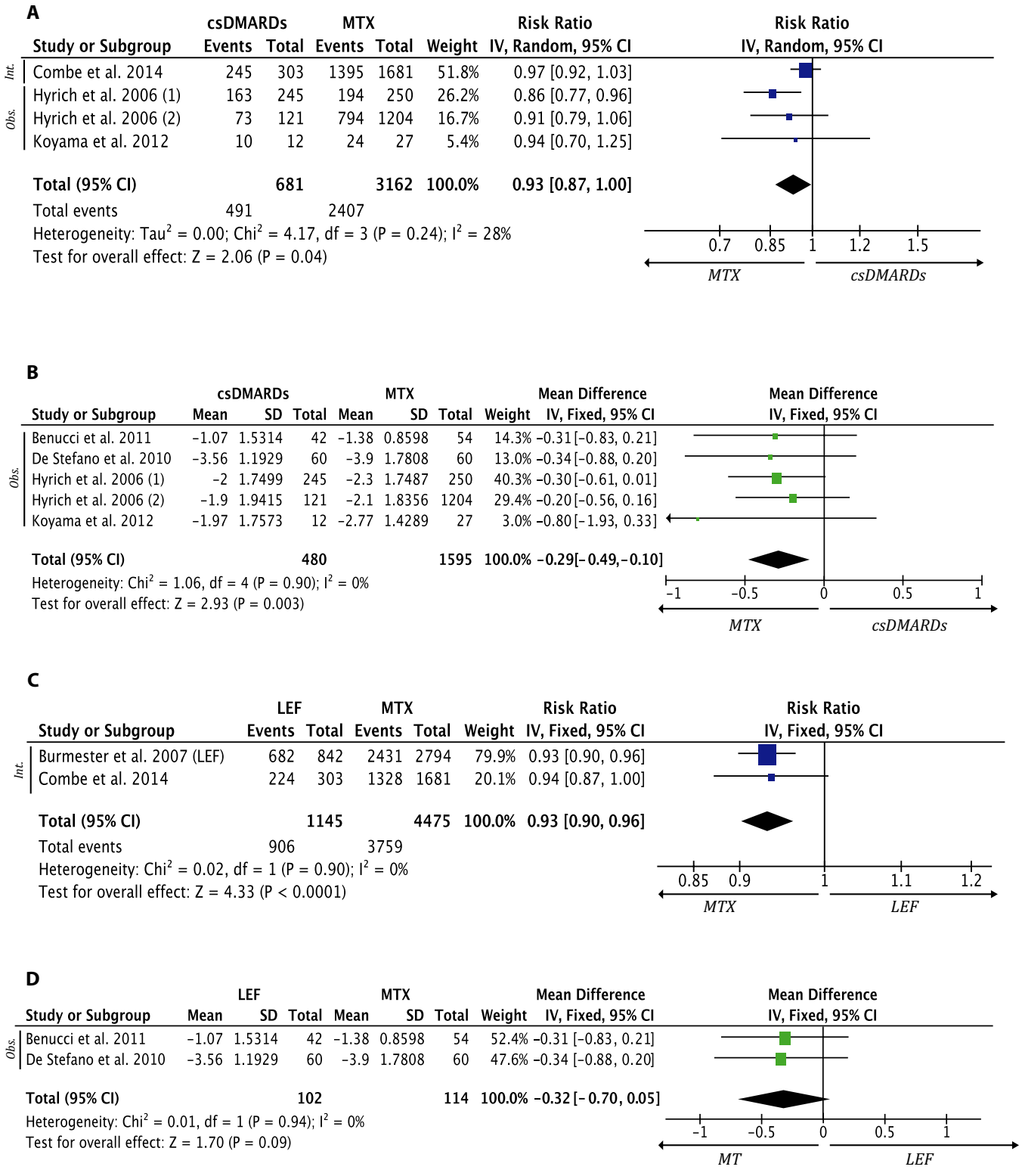

Figure 2. Effectiveness for tumor necrosis factor inhibitor (TNFi) with non-methotrexate (MTX) conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) or leflunomide (LEF) versus MTX–TNFi. **A**, Risk ratio of good and moderate European Alliance of Associations for Rheumatology (EULAR) response at 6 months for csDMARDs; **B**, Mean difference of delta Disease Activity Score in 28 joints (DAS28) at 6 months for csDMARDs; **C**, Risk ratio of good and moderate EULAR response at 6 months for LEF; and **D**, Mean difference of delta DAS28 at 6 months for LEF. 95% CI = 95% confidence interval; Int = interventional study (post hoc analysis); Obs = observational study (cohort). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract.

MTX–TNFi (RR 0.91 [95% CI 0.87, 0.95], $P < 0.0001$; n = 7,900; 3 studies; $I^2 = 0\%$) (see Supplementary Table 4B, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract) (13,15,16). We found no publication bias (Egger test: Supplementary Table 4A, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract: $P = 0.529$; Supplementary Table 4BB, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract: $P = 0.45$). Fewer data were available for the long-term. In the Strangfeld et al study, discontinuation rates at 36 months were higher with LEF than with MTX, with 53.4%, 63.1%, and 67.1% for etanercept, ADA, and INF, respectively, in combination with LEF, and 46.3%, 51.3%, and 61.5% with MTX (16).

Safety. The risk of SAEs was higher with non-MTX csDMARDs than with MTX (RR 1.63 [95% CI 1.32, 2.01], $P < 0.00001$; 6 studies; $I^2 = 0\%$, n = 5,995) (Figure 3A). In the non-MTX csDMARDs group, 1,133 of 1,442 patients took LEF (7,9,11,12,17,18). Similarly, we found an increased risk with LEF versus MTX in association with TNFi (RR 1.74 [95% CI 1.33, 2.28], $P < 0.0001$; 2 studies; $I^2 = 0\%$, n = 3,756) (Figure 3B) (9,12). A meta-analysis of total AEs was not possible because of lack of data. No publication bias was found (Egger test: Figure 3A: $P = 0.381$; Figure 3B: not significant).

ABA. *Effectiveness.* We found only 2 studies with a delta DAS28 score and 3 with change in HAQ DI. Alten et al (19) recently retrieved post hoc data from 3 clinical trials and 1 observational study (20–23). The 2 groups (non-MTX csDMARDs and MTX groups) did not differ in the delta DAS28 score (mean difference = 0.08 [95% CI –0.12, 0.28], $P = 0.43$; $I^2 = 0\%$, n = 883) (Figure 4A) or change in HAQ DI (mean difference = 0.01 [95% CI –0.04, 0.07], $P = 0.59$; $I^2 = 0\%$, n = 1,561) (Figure 4B). LEF represented 35% of the non-MTX csDMARDs group (79 of 218 patients) for the delta DAS28 score and 40% (144 of 361 patients) for change in HAQ DI. For the delta DAS28 score and HAQ DI change, respectively, other csDMARDs were SSZ (42 of 218 [19.3%] and 78 of 361 [21.6%]), hydroxychloroquine (60 of 218 [27.5%] and 95 of 361 [26.3%]) and azathioprine (34 of 218 [15.6%] and 44 of 361 [12.2%]). No publication bias was found (Egger test: Figure 4A: not significant, Figure 4B: $P = 0.382$). Long-term results at 12 and 24 months in the ACTION study did not show any differences in DAS28 with the C-reactive protein (CRP) level or change in HAQ DI for ABA combined with non-MTX csDMARDs versus MTX (23).

Safety. We found no difference in the risk of total AEs between non-MTX csDMARDs–ABA and MTX–ABA (RR 0.96 [95% CI 0.85, 1.09], $P = 0.53$; 3 studies; $I^2 = 11\%$, n = 1,561) (Figure 4C) (20–22). The most common AEs were infections, followed by skin and subcutaneous tissue disorders (no difference). A lower number of patients was included as compared with the analysis with TNFi (n = 1,561 versus n = 5,995 for TNFi). We found no differences in the risk of SAEs (RR 0.96 [95% CI 0.21, 4.46], $P = 0.96$; 3 studies; $I^2 = 59\%$, n = 1,561) (Figure 4D), but with high heterogeneity between the 3 studies because of the small sample sizes. Of the 361 patients with non-MTX csDMARDs, LEF represented 40%

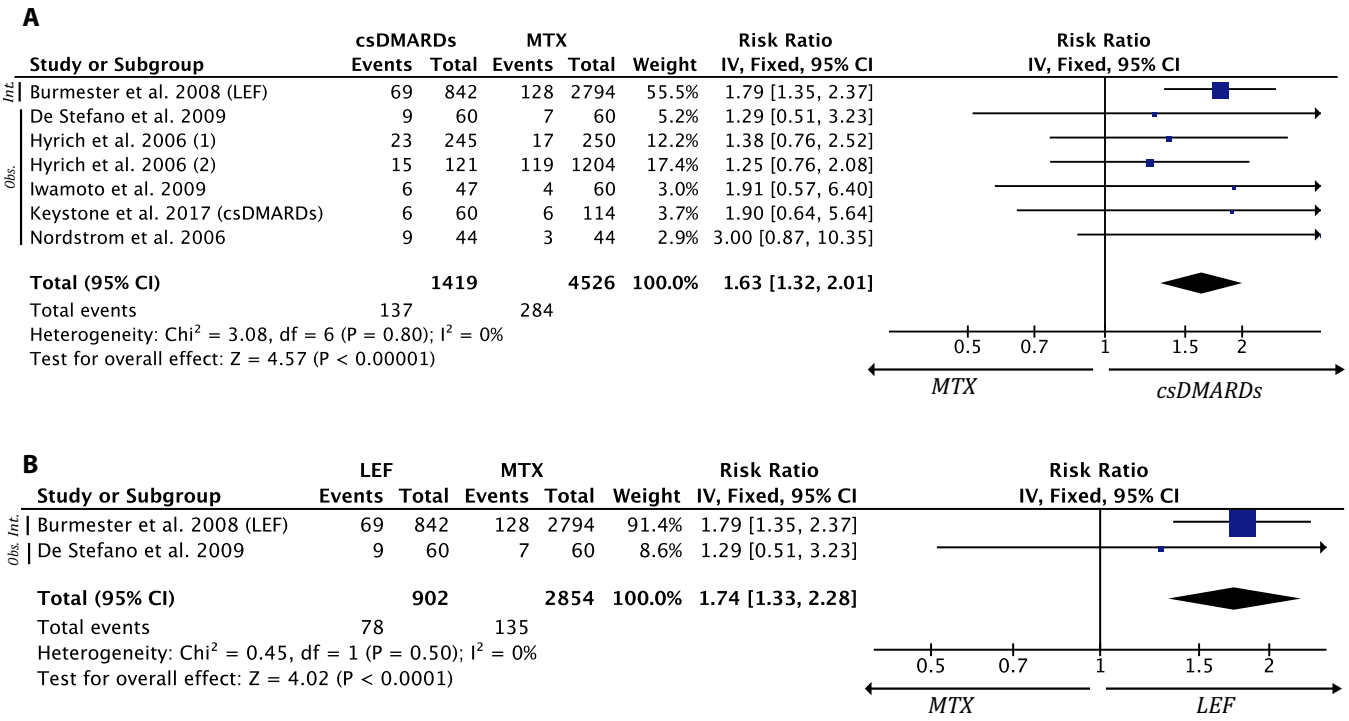

Figure 3. Serious adverse events rate for tumor necrosis factor inhibitor (TNFi) with non-methotrexate (MTX) conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) or leflunomide (LEF) versus MTX–TNFi. **A**, Non-MTX csDMARDs versus MTX; **B**, LEF versus MTX. 95% CI = 95% confidence interval; Int = interventional study (post hoc analysis); Obs = observational study (cohort). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract.

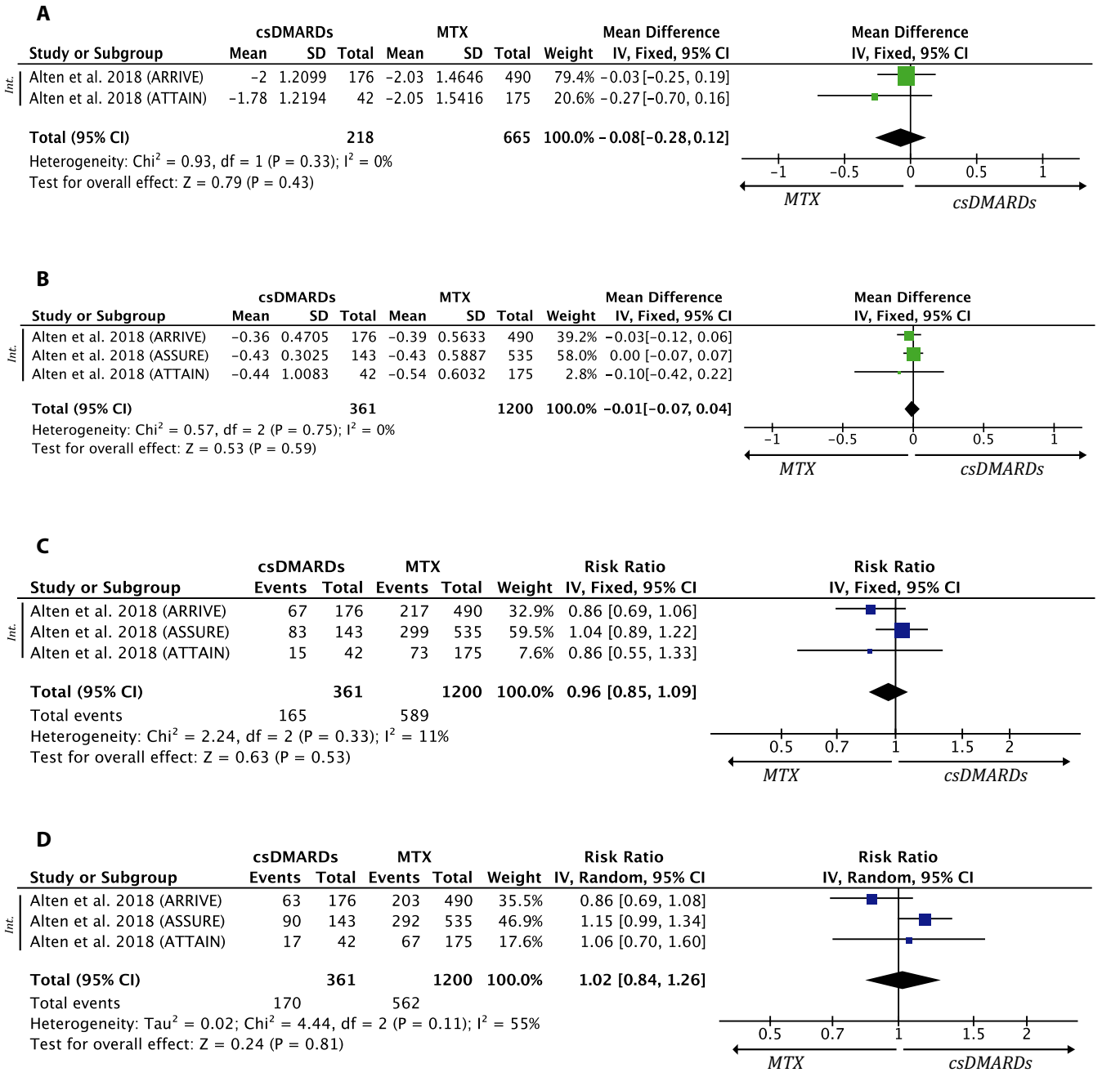

Figure 4. Effectiveness and safety at 6 months for abatacept (ABA) with non-methotrexate (MTX) conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) versus MTX–ABA. **A**, Mean difference of delta Disease Activity Score in 28 joints; **B**, Change in Health Assessment Questionnaire disability index score; **C**, Risk ratio of total adverse events rate; and **D**, Risk ratio of serious adverse events rate. 95% CI = 95% confidence interval; Int = interventional study (post hoc analysis). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract.

of patients (144 of 361). Other csDMARDs were SSZ (78 of 361 [21.6%]), hydroxychloroquine (95 of 361 [26.3%]), and azathioprine (44 of 361 [12.2%]). No publication bias was found (Egger test: Figure 4C: $P = 0.728$, Figure 4D: $P = 0.394$).

RTX. *Effectiveness.* The relative risk of good EULAR response at 6 months was higher with non-MTX csDMARDs than with MTX (RR 1.30 [95% CI 1.09, 1.55], $P = 0.004$; 5 studies; $I^2 = 0\%$, n = 2,581) (Figure 5A) (24–28). The delta DAS28 score did not differ between the 2 groups (mean difference = 0.10 [95% CI –0.06, 0.26], $P = 0.21$; 3 studies; $I^2 = 14\%$, n = 1,952) (see Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract) (24–26). For patients receiving LEF (317 of 317 for 3 studies, and not applicable of 160 for Soliman et al) combined with RTX, the relative risk of good EULAR response

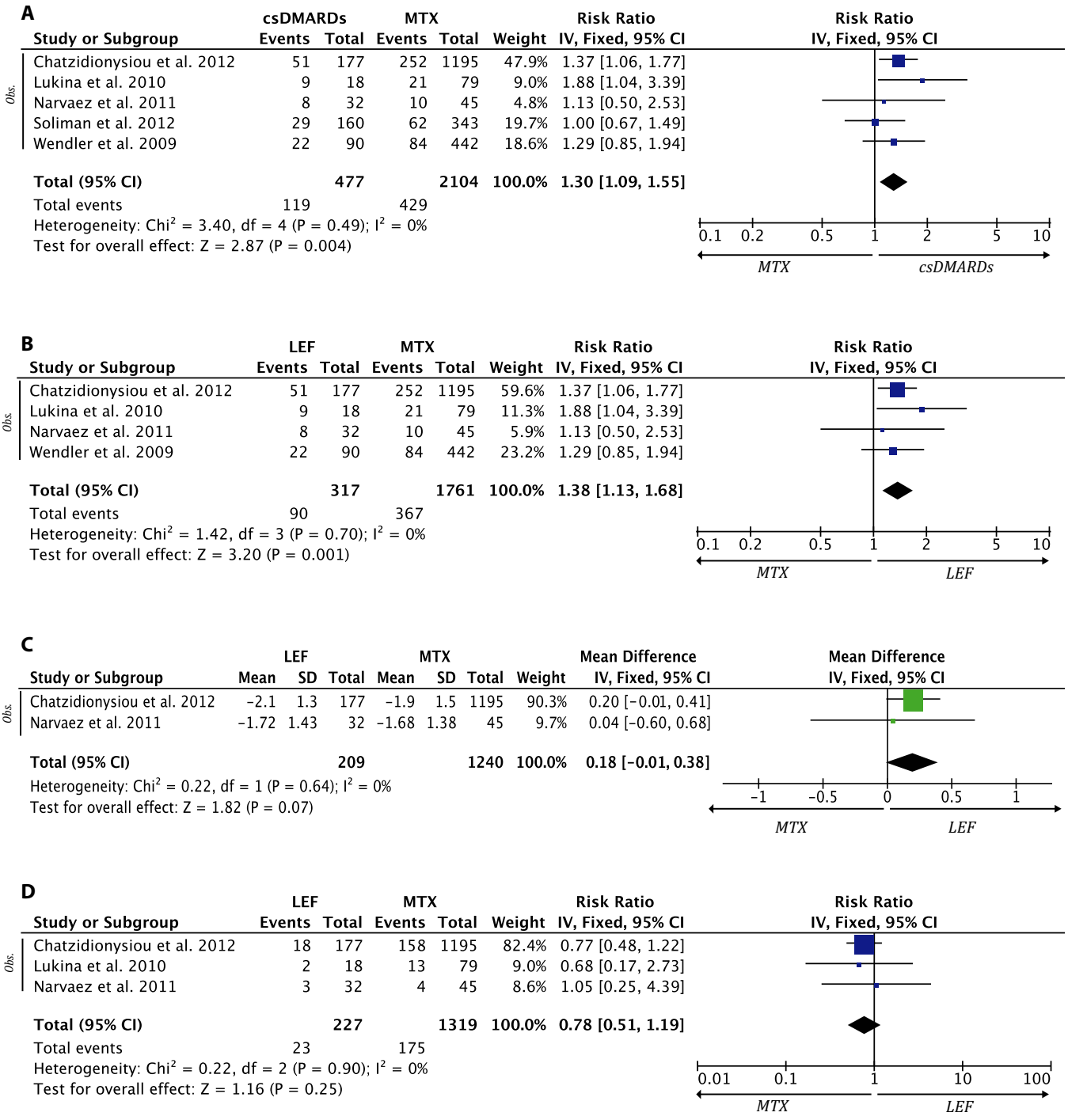

Figure 5. Effectiveness or safety for rituximab (RTX) with non-methotrexate (MTX) conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) or leflunomide (LEF) versus MTX–RTX. **A**, Risk ratios of good European Alliance of Associations for Rheumatology (EULAR) response at 6 months for csDMARDs; **B**, Risk ratios of good EULAR response at 6 months for LEF; **C**, Mean difference of delta Disease Activity Score in 28 joints at 6 months for LEF; and **D**, Risk ratios of serious adverse event rate for LEF. 95% CI = 95% confidence interval; Obs = observational study (cohort). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract.

was 38% higher at 6 months with LEF than with MTX (RR 1.38 [95% CI 1.13, 1.68], $P = 0.001$; 4 studies; $I^2 = 0\%$, n = 2,078) (Figure 5B) (24,25,27,28). Moreover, the delta DAS28 score was decreased more with LEF than with MTX but not significantly (mean difference = 0.18 [95% CI –0.01, 0.38], $P = 0.07$; 2 studies; $I^2 = 0\%$, n = 1,449) (Figure 5C) (24,25). No publication bias was found (Egger test: Figure 5A: $P = 0.878$; Figure 5B: $P = 0.834$; Figure 5C: not significant; Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24195/abstract): $P = 0.715$).

RTX retreatment. In the Richter et al study, second treatment with RTX was 62.4% in the LEF group and 55.6% in the MTX group (29). In the CERRERA registry, retreatment was lower during the first 12 months in the LEF than in the MTX group (21.5% versus 31.9%) (24).

Safety. The LEF–RTX and MTX–RTX groups did not differ in the risk of AEs (RR 0.78 [95% CI 0.51, 1.19], $P = 0.25$; 3 studies, n = 1,546) (Figure 5D) (24,25). One study represented 82.4% of the weight of this analysis (24), but results were similar in the 3 studies, without heterogeneity ($I^2 = 0\%$). No publication bias was found (Egger test: Figure 5D: $P = 0.77$). No difference in the SAE rate was found.

TCZ. *Effectiveness.* Three articles were retrieved, but the heterogeneity of outcome measures did not allow for meta-analysis. In the Genovese et al study, ACR 20% improvement (ACR20) at 6 months represented 51 of 78 patients (65.4%) and 269 of 456 patients (59%), respectively, for TCZ–LEF and TCZ–MTX (30). In the Narváez et al study, ACR 50% improvement (ACR50) at 6 months represented 11 of 26 patients (42%) in the LEF group and 24 of 55 (44%) in the MTX group (31). EULAR response did not differ, with good EULAR response in 16 of 26 patients (62%) versus 34 of 55 (62%) with LEF and MTX, and good-moderate EULAR response in 23 of 26 (88%) versus 51 of 55 (93%). No differences were found in the mean delta DAS28 score (2.17 ± 1.43 versus 2.23 ± 1.38) and for DAS28 remission (<2.6: 11 of 26 (42%) versus 24 of 55 (44%) for LEF versus MTX). Inanc et al observed a similar rate of remission between TCZ with LEF and with MTX (42% versus 35%, $P > 0.05$) in the TURKBIO registry at 6 months (32).

Persistence rate. In the Jones et al study, the persistence rate at 12 months was 75 of 97 (77%) and 99 of 139 (71%) with non-MTX csDMARDs and MTX, respectively (33). No long-term data were available.

Safety. Narváez et al did not find any differences in total AEs (17 of 29 [59%] and 32 of 62 [52%] with LEF and MTX, respectively, $P = 0.65$) or SAEs (3 of 29 [10%] and 7 of 62 [11%], respectively, $P = 0.89$) (31). In another study, the AE rate was 52 of 78 (67%) and 332 of 456 (73%) with LEF and MTX (30).

JAK inhibitors. *Effectiveness.* Only 2 studies for tofacitinib and 1 for baricitinib were available to compare the effectiveness of combining non-MTX csDMARDs and MTX. In ORAL SYNC, the ACR20 response with tofacitinib 5 mg twice a day was similar with LEF or MTX (70.6% versus 71.7%), but a small number of patients was included (37 versus 159) (34). We found no difference in the mean change in HAQ score at month 3 (–0.53 and –0.45, respectively). In the TURKBIO registry, the level of remission and low disease activity rate at 6 months were lower with LEF than with MTX (21% versus 42%) combined with tofacitinib, but the difference was not statistically significant because of the small sample size of patients with tofacitinib in this registry (n = 33). No baseline characteristics were available to compare the 2 groups (32). In the RA-BUILD RCT evaluating baricitinib (4 mg per day), ACR20 at 6 months was 53% (23 of 43) versus 67% (76 of 114) with non-MTX csDMARDs and MTX. The ACR50 was 40% (17 of 43) versus 42% (48 of 114) and DAS28-CRP <3.2 (remission plus low disease activity) was obtained in 44% (19 of 43) versus 52% (59 of 114) (35).

Safety. In the ORAL SYNC study, safety did not differ between combining tofacitinib 5 mg twice a day with MTX or LEF, but conclusions were limited due to the small sample size in each group (159 versus 37) (34). No data were available for baricitinib.

DISCUSSION

In this study we compared the safety and efficacy of MTX and other csDMARDs combined with advanced therapies. As expected, MTX was associated with better outcomes compared to other csDMARDs, including LEF, when associated with TNFi, although the difference was modest: 7% higher EULAR response, a difference of (baseline minus 6 months) delta DAS28 score in favor of the MTX group of 0.3 at 6 months and a 7% higher persistence rate compared to non-MTX csDMARDs. MTX is thus a good option to use when possible in association with TNFi. Conversely, LEF seemed superior to MTX when combined with RTX, with 38% higher EULAR response and similar tolerance. For association with ABA, we found no difference between non-MTX csDMARDs and MTX for effectiveness or AE rate. Similarly for association with TCZ, no difference was found between non-MTX csDMARDs and MTX, although the literature is sparse. For baricitinib, the small number of patients included in the csDMARDs group and the lack of safety data preclude conclusions.

The superiority of LEF over MTX combined with RTX might be explained by a possible synergistic effect between the 2 drugs. LEF inhibits de novo pyrimidine synthesis by targeting the cellular dihydroorotate dehydrogenase, thereby decreasing T- and B-cell proliferation in vitro (36,37). However, LEF was also shown as able to interfere with the JAK/STAT pathway and to inhibit interleukin (IL)-17 and TNF (38), which could reinforce TNFi effects. LEF also inhibits B-cell proliferation in vivo, inducing B-cell apoptosis in chronic lymphocytic leukemia (39). Thus, LEF could reinforce the effect of RTX on B cells.

However, both MTX and LEF could potentiate TNFi effects. Indeed, MTX may inhibit IL-6 production by macrophages, which could facilitate a synergistic effect with TNFi (40). Many studies have demonstrated that MTX decreases the formation of antidrug antibodies in RA (41). Conversely, the effect of LEF on immunogenicity has been poorly studied, and we lack data on a direct comparison of MTX and LEF. Dénarié et al explored TNF bioactivity in RA patients receiving INF: 39 women with RA and active disease

despite csDMARDs received INF combined with csDMARDs (24 for MTX, 13 for LEF, and 2 for hydroxychloroquine). At 22 weeks, the level of TNF bioactivity and antibody toward INF content were lower with MTX than with other csDMARDs (42).

LEF was the most common drug in the non-MTX csDMARDs group. SSZ was less often used Specifically analyzing patients with SSZ was not possible due to the small number of patients. Addressing the effect of SSZ would require further studies. For hydroxychloroquine, other data are needed, and this meta-analysis could not answer this question because of the small number of patients receiving hydroxychloroquine. Patients included in registries received other csDMARDs such as bucillamine, azathioprine, D penicillamine, gold salts, and tacrolimus, but we lack data to recommend their use in clinical practice. In recent international guidelines, these drugs were not recommended in combination with advanced therapies (1,2).

This work has some limitations. Studies included were mostly observational (cohort studies). The choice of drugs probably depended on confounding factors that were not taken into account in the analysis, which might be a major bias. However, the choice of drugs does not appear to be explained by differences in baseline characteristics of patients. The only numerical difference was the use of glucocorticoids, with higher use in the MTX versus the non-MTX csDMARDs groups (96% versus 84%). Baseline disease activity (identical DAS28 score) or severity (age, sex, rheumatoid factor, ACPAs) were similar between patients receiving LEF or MTX. For the RTX analysis, the study by Chatzidionysiou et al represented a large part of the weight, but results were comparable among studies, without heterogeneity ($I2 = 0\%$) (24).

Wendler et al found a similar trend for mean difference of the delta DAS28 score of approximately 0.10 between LEF and MTX associated with RTX, which was approximately the mean difference we showed (43). This study could not be included in our meta-analysis because of lack of confidence intervals in the article. Another limitation of our study was lack of data for structural efficacy of advanced therapies in combination with MTX or non-MTX csDMARDs. These data would be important to assess before claiming a potential superiority of LEF over MTX in combination with RTX.

Another bias could be that MTX is the first choice for most patients with RA. Therefore, other csDMARDs are used as second or third DMARDs when MTX has failed. However, looking at baseline characteristics of the included studies, we found that the number of anterior previous csDMARDs, when data were available, did not differ between MTX and the non-MTX csDMARDs group, in association with RTX: 2.1 ± 0.36 in the non-MTX csDMARDs group and 2.16 ± 0.37 in the MTX csDMARDs group. For TNFi, the number of anterior previous csDMARDs was 3.84 ± 0.92 in the non-MTX csDMARDs group and 3.14 ± 0.86 in the MTX group. In addition, most registries were European, from different countries, which could limit the generalizability, especially for Asian populations. The mean dose of MTX was 15 mg/week, and most patients received 20 mg/day of LEF, which represents optimal doses in routine practice. The MTX dose was usually lower for Japanese patients than for Europeans, and in 1 study that was included, the mean MTX dose in Japanese RA patients was approximately 6 mg/week (6).

In conclusion, non-MTX csDMARDs, especially LEF, are safe and efficient alternatives to MTX in association with advanced therapies. In this work, LEF appeared to be superior to MTX when combined with RTX. RCTs should be conducted to confirm whether the LEF–RTX combination is superior to MTX–RTX.

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. Daien 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. Decarriere, Combe, Gaujoux-Viala, Lukas, Morel, Daien.

Acquisition of data. Decarriere, Barnetche, Daien.

Analysis and interpretation of data. Decarriere, Barnetche, Combe, Lukas, Morel, Daien.

ROLE OF THE STUDY SPONSOR

AbbVie France 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 AbbVie France.

REFERENCES

1. Smolen JS, Landewé R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Ann Rheum Dis 2017;76:960–77.
2. Combe B, Landewe R, Daien CI, Hua C, Aletaha D, Álvaro-Gracia JM, et al. 2016 update of the EULAR recommendations for the management of early arthritis. Ann Rheum Dis 2017;76:948–59.
3. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. Lancet 2004;363:675–81.
4. Golicki D, Newada M, Lis J, Pol K, Hermanowski T, Tłustochowicz M. Leflunomide in monotherapy of rheumatoid arthritis: meta-analysis of randomized trials. Pol Arch Med Wewn 2012;122:22–32.
5. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.
6. Koyama Y, Shiraishi H, Ohta T, Uchino A. Etanercept in combination with conventional disease-modifying antirheumatic drugs (DMARDs) in the treatment of rheumatoid arthritis patients intolerant to methotrexate. Mod Rheumatol 2012;22:100–8.
7. Hyrich KL, Symmons DP, Watson KD, Silman AJ, on behalf of the British Society for Rheumatology Biologics Register. Comparison of the response to infliximab or etanercept monotherapy with the response to cotherapy with methotrexate or another disease-modifying antirheumatic drug in patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. Arthritis Rheum 2006;54:1786–94.

8. Combe B, Dasgupta B, Louw I, Pal S, Wollenhaupt J, Zerbini CA, et al. Efficacy and safety of golimumab as add-on therapy to disease-modifying antirheumatic drugs: results of the GO-MORE study. Ann Rheum Dis 2014;73:1477–86.
9. De Stefano R, Frati E, Nargi F, Baldi C, Menza L, Hammoud M, et al. Comparison of combination therapies in the treatment of rheumatoid arthritis: leflunomide-anti-TNF-alpha versus methotrexate-anti-TNF-alpha. Clin Rheumatol 2010;29:517–24.
10. Benucci M, Saviola G, Baiardi P, Manfredi M, Sarzi-Puttini P, Atzeni F. Efficacy and safety of leflunomide or methotrexate plus subcutaneous tumour necrosis factor-alpha blocking agents in rheumatoid arthritis. Int J Immunopathol Pharmacol 2011;24:269–74.
11. Iwamoto N, Kawakami A, Fujikawa K, Aramaki T, Kawashiri SY, Tamai M, et al. Prediction of DAS28-ESR remission at 6 months by baseline variables in patients with rheumatoid arthritis treated with etanercept in Japanese population. Mod Rheumatol 2009;19:488–92.
12. Burmester GR, Mariette X, Montecucco C, Monteagudo-Sáez I, Malaise M, Tzioufas AG, et al. Adalimumab alone and in combination with disease-modifying antirheumatic drugs for the treatment of rheumatoid arthritis in clinical practice: the Research in Active Rheumatoid Arthritis (ReAct) trial. Ann Rheum Dis 2007;66:732–9.
13. Finckh A, Dehler S, Gabay C, on behalf of the SCQM doctors. The effectiveness of leflunomide as a co-therapy of tumour necrosis factor inhibitors in rheumatoid arthritis: a population-based study. Ann Rheum Dis 2008;68:33–9.
14. Kristensen LE, Saxne T, Nilsson JA, Geborek P. Impact of concomitant DMARD therapy on adherence to treatment with etanercept and infliximab in rheumatoid arthritis: results from a six-year observational study in southern Sweden. Arthritis Res Ther 2006;8:R174.
15. Soliman MM, Ashcroft DM, Watson KD, Lunt M, Symmons DP, Hyrich KL, et al. Impact of concomitant use of DMARDs on the persistence with anti-TNF therapies in patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. Ann Rheum Dis 2011;70:583–9.
16. Strangfeld A, Hierse F, Kekow J, von Hinueber U, Tony HP, Dockhorn R, et al. Comparative effectiveness of tumour necrosis factor alpha inhibitors in combination with either methotrexate or leflunomide. Ann Rheum Dis 2009;68:1856–62.
17. Nordström DC, Konttinen L, Korpela M, Tiippana-Kinnunen T, Eklund K, Forsberg S, et al. Classic disease modifying anti-rheumatic drugs (DMARDs) in combination with infliximab: the Finnish experience. Rheumatol Int 2006;26:741–8.
18. Keystone E, Suboticki J, Griffith J, Zhang Y, Kremer J. Adalimumab in combination with non-methotrexate conventional synthetic disease modifying rheumatic drugs in a clinical trial setting [abstract]. Ann Rheum Dis 2017;76 Suppl 2:828.
19. Alten R, Burkhardt H, Feist E, Krüger K, Rech J, Rubbert-Roth A, et al. Abatacept used in combination with non-methotrexate disease-modifying antirheumatic drugs: a descriptive analysis of data from interventional trials and the real-world setting. Arthritis Res Ther 2018;20:1.
20. Weinblatt M, Combe B, Covucci A, Aranda R, Becker JC, Keystone E. Safety of the selective costimulation modulator abatacept in rheumatoid arthritis patients receiving background biologic and nonbiologic disease-modifying antirheumatic drugs: a one-year randomized, placebo-controlled study. Arthritis Rheum 2006;54:2807–16.
21. Genovese MC, Becker JC, Schiff M, Luggen M, Sherrer Y, Kremer J, et al. Abatacept for rheumatoid arthritis refractory to tumor necrosis factor α inhibition. N Engl J Med 2005;353:1114–23.
22. Schiff M, Pritchard C, Huffstutter JE, Rodriguez-Valverde V, Durez P, Zhou X, et al. The 6-month safety and efficacy of abatacept in patients with rheumatoid arthritis who underwent a washout after anti-tumour necrosis factor therapy or were directly switched to abatacept: the ARRIVE trial. Ann Rheum Dis 2009;68:1708–14.
23. Nüßlein HG, Alten R, Galeazzi M, Lorenz HM, Boumpas D, Nurmohamed MT, et al. Real-world effectiveness of abatacept for rheumatoid arthritis treatment in European and Canadian populations: a 6-month interim analysis of the 2-year, observational, prospective ACTION study. BMC Musculoskelet Disord 2014;15:14.
24. Chatzidionysiou K, Lie E, Nasonov E, Lukina G, Hetland ML, Tarp U, et al. Effectiveness of disease-modifying antirheumatic drug co-therapy with methotrexate and leflunomide in rituximab-treated rheumatoid arthritis patients: results of a 1-year follow-up study from the CERERRA collaboration. Ann Rheum Dis 2012;71:374–7.
25. Narváez J, Díaz-Torné C, Ruiz JM, Hernández MV, Torrente-Segarra V, Ros S, et al. Comparative effectiveness of rituximab in combination with either methotrexate or leflunomide in the treatment of rheumatoid arthritis. Semin Arthritis Rheum 2011;41:401–5.
26. Soliman MM, Hyrich KL, Lunt M, Watson KD, Symmons DP, Ashcroft DM, et al. Effectiveness of rituximab in patients with rheumatoid arthritis: observational study from the British Society for Rheumatology Biologics Register. J Rheumatol 2012;39:240–6.
27. Wendler J, Sørensen H, Tony H, Richter C, Krause A, Rubbert-Roth A, et al. Effectiveness and safety of rituximab (RTX) monotherapy compared to RTX combination therapy with methotrexate or leflunomide in the German RTX-treatment of active rheumatoid arthritis in daily practice trial [abstract]. Ann Rheum Dis 2009;68:76.
28. Lukina G. Rituximab combinations in rheumatoid arthritis patients : methotrexate versus leflunomide [abstract]. Ann Rheum Dis 2010;69 Sup 3:536.
29. Richter A, Strangfeld A, Herzer P, Wilden E, Bussmann A, Listing J, et al. Sustainability of rituximab therapy in different treatment strategies: results of a 3-year followup of a German biologics register. Arthritis Care Res (Hoboken) 2014;66:1627–33.
30. Genovese MC, McKay JD, Nasonov EL, Mysler EF, da Silva NA, Alecock E, et al. Interleukin-6 receptor inhibition with tocilizumab reduces disease activity in rheumatoid arthritis with inadequate response to disease-modifying antirheumatic drugs: the tocilizumab in combination with traditional disease-modifying antirheumatic drug therapy study. Arthritis Rheum 2008;58:2968–80.
31. Narváez J, Díaz-Torné C, Magallares B, Hernández MV, Reina D, Corominas H, et al. Comparative effectiveness of tocilizumab with either methotrexate or leflunomide in the treatment of rheumatoid arthritis. PloS One 2015;10:e0123392.
32. Inanc N, Ozen G, Yalçınkaya Y, Dalkilic E, Koca SS, Can G, et al. Is there any difference in RA patients for methotrexate use vs. leflunomide use as a concomitant treatment with biological and targeted synthetic DMARDs in Turkbio Registry? [abstract]. Arthritis Rheumatol 2017;69 Suppl 10.
33. Jones G, Hall S, Bird P, Littlejohn G, Tymms K, Youssef P, et al. A retrospective review of the persistence on bDMARDs prescribed for the treatment of rheumatoid arthritis in the Australian population. Int J Rheum Dis 2018;21:1581–90.
34. Kremer J, Li ZG, Hall S, Fleischmann R, Genovese M, Martin-Mola E, et al. Tofacitinib in combination with nonbiologic disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis: a randomized trial. Ann Intern Med 2013;159:253–61.
35. Dougados M, van der Heijde D, Chen YC, Greenwald M, Drescher E, Liu J, et al. Baricitinib in patients with inadequate response or intolerance to conventional synthetic DMARDs: results from the RA-BUILD study. Ann Rheum Dis 2017;76:88–95.
36. Greene S, Watanabe K, Braatz-Trulson J, Lou L. Inhibition of dihydroorotate dehydrogenase by the immunosuppressive agent leflunomide. Biochem Pharmacol 1995;50:861–7.

37. Siemasko K, Chong AS, Jäck HM, Gong H, Williams JW, Finnegan A. Inhibition of JAK3 and STAT6 tyrosine phosphorylation by the immunosuppressive drug leflunomide leads to a block in IgG1 production. J Immunol 1998;160:1581–8.
38. González-Alvaro I, Ortiz AM, Domínguez-Jiménez C, Aragón-Bodi A, Díaz Sánchez B, Sánchez-Madrid F. Inhibition of tumour necrosis factor and IL-17 production by leflunomide involves the JAK/STAT pathway. Ann Rheum Dis 2009;68:1644–50.
39. Ringshausen I, Oelsner M, Bogner C, Peschel C, Decker T. The immunomodulatory drug leflunomide inhibits cell cycle progression of B-CLL cells. Leukemia 2008;22:635–8.
40. Municio C, Dominguez-Soto Á, Fuentelsaz-Romero S, Lamana A, Montes N, Cuevas VD, et al. Methotrexate limits inflammation through an A20-dependent cross-tolerance mechanism. Ann Rheum Dis 2018;77:752–9.
41. Breedveld FC, Weisman MH, Kavanaugh AF, Cohen SB, Pavelka K, van Vollenhoven R, et al. The PREMIER study: a multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. Arthritis Rheum 2006;54:26–37.
42. Dénarié D, Rinaudo-Gaujous M, Thomas T, Paul S, Marotte H. Methotrexate reduced TNF bioactivity in rheumatoid arthritis patients treated with infliximab. Mediators Inflamm 2017;2017:1–7.
43. Wendler J, Burmester GR, Sörensen H, Krause A, Richter C, Tony HP, et al. Rituximab in patients with rheumatoid arthritis in routine practice (GERINIS): six-year results from a prospective, multicentre, non-interventional study in 2,484 patients. Arthritis Res Ther 2014;16:R80.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 885–892
DOI 10.1002/acr.24189

American College of Rheumatology
Empowering Rheumatology Professionals

Fatherhood Experiences of Men With Inflammatory Arthritis: A Preliminary Grounded Theory

Michal Avrech Bar,[1] Thao T. Dao,[2] Leah R. DeBlock Vlodarchyk,[3] and Catherine L. Backman[4]

Objective. Minimal prior research has examined the impact of inflammatory arthritis (IA) on men's perspectives on parenting. We aimed to describe fathering roles and experiences, the effect of IA on parenting activities, and strategies used by fathers with IA to fulfill this role.

Methods. A grounded theory approach guided data gathering and analysis. Nine men with IA, parenting at least 1 child age <19 years, were recruited through rheumatology practices, therapy clinics, and social media. Each engaged in 1 in-depth personal interview. Transcripts were analyzed using inductive and iterative steps to identify key themes and a preliminary explanatory framework of fathering experiences of men with IA.

Results. All men were married, ages 31–62 years, with 1 to 5 children ages 6 months to 28 years. "Being an involved father" describes participants' perspectives on fulfilling their role as hands-on parents, role models, and financial providers. "Taking ownership" explains how participants managed daily life, comprising 2 subthemes, "taking care of yourself," using strategies like exercise and communicating with loved ones, and "redefining yourself," a process of adapting to reframed identity and lifestyle adjustments. "Accessing support" indicates men who felt well-supported by social networks (most critically their wives), health care providers, and informational and educational resources.

Conclusion. This small, grounded theory study offers an enriched understanding of fatherhood experiences of men with IA. When social, practical, and educational supports are in place, these men found parenting joyful and rewarding. Despite task limitations, their perspectives on being involved fathers was unrestricted by arthritis.

INTRODUCTION

From the time of industrialization, the primary focus of fathers has been economic support for the family (1), while child-rearing has rested primarily on mothers. However, the rise of women joining the workforce shifted family roles and paternal involvement has evolved (2–3). Contemporary fathers are increasingly described as equal partners in parenting (3–4), as care providers, protectors, models, moral guides, and breadwinners (1). Several studies have described men as integral to domestic and child-rearing activities, taking more responsibility for organizing and planning their children's lives (2,5–7), yet statistics show this shifting responsibility remains gradual, with men still providing the majority of household income and doing less household work and child care than women (8).

Less is known about the fathering experiences of men with arthritis. Inflammatory types of arthritis (e.g., rheumatoid arthritis [RA], ankylosing spondylitis [AS]) are chronic autoimmune diseases accompanied by significant pain, fatigue, and potential loss of mobility (9), contributing to difficulties at work and in relationships (10). Although there are no definitive prevalence figures for inflammatory arthritis (IA) as a group of conditions, 3–5% of Canadians are estimated to be living with some form of IA (11). While most types of IA affect approximately twice as many women as men (11), AS affects 2 to 3 times as many men as women (12). IA leads to reduced work capacity, promotional opportunities, educational achievement, and performance of nonpaying jobs and housework (9–10), but regarding men, little is known about IA and parenting.

Women have qualitatively described negative impacts of IA on their role as mothers, such as pain and fatigue disrupting the performance of parenting tasks and limitations leading to feelings of loss, yet also suggesting how new routines to accommodate IA can positively influence family relationships and nurturing

[1]Michal Avrech Bar, PhD: School of Health Professions, Tel Aviv University, Tel Aviv, Israel; [2]Thao T. Dao, MOT: WorkSafeBC, Vancouver, British Columbia, Canada; [3]Leah R. DeBlock Vlodarchyk, MOT: Fort Saskatchewan Community Hospital and Alberta Health Services, Fort Saskatchewan, Alberta, Canada; [4]Catherine L. Backman, PhD: University of British Columbia, Vancouver, and Arthritis Research Canada, Richmond, British Columbia, Canada.

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

Address correspondence to Catherine L. Backman, PhD, Department of Occupational Science and Occupational Therapy, The University of British Columbia, T325-2211 Wesbrook Mall, Vancouver, BC, Canada V6T 2B5. Email: catherine.backman@ubc.ca.

Submitted for publication July 31, 2019; accepted in revised form March 10, 2020.

SIGNIFICANCE & INNOVATIONS

- This preliminary grounded theory fills a knowledge gap on the impact of arthritis on fathering, describing how married men experience parenting in the presence of inflammatory arthritis (IA).
- While men in this study reported some task limitations due to arthritis, overall, they believed IA did not affect their sense of being involved, engaged, and committed fathers.
- This group of men uniformly "took ownership" of their health and "redefined" themselves as competent fathers able to concurrently manage their disease and their parenting responsibilities, with examples that differ from the range of women's experiences with arthritis and motherhood reported in prior studies.

children (13–16). Men with IA may have similar parenting issues, but few studies have examined their perspectives. Pain, fatigue, and physical functioning create parenting challenges for both men and women, but mothers have reported more difficulty caring for preschool children, while fathers had more difficulty with older children (17). A narrative review on the psychosocial impact of RA found sex differences regarding the impact on quality of life, work, distress, coping, and support, but cited the lack of studies including men's experiences (18). A qualitative study of fathers with scleroderma reported that fatigue, physical limitations, and vascular changes prevented men from engaging in physical play with their children (19). Men with cancer reported that managing multiple roles as working men, husbands, and fathers became increasingly complex when dealing with illness, and the ongoing emotional impact included feelings of frustration, isolation, and fear of mortality (20), similar to men with scleroderma (19).

A gap exists regarding the impact of IA on fathers. Qualitative investigations of men with IA will offer insight into perspectives on fatherhood and difficulties encountered and will inform health services and resources to support parenting. Therefore, our study purpose was to describe fathering roles and experiences, the effect of IA on parenting activities, and strategies used by fathers to fulfill this role in the presence of arthritis.

MATERIALS AND METHODS

Design. The Standards for Reporting Qualitative Research (21) guide this report. A qualitative interview study informed by grounded theory (22–23) is selected as an appropriate starting point when little is known about a phenomenon. Constructivist grounded theory (23–24) is congruent with the research purpose, to understand a social process (fathering) grounded in real-life experiences. Data gathering and analysis occur concurrently and support co-construction of concepts reflecting participants' subjective experiences of fathering in the presence of IA. Interview studies in our region attract few men, and this was an initial, unfunded investigation; consequently, we planned to recruit 10–12 men to justify the need for further study. While some authors recommend samples of 20–30 participants for grounded theory, congruence with study purpose and data quality are equally important considerations in determining sample adequacy (24). The systematic yet flexible grounded theory methodology, applied well, scales down to small samples (25).

Participant selection. Potential participants were fathers who: 1) had at least 1 child age <19 years living at home, at least 1 week per month (if shared custody); 2) had an IA diagnosis confirmed by a rheumatologist (RA, AS, psoriatic arthritis, or juvenile idiopathic arthritis [JIA]); and 3) were residing within 100 km of Vancouver, Canada. We sought variation in age/number of children, disease duration, and disease onset before/after parenthood as primary considerations and kept eligibility criteria broad to foster recruitment. Recruitment notices were distributed via rheumatologists' offices, a multidisciplinary arthritis outpatient clinic, and social media. The University of British Columbia Behavioral Research Ethics Board approved the study and all participants provided informed consent.

Data collection and analysis. One in-person, in-depth semistructured interview occurred with each participant at a location of his choice (home, workplace, coffee shop). Interviews discussed households, a typical day, what it was like to be a father, the impact of arthritis on daily life, and parenting strategies. Follow-up probes encouraged elaboration. A patient research partner (a parent with arthritis) helped develop the interview guide (Table 1). Questions and probes evolved as interviews occurred to incorporate insights from prior interviews and test preliminary categories, consistent with grounded theory approaches toward

Table 1. Interview guide as it existed for interviews 4 and 5*

1) Basic demographic queries: age, marital status, children, diagnosis.
2) Tell me about your family. What interests do you have/what do you like to do as a family?
3) Tell me about a typical day. Explore work, household activities, hobbies, chores, child care. What's working well/isn't working?
4) Tell me about your arthritis. Explore impact, if any: on what you do, on your activities, on your decisions as a family.
5) How would you define the role of a father? What do dads do, or are they expected to do? (Probe for personal, family, societal expectations). Follow-up: What does being a dad mean to you/how do you feel about being a father? What makes you proud?
6) Are you experiencing any challenges with your arthritis or your role as a father, physical, emotional, or otherwise? What strategies do you use to manage challenges or demands?
7) What kind of support or information do you look for? Have you found the support you need? Who provides it?
8) Any advice for other dads?
9) Anything you'd like to add?

* Interviews were conversational in nature and used this list as a topic guide, not as a script.

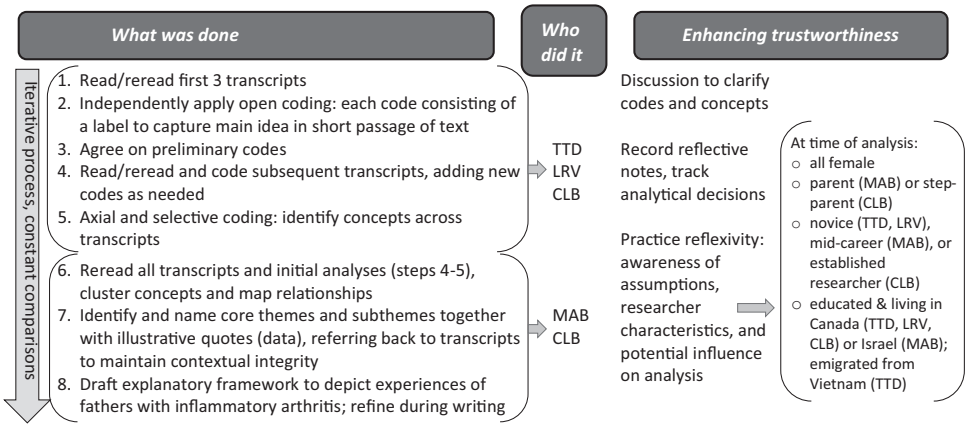

Figure 1. Summary of data analysis procedures. Initials indicate the authors.

theoretical sampling and deeper understanding. All interviews were conducted by the second author (TTD), audiotaped, and transcribed verbatim.

An iterative analytical process (detailed in Figure 1) constructed main themes and a preliminary grounded theory (22–24,26). Repetition, coherence, and support of conceptual categories across successive interviews suggested theoretical saturation for identified themes, but the wholly transparent reason that recruitment stopped was lack of resources.

RESULTS

Participants. Nine men, ages 31–62 years, were interviewed (Table 2). Eight men were married with children at home full-time; 1 divorced participant was remarried with stepchildren living with him and his wife part-time. Participants had 1 to 5 children, ages 6 months to 28 years. Three became fathers before their IA diagnosis, 5 had IA before having children, and 1 had 2 children before his diagnosis and 3 afterward. Two men immigrated to Canada (from Holland and Ethiopia), 1 was of Chinese descent, and the remainder White. Participants selected their own pseudonyms. Eight interviews lasted from 46 to 72 minutes (mean = 60), and 1 interview was 2 hours.

Core themes. Although participants reported varied views on fatherhood and IA impact, men in this study universally loved being fathers. Paul-2 stated, "I love being a dad…it means the world to me," and TMC noted, "I love it…I always wanted to be a dad. I wanted to have kids when I was 21 years old." Colin declared, "It's the best role you can have. Being a parent is the most important thing there is. It comes ahead of work, it comes ahead of social time, it comes ahead of everything else."

Table 2. Participant characteristics*

Name†	Age	Diagnosis/ duration	Children	Employment status	Other
Chuck	37	AS, 13 years	Son, age 6 months	Full-time graduate student	Father had AS
TMC	44	AS, 27 years	Sons, ages 12, 14 years	Firefighter and shopkeeper	Father had AS
Colin	50	PsA, 20 years	Daughters, ages 16, 18, 20 years	Marine terminal manager	–
Dan	62	RA with SS, 3 years	Stepsons at home, ages 18, 20 years; sons no longer at home, ages 26, 28 years	Aircraft mechanic instructor on short-term disability leave	Lives in Canada and Thailand
Joe	48	RA, 1 year	Daughters, ages 7, 13 years	Produce manager on sick leave	Chinese ethnicity
Paul-1	57	PsA, 4 years	Son, age 6 years	Bookbinder on short-term disability leave	Emigrated from Ethiopia
Paul-2	42	AS, 7 years plus childhood arthritis	Five children, ages 1.5, 4, 6, 9, 15 years	Poultry farm manager	Emigrated from Holland
Trent	31	JIA, 29 years	Daughter, age 6 months	Disability pension	Wife has JIA
Sean	30s‡	AS, 9 years	Daughter, age 6 months	Environmental scientist, on parental leave	–

* AS = ankylosing spondylitis; JIA = juvenile idiopathic arthritis; PsA = psoriatic arthritis; RA = rheumatoid arthritis; SS = Sjögren's syndrome.
† Participants chose their own pseudonym.
‡ Precise age not stated.

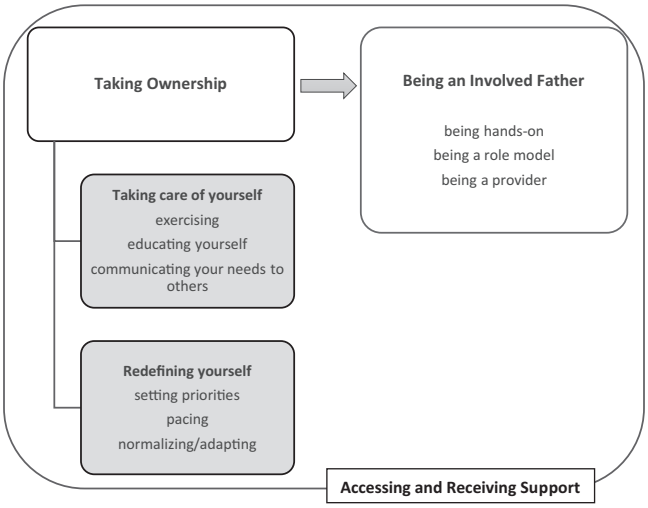

Figure 2. A preliminary grounded theory explaining fatherhood experiences in the presence of inflammatory arthritis, showing relationships among core themes.

Our analysis generated 3 core themes (depicted in Figure 2): being an involved father, taking ownership (with 2 subthemes), and accessing support.

Theme 1: being an involved father. Men discussed fatherhood as being involved in their children's lives, depicted in 3 categories: "hands-on parent," "role model," and "financial provider." Participation in fathering occupations entailed doing tasks required to care for children, which varied with life stages. "You do everything that needs to be done," said Dan. Chuck's examples of being a hands-on father included: "I change diapers, I get him dressed, I expect to do all that stuff." Others described taking children to school, preparing lunches, or accompanying them to after-school activities: "I love to take her [youngest daughter] to the ballet…and love to go to hockey games with my middle daughter," said Colin, who also coached his daughters' sports teams.

Being an involved father included sharing and negotiating caregiving with their wives. "[Dads do] everything. I have to share the main importances [sic] to raise the child together with the wife. It's not my duty only to raise him, but my wife's too. We share everything" (Paul-1). Sean, on paternity leave while his wife had returned to work, described "stay-at-home parenting" as:

> "A really unique fathering role, you know, just as the primary caregiver. I'm really taking on more of a mothering role, other than the obvious biological things. I tend to be the more physical parent as far as play with her…I mean I feed her more often than her mother does. We're working on solid foods, and that's a surprising amount of work."

Although all participants described periods where arthritis restricted activity performance and choice, they generally agreed with Sean: "There's not much I've been prevented from doing" as a parent. Participants relayed past, present, temporary, and persistent limitations to performing hands-on tasks due to pain, fatigue, and reduced strength or mobility (Table 3: Chuck, Paul-1, Trent). Colin reflected back to when his children were young:

> "There were certainly times where I was frustrated as a parent, there were certainly times where…we had to cut a walk

Table 3. Additional illustrations for selected themes or subthemes

Theme or subtheme	Sample quotes
Being an involved father (limitations being hands-on)	I've had a few moments where I've felt bad because I needed to put him down, he weighs less than 20 pounds still, but that can be hard sometimes if I'm having a bad day. (Chuck) I can't play Frisbee with my son…so he gets upset a bit because he thinks I'm holding back or something, but he understands that I'm in pain. (Paul-1) It's the physical things, like not being able to bathe her because the bathtub's too low. (Trent)
Being a role model, guiding, mentoring	You want to kind of point them in the right direction and give them all the stimulation and opportunities that they need. You want to provide that guidance for them and that assistance. (Dan) The importance of being a good father is that there is this linear progression of everything, so you're going to create a lineage of your children who have children, who have children. Just like if you think back to your father and your father's father. It's an important thing to do because you set their morals; how that individual [lives] will affect generations after you. (Colin)
Taking care of yourself	Since my diagnosis, I've actually increased my mobility, largely through…being active and doing a lot of yoga and snowboarding…just making sure that I stay active on a regular basis is sort of the number one thing I find for me that'll really help me get through the day. (Sean) I definitely need regular exercise, it makes a big difference. Now, I've been particularly lucky in that I come from an athletic background and I know how and I know what to do and I came into the disease in fairly good shape. I was lifting weights and doing aerobics regularly, every day, pretty much, when I got the illness and the exercise makes a huge difference. (Dan)
Redefining yourself	I have to change my lifestyle…I have to look after myself and so my primary concern is my physical and mental health and I need to make it a priority over and above everything else. (Dan) …trying to find…something alternative to what I used to do or living with the new reality. (Trent) You need to pace yourself because if you don't pace yourself, you don't realize you're not the same person. (Joe) It was a tough sort of 5-year period [with active inflammatory arthritis], but my children were small so I didn't miss out on a lot…I didn't have to worry about kicking a ball hard because they were small. As they got older and my fitness increased I was able to relatively confidently coach, so I coached them in soccer and softball and some tennis. (Colin)

short because of me, not my 3-year-old. There were times when my children said 'Daddy, up on shoulders' and I just couldn't do it when they were small, so that's disappointing."

Despite arthritis interfering with specific tasks, men emphasized that IA did not affect their ability to be good fathers. Trent, physically limited by JIA for many years, said, "Just because you're physically disabled doesn't mean you're not going to be a good father." His reflections on being an involved father addressed the balance between being hands-on and a positive influence:

> "Always be there for your child, emotionally, supportive. I mean be there for advice, as a shoulder to cry on, and be there to share your triumphs too. I mean father is the play figure, too, right? And that's the biggest part I have struggling with this. How do I be that play figure, that big tumbly, wrestly dad?"

Being an involved father included guiding children's development as human beings and valuing opportunities to "spend as much time as possible" with them (Joe). "Parenting is about being a role model and living by example…so arthritis has nothing to do with that" (TMC). When describing important aspects of fatherhood, all men talked about providing guidance for the future (Table 3, Dan) and being a mentor, facilitator, or teacher, by explaining to their children "what life is all about" (Paul-1). Colin took a longitudinal view, noting how fathers influence multiple generations (Table 3).

Five fathers spoke about being the family's primary financial provider. For some, like Paul-2, being the provider was crucial to his involvement: "My first role is to bring money on the table." He referenced the cultural context: "In our community circle, it [being a good father] means to take care of your family's financial aspect." Paul-1 also relayed traditional sociocultural expectations: "I grew up in a society where the father is the breadwinner, you know, and mom is a homemaker"; consequently, he felt primary responsibility for his family's financial security. Dan described pressures from financing a house and meeting his family's material needs. In contrast, TMC explained that the traditional financial provider role he saw his father fulfill was shifting; while TMC still provided the majority of his family's income, he observed more couples sharing in financial, nurturing, and household work, something he attributed in part to his 4 days on, 4 days off employment schedule.

Theme 2: taking ownership. The second core theme explains assuming responsibility to effectively manage daily life and adapt to changing circumstances. Taking ownership has 2 subthemes: "taking care of yourself" and "redefining yourself."

Taking care of yourself. "Take care of yourself first" was repeated within and across interviews to describe adjustments learned through experience and formal health education. It was a strategy for staying well:

> "You have to take care of yourself first. If you don't take care of yourself first then you're not going to be there. Make sense? Take care of yourself first. If you need to rest, if you need to do that stuff, then you'll be that much better a parent when you do feel good" (TMC).

Self-care strategies used were regular exercise, obtaining arthritis information or treatment, and communicating with family and friends. Most fathers stressed taking care of themselves physically (Table 3, Sean), and staying active was an easily adopted self-management strategy for those engaged in sports before their diagnosis (Table 3, Dan).

Participants learned about exercising from their physicians, physical therapists, counselors, or the Internet and books, which they subsequently verified with health professionals. Paul-2 suggested that it is "really important that you go hunt for information. Be very open-minded and don't just take 1 source. Take several sources." As Dan explained:

> "With a chronic illness, you have to educate yourself and you have to take care of yourself and you have to learn to be very proactive and you have to do that now. Nowadays, we're lucky there's so much good stuff on the Internet, and there's good books available but you can't just limit it to the physical aspect…you also have to educate yourself about the psychological aspects…you have to be proactive."

Another beneficial strategy was communicating with loved ones. Chuck said that talking with his wife about his illness reflects openness in their relationship, "so that she knows what I need and what I can give." Joe communicated with his daughters:

> "As a father, you know, if they want me to play rough or jump and that sort of thing, I still have my limitations in what I can and can't do…I just tell them that unfortunately I can't do that right now, because Dad's knees or ankles are still in pain…Dad will get better one day. So she understands."

Redefining yourself. A second aspect of taking ownership was reflecting on self and identity. "Trying to figure out what I can do to redefine myself, if you will, and find new interests and things I can do," was the way Trent explained his shifting sense of self. By changing priorities, pacing activities, and normalizing their disease, fathers redefined their interests and lifestyle, thus adapting to both arthritis and the evolving role of father (Table 3, Dan, Trent, Joe). Men occasionally changed the way they did tasks and found "creative ways of doing it" (Paul-1) as part of redefining self-expectations. While actions aimed at redefining themselves overlapped with the prior subtheme of taking care of yourself, the second subtheme was differentiated by reflection on self-identity rather than health. It was enacted by planning ahead, listening to their bodies, and taking pause, as Dan explained:

> "You really have to learn how to pace yourself, as to your overall energy levels, you find that you can't try and do too much. If you feel good in the morning and you overdo it, you're really wiped. Luckily, I'm used to listening to my body… initially when I first got the illness I was taking 4 naps a day. Now I'm typically around 1 nap a day. I knew that one of the things I had to do was accept the fact that I have a chronic illness that is, you know, reasonably major…I need to [set priorities], otherwise I'm not able to perform."

Participants described adjusting to changing life situations. Most fathers normalized their disease as part of life and 2 dismissed it as having minimal impact on their lives. Sean said AS was "a small part of my life, doesn't affect much, and parenting-wise it certainly doesn't," while other fathers were redefining self by making sense of their situation and adapting routines associated with being a father with IA (Table 3, Colin).

Theme 3: accessing and receiving support. Participants accessed support from family, friends, and health and social systems. Without exception, they relied on their wives as their primary source of support. Couples negotiated parenting tasks and men shared how their wives were essential to their capacity to take ownership. Men described reciprocity with their wives:

> "We all have things that we're good at and things that we're bad at…so if I'm limited by fatigue, then that's where I'd ask her for support. And I'll just stick a little joke in here: when it comes to map reading she'll ask me, and I'll teach the little guy how to read a map because she's not good at that." (Chuck)

Support ranged from informal and practical assistance with childcare or household tasks to formal services within social systems (e.g., insurance, paid leave, health education) (Table 4). No participants were hesitant about seeking or accepting support, but a few relayed unmet needs when seeking information specific to their situation. Sean said, "One thing…I found very lacking early on was support for someone with AS who's in my demographic… you know, early 30s, late 20s, athletic-type people. [Everything] was designed for people who are old…it didn't feel appropriate and wasn't helpful." The core themes of involved father and taking ownership are situated against this backdrop of accessing and receiving support, illustrated in the preliminary grounded theory (Figure 2).

DISCUSSION

Because little prior research has focused on men's experiences of arthritis and parenting (17–19), we sought to describe, in some depth, fatherhood perspectives among men with IA. Like Poole et al (19), we did not achieve our sampling goal, but variation in disease duration, functional ability, employment, and

Table 4. Examples of supports used by fathers with inflammatory arthritis

Source of support	Sample quotes
Spouse	My wife, yeah, she's definitely the most key thing for helping through on day-to-day stuff. It depends on the issue though. If my issue has something to do with arthritis symptoms, I'll probably not go to her....but day-to-day absolutely she's the most important support. (Sean) She is wonderfully nurturing, very supportive...absolutely, she's wonderful. Whatever I need, I wouldn't have to ask, it would just be offered type of thing. (Chuck)
Family members	We were also very lucky that we have both my wife's mother and mine that help babysit and so they're hanging around the house all the time, so we had a good support structure around us. (Colin)
Friends and coworkers	[My biggest supporters are] family and friends. But mostly friends. The guys that I work with, yeah, my friends in the fire department. They're like brothers. (TMC) Last weekend we had a friend offer to watch him so that we could go mountain biking. But to drive to the trail and get a good ride in is more than the 2-hour window between breast feeds, so [my friend] actually came in the car with us and walked around with [the baby] while we did a 2-hour bike ride. (Chuck)
Other people with inflammatory arthritis	You need to network with other people with arthritis, because there's a lot to be learned from that, you know. I've taken some of the courses the Canadian Arthritis Society provides which have been very helpful...very supportive. (Dan)
Financial	I'm very fortunate that between my insurance coverage, my company, and my wife's, we don't pay for any medical expenses. My [medication] is, I think I was told, about $17,000 a year now. I don't pay for a penny of that. (Colin)
Environmental adaptations and access to assistive devices	Definitely an eye-opener, like at first I couldn't sit down on the toilet. The seat was so low and I just didn't have the strength, so the occupational therapist was great, they rented me a toilet seat from the Red Cross and at first I was, to be honest, I was embarrassed about it...it got to the point my wife didn't want to use the bathroom in our bedroom. It did assist me in a lot of ways to get stronger with my legs, in the 3-month period that I did have it, it helped. (Joe) I've had an occupational therapist come into the house to help with adaptive equipment...supports for things to help me in my daily life like raised toilet seats, the bathing chair, little hand grip thing to pick stuff up off the floor, can openers. Stuff that I'm having quite a bit of difficulty with right now. (Trent)
Internet, books	Anything you need to know, you learn a lot of things from it [the Internet], then you see a way, you know, whether or not it's going to help you, then you ask professionals whether you should do that. Everything on the Internet, I don't believe it's true, it's somebody's opinion. Then you ask the professionals and see whether it works for you. (Paul-1) I have a whole box full of magazines and books and articles that I've read in the past that I think are interesting and I save them. (Paul-2)

children's life stage within our sample provided a range of fathering experiences. Figure 2 reflects the socially constructed reality of fatherhood for this small group of men with IA. Participants loved their role as fathers and were well-supported by partners, family, and friends, and thus, the grounded theory generated is specific to these men and may not reflect other men's perspectives.

As involved fathers, the participants regularly engaged in tasks such as feeding, dressing, bathing, and playing with their children. They saw themselves as role models shaping their children's morals and future. Although participants identified periodic frustrations with some hands-on activities, they did not convey the emotional distress reported by fathers with scleroderma, where fear of death, social isolation, and frustration with inability to play with children were prominent features (19). The nature of "involved father" among our participants with IA is similar to contemporary descriptions of fathers without disabilities (2,27), where being committed to parenting and household tasks and authentic coparenting are prominent themes. Vocabulary used by our participants was consistent with phrases used by nondisabled fathers, like "spending time with" and "being there for" their children (28), suggesting similar perceptions of what it means to be an involved father. While financial provider was a stated role component for some fathers, it was not the most important (28), illustrating evolving views of masculinity, with increased confidence and identity as caregivers, an observation corroborated in our study.

Some men with RA seek to preserve hegemonic ideals of masculinity (physical strength, breadwinner); other men renegotiate masculinity (replacing former masculine activities with redefined masculine activities) or are sufficiently comfortable to reject acting in hegemonic masculine ways (29). Parts of our analysis support all 3 typologies. Like 2 case studies shared by Flurey et al (29), the use of third-person language in some quotes from our participants suggests they may be protecting hegemonic masculine identities in emotionally challenging situations by separating themselves from events. The preference for exercise as a key self-management strategy is also consistent with a traditional masculine characteristic of physical strength. The focus on physical activity for taking care of self is consistent with a review reporting that men valued physical activity (18). However, our core themes of being an involved father and taking ownership are most consistent with renegotiating masculinity, as men in our study described actions taken to redefine themselves as fathers. Future research on fathering experiences would benefit from directly querying masculinity ideals during data generation.

In contrast to fathers without disabilities who prioritized their family's needs over their own (27), our participants identified the importance of taking ownership, to take care of themselves first, to be well enough to take care of their children. Men in both studies share a similar belief, that family comes above all, but men with IA emphasized managing their illness as the first step to upholding this belief. The integration of IA into redefining self as part of adapting to life with arthritis supports a tentative finding from a prior literature review (18).

Our participants readily accepted practical and social support and spoke confidently of communicating with others to access support, whereas a review reported that men were less likely than women to seek support (18). This discrepancy may reflect differences in the study samples, and further research could identify trends in support sought by men and women.

Compared to prior studies of mothers with IA (13,16), fathers in the current study perceived participation restrictions differently. For example, both mothers (13,16) and fathers reported difficulties with or inability doing physical activities like lifting/carrying children and playing sports, but the impact on their perception of parenting differed. Some mothers felt a profound sense of loss in doing things differently and worried about the impact of their arthritis on their children (13,16), whereas fathers explicitly stated that these limitations had little to do with being a father. Some mothers suggested that arthritis had positive attributes in nurturing children to be caring, thoughtful adults (13,15–16), which was less apparent in fathers' accounts.

Differences could be explained by more within-sample variation in the mothering studies or gendered perspectives on parenting responsibilities. Possibly, some women worry more than men about the gap between their desired participation and actual participation as parents; they report more ups and downs in participation attributed to IA (13). Or men in the current study may not be fully aware of tasks naturally assumed by their wives and do not experience participation in the same way as mothers with IA. A similarity observed between mothers and fathers is the strong sense of "being" parents, regardless of ability.

As one of the few descriptions of fathering and IA, our findings may inform practice. Health professionals should inquire about the impact of IA on fathering and the presence of spousal support, since these men relied on their wives. Taking ownership, as described in this study, may illustrate ways to promote self-management to preserve participation in valued life roles like parenting.

Our participants may have volunteered because they valued parenting and were more highly involved than other men. They were willing to talk about their experiences as fathers with IA and were active self-managers; we may be missing voices of "strong and silent" men (29–30). The taking ownership theme suggests they were committed to managing their health, and all were well-supported by their wives. Therefore, the initial grounded theory presented here may have less relevance to single fathers, men with same-sex spouses, men who are struggling with their disease, or men lacking support. Nevertheless, findings suggest that when social, practical, and educational supports are in place, men with IA enjoy parenthood, find it rewarding, and believe they are largely unrestricted by arthritis in being good fathers.

ACKNOWLEDGMENTS

The authors gratefully acknowledge staff at the Mary Pack Arthritis Program for assistance with participant recruitment, and we thank patient partner Nadia Prestley for comments on study design, recruitment notices, and interview guide.

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. Backman 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. Dao, DeBlock Vlodarchyk, Backman.
Acquisition of data. Dao, DeBlock Vlodarchyk, Backman.
Analysis and interpretation of data. Avrech Bar, Dao, DeBlock Vlodarchyk, Backman.

REFERENCES

1. Lamb ME. The role of the father in child development. 4th ed. Hoboken (NJ): Wiley; 2014.
2. Cabrera N, Tamis-LeMonda CS, Bradley RH, Hofferth S, Lamb ME. Fatherhood in the twenty-first century. Child Dev 2000;71:127–36.
3. McGill BS. Navigating new norms of involved fatherhood: employment, fathering attitudes and father involvement. J Family Issues 2014;35:1089–106.
4. Banchefsky S, Park B. The "new father:" Dynamic stereotypes of fathers. Psychol Men Masc 2016;17:103–7.
5. Chin R, Daiches A, Hall P. A qualitative exploration of first-time fathers' experiences of becoming a father. Community Pract 2011;84:19–24.
6. Fägerskiöld A. A change in life as experienced by first-time fathers. Scand J Caring Sci 2008;22:64–71.
7. Hamilton A, de Jonge D. The impact of becoming a father on other roles: an ethnographic study. J Occup Sci 2010;17:40–6.
8. Johansson T. Fatherhood in transition: paternity leave and changing masculinities. J Fam Commun 2011;11:165–80.
9. Kvien TK. Epidemiology and burden of illness of rheumatoid arthritis. Pharmacoeconomics 2004;22 Suppl 1:1–12.
10. Felthusen C, Björk M, Forsblad-d'Elia H, Mannerkorpi K. Perception, consequences, communication and strategies for handling fatigue in persons with rheumatoid arthritis of working age: a focus group study. Clin Rheumatol 2013;32:557–66.
11. Gladman DD. Trauma and inflammatory arthritis. 2008. Workplace Safety and Insurance Appeals Tribunal. URL: https://www.wsiat.on.ca/en/MedicalDiscussionPapers/arthritis.pdf
12. BMJ Publishing Group. BMJ best practice ankylosing spondylitis. 2018. URL: https://bestpractice.bmj.com/topics/en-gb/366/pdf/366/.pdf.
13. Backman CL, Del Fabro Smith L, Smith S, Montie PL, Suto M. Experiences of mothers living with inflammatory arthritis. Arthritis Rheum 2007;57:381–88.
14. Feddersen H, Kristiansen TM, Andersen PT, Hørslev-Petersen K, Primdahl J. Construction of meaningful identities in the context of rheumatoid arthritis, motherhood and paid work: a meta-ethnography. J Clin Nurs 2017;26:4117–28.
15. Poole J, Willer K, Mendelson C. Occupation of motherhood: challenges for mothers with scleroderma. Am J Occup Ther 2009;63:214–9.
16. Del Fabro Smith L, Suto M, Chalmers A, Backman CL. Belief in doing and knowledge in being mothers with arthritis. OTJR (Thorofare, N J) 2011;31:40–8.
17. Barlow JH, Cullen LA, Foster NE, Harrison K, Wade M. Does arthritis influence perceived ability to fulfill a parenting role? Perceptions of mothers, fathers and grandparents. Patient Educ Couns 1999;37:141–51.
18. Flurey CA, Hewlett S, Rodham K, White A, Noddings R, Kirwan J. Men, rheumatoid arthritis, psychosocial impact and self-management: a narrative review. J Health Psychol 2016;21:2168–82.
19. Poole JL, Haygood D, Mendelson C. "I'm still dad:" the impact of scleroderma on being a father. Occup Ther Health Care 2018;32:1–13.
20. Elmberger E, Bolund C, Lützén K. Men with cancer: changes in attempts to master the self-image as a man and as a parent. Cancer Nurs 2002;25:477–85.
21. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014;89:1245–51.
22. Strauss A, Corbin J. Basics of qualitative research: techniques and procedures for developing grounded theory. 2nd ed. Thousand Oaks (CA): Sage; 1998.
23. Charmaz K. Constructing grounded theory. London: Sage; 2006.
24. Charmaz K. The power and potential of grounded theory. Medical Sociology Online. 2012;6. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1062.8596&rep=rep1&type=pdf.
25. Treem JW, Browning L. Grounded theory. In: Scott C, Lewis L, editors. The international encyclopedia of organizational communication, 4 volume set. West Sussex (UK): John Wiley; 2017. p. 1045–59.
26. Shkedi A. The meaning behind the words. Methodologies of qualitative research: theory and practice. Tel Aviv: Ramot; 2011.
27. Wada M, Backman CL, Forwell SJ. Men's discursive constructions of balance in everyday life. Community Work Fam 2015;18:117–33.
28. Williams RA. Masculinities and fathering. Community Work Fam 2009;12:57–73.
29. Flurey C, White A, Rodham K, Kirwan J, Noddings R, Hewlett S. "Everyone assumes a man to be quite strong": men, masculinity and rheumatoid arthritis, case-study approach. Sociol Health Illn 2018;40:115–29.
30. O'Brien R, Hunt K, Hart G. "It's caveman stuff, but that is to a certain extent how guys still operate": men's accounts of masculinity and help seeking. Soc Sci Med 2005;601:503–16.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 893–897
DOI 10.1002/acr.24200

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Ustekinumab for the Treatment of Giant Cell Arteritis

Mark A. Matza, Ana D. Fernandes, John H. Stone, and Sebastian H. Unizony

Objective. To evaluate the efficacy and safety of ustekinumab (UST) in giant cell arteritis (GCA).

Methods. We conducted a prospective, open-label trial of UST in patients with active new-onset or relapsing GCA. Active disease was defined as the presence of GCA symptoms and elevation of erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP) level within 6 weeks of baseline. All patients received a 24-week prednisone taper and subcutaneous UST 90 mg at baseline and at weeks 4, 12, 20, 28, 36, and 44. The primary endpoint, prednisone-free remission, was defined as the absence of relapse through week 52 and normalization of the ESR and CRP level. Relapse was defined as the recurrence of GCA symptoms requiring treatment intensification. A sensitivity analysis excluding ESR/CRP level normalization from the prednisone-free remission definition was performed.

Results. The study enrolled 13 patients (target sample size 20). Enrollment was closed prematurely after 7 of the initial 10 patients relapsed. Five patients (39%) had new-onset disease. The initial prednisone doses were 20 mg (1 patient), 40 mg (9 patients), and 60 mg (3 patients). All patients entered disease remission within 4 weeks of baseline. Only 3 (23%) achieved the primary endpoint. Of the 10 patients (77%) who failed to achieve the primary endpoint, 7 relapsed after a mean period of 23 weeks. The remaining 3 patients met the alternative definition of prednisone-free remission that did not require ESR/CRP level normalization. One serious adverse event occurred.

Conclusion. UST combined with 24 weeks of prednisone was associated with a high rate of treatment failure in this prospective GCA trial.

INTRODUCTION

Giant cell arteritis (GCA) is the most common primary form of vasculitis in adults in Western countries, where the lifetime risks in women and men age ≥50 years are 1% and 0.5%, respectively (1). Signs and symptoms associated with this chronic inflammatory disorder include headaches, jaw claudication, scalp tenderness, visual impairment, polymyalgia rheumatica (PMR), overwhelming fatigue, fever, and weight loss. The most frequent disease complications are blindness, which occurs in 15–20% of the patients (2), and aortic aneurysm, which develops in 20–30% of the patients (3).

Until recently, the treatment for GCA was limited to lengthy courses of glucocorticoids (1). However, 40–80% of patients treated with these regimens develop disease relapse, depending on the duration of follow-up, the definition of relapse used, and most importantly, the rate of glucocorticoid dose reduction and whether this medication is tapered to discontinuation (4,5). Prolonged use of glucocorticoids in GCA is associated with several potential side-effects, including infection, osteoporosis, cataracts, insomnia, psychosis, weight gain, diabetes mellitus, and hypertension (6).

Following 2 randomized controlled trials demonstrating that interleukin (IL)-6 signaling blockade with tocilizumab is an effective therapy for GCA (7,8), treatment approaches are changing. Unfortunately, IL-6 inhibition is not curative, and ~30% of patients treated with tocilizumab relapse within 12 months (7,9). In addition, another 5% of patients must discontinue tocilizumab due to treatment-related adverse events (10). Therefore, a great need exists in GCA for additional treatment options to prevent disease relapse in a manner that is safe and permits a reduction in the use of glucocorticoids.

IL-12 and IL-23 have been implicated in the pathogenesis of GCA by fostering the T helper (Th)1 and Th17 CD4-positive

ClinicalTrials.gov identifier: NCT02955147.
Supported by Janssen Pharmaceuticals.
Mark A. Matza, MD, MBA, Ana D. Fernandes, BA, John H. Stone, MD, MPH, Sebastian H. Unizony, MD: Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts.
Dr. Unizony has received consulting fees from Sanofi-Aventis and Kiniksa Pharmaceuticals (less than $10,000 each) and grants from Genentech. No other disclosures relevant to this article were reported.

Address correspondence to Sebastian H. Unizony, MD, Co-Director, Vasculitis and Glomerulonephritis Center, Division of Rheumatology, Allergy, and Immunology, Massachusetts General Hospital, Assistant Professor of Medicine, Harvard Medical School, Yawkey Center, Suite 2C, 55 Fruit Street Boston, MA 02114. Email: sunizony@mgh.harvard.edu.
Submitted for publication December 23, 2019; accepted in revised form March 24, 2020.

SIGNIFICANCE & INNOVATIONS

- Interleukin-12/23p40 blockade with ustekinumab combined with 24 weeks of prednisone was associated with a high rate of treatment failure in this prospective giant cell arteritis trial.
- Ustekinumab was well tolerated.

lymphocyte pathways. These cytokines are heterodimers that share a common p40 subunit (IL-12/23p40). In addition, a functional IL-12 molecule requires a p35 subunit, and a functional IL-23 molecule includes a p19 subunit. Preliminary evidence suggests that IL-12/23p40 blockade therapy with ustekinumab (UST) may be effective in controlling GCA activity (11). In this pilot trial, we aimed to assess the efficacy and safety of UST in combination with 6 months of prednisone in a group of patients with GCA.

PATIENTS AND METHODS

Study design and patients. We conducted a prospective, single-center, single-arm, open-label trial to evaluate the efficacy and safety of UST in combination with prednisone in patients with active GCA. Patients were required to be age ≥50 years and could have either new-onset (diagnosis within 6 weeks of baseline) or relapsing disease (diagnosis >6 weeks from baseline). GCA diagnosis required 1) the current or historical presence of characteristic cranial symptoms (e.g., new-onset headache, jaw claudication, scalp tenderness, visual disturbances) or PMR symptoms; 2) elevated inflammatory markers (erythrocyte sedimentation rate [ESR] ≥50 mm/hour or C-reactive protein [CRP] level ≥10 mg/liter); and 3) temporal artery biopsy or vascular imaging confirmation of the diagnosis. The types of vascular imaging studies considered for the purpose of diagnosis were computed tomography angiography, magnetic resonance angiography, and positron emission tomography.

Disease activity for study enrollment was defined as the presence of cranial or PMR signs or symptoms along with elevation of the ESR (≥40 mm/hour) or CRP level (≥10 mg/liter) within 6 weeks of baseline. Patients with uncontrolled comorbidities, active infection, severe infection or infection requiring antibiotics within 4 weeks of baseline, a history of recurrent or opportunistic infection, or malignancy within the previous 5 years were excluded. Patients receiving methotrexate were eligible after a 2-week washout period.

Treatment regimen. All patients received UST 90 mg, which was administered subcutaneously at baseline and then at weeks 4, 12, 20, 28, 36, and 44. In addition, all patients received a prespecified 24-week prednisone taper starting at 60 mg, 40 mg, or 20 mg daily (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract). The initial prednisone dose was selected by the investigators according to best clinical judgment.

Disease assessments and study endpoints. After baseline, patients were evaluated at weeks 4, 12, 20, 28, 36, 44, and 52 (treatment phase). A safety follow-up visit occurred at week 60. The primary study endpoint, prednisone-free remission, was defined as: 1) the absence of relapse from the time that remission was achieved through week 52; 2) normalization of ESR (<40 mm/hour) and CRP level (<10 mg/liter); and 3) adherence to the protocol prednisone taper. Disease relapse was defined as the recurrence of signs or symptoms of GCA (e.g., cranial or PMR) that required treatment intensification, regardless of the ESR and CRP levels. A sensitivity analysis that excluded ESR and CRP level from the definition of prednisone-free remission was also performed. Other study endpoints included time to disease relapse, cumulative prednisone dose, and safety.

Ethical considerations. The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. The study protocol was approved by the Institutional Review Board of Massachusetts General Hospital, and all patients provided written informed consent. The investigators designed the study and gathered and analyzed the data.

Statistical analysis. A convenience sample of 20 consecutive patients was planned. However, enrollment was prematurely closed after 7 of the first 10 patients entering the study relapsed. Here we report the outcomes of 13 patients enrolled in the trial. All data were analyzed using descriptive methods. Means ± SDs were reported for continuous variables. Counts and percentages were reported for categorical variables.

RESULTS

Patients. Between December 2016 and August 2018, we screened 16 GCA patients for this trial. Three patients failed the screening process (dementia [n = 1], severe prednisone-induced depression [n = 1], and positive hepatitis B core antibody [n = 1]), and 13 patients were enrolled. The mean age of the cohort was

Table 1. Patient baseline characteristics*

Characteristic	GCA patients (n = 13)
Age, mean ± SD years	71 ± 7
Female	11 (85)
White	13 (100)
New-onset disease	5 (39)
Biopsy-proven disease	11 (85)
Imaging-proven disease	4 (31)
Cranial signs or symptoms	13 (100)
PMR symptoms	8 (62)
ESR, mean ± SD mm/hour	41 ± 16
CRP, mean ± SD mg/liter	50 ± 39

* Values are the number (%) unless indicated otherwise. CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; GCA = giant cell arteritis; PMR = polymyalgia rheumatica.

Table 2. Efficacy outcomes*

Outcome	GCA patients (n = 13)
Prednisone-free remission by week 52	3 (23)
Alternative definition of prednisone-free remission by week 52†	6 (46)
Disease flare	7 (54)
Clinical features at disease relapse‡	
Cranial signs or symptoms	3 (43)
PMR symptoms	7 (100)
ESR, mean ± SD mm/hour	49 ± 26
CRP, mean ± SD mg/liter	40 ± 34
Time to flare, mean ± SD weeks	23 ± 7
Number of UST doses received, mean ± SD	4 ± 1
Prednisone dose, mean ± SD mg/day	3 ± 3

* Values are the number (%) unless indicated otherwise. CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; GCA = giant cell arteritis; PMR = polymyalgia rheumatica; UST = ustekinumab.
† The alternative definition of prednisone-free remission required the absence of disease relapse up to week 52 while adhering to the protocol prednisone taper, regardless of the level of ESR and CRP.
‡ Analyses limited to the 7 patients who relapsed.

71 years, and 11 (85%) were women. GCA was confirmed by temporal artery biopsy in 11 patients (85%) and by vascular imaging in 4 patients (31%). Five patients (39%) had new-onset disease. Other demographic characteristics and clinical features of the cohort are shown in Table 1. The initial prednisone dose was 60 mg in 3 patients, 40 mg in 9 patients, and 20 mg in 1 patient. At baseline, none of the patients were receiving other immunomodulatory treatments for GCA besides glucocorticoids. One subject had previously failed a trial of tocilizumab, and another subject had previously failed trials of methotrexate, tocilizumab, and abatacept (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract).

Efficacy. Efficacy outcomes are shown in Table 2 and Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract. All 13 patients entered disease remission within 4 weeks of baseline. The primary endpoint was achieved by 3 patients (23%). Of the 10 patients (77%) who failed to achieve the primary endpoint, 7 experienced relapses after a mean ± SD period of 23 ± 7 weeks and 4 ± 1 UST injections. Six of the 7 patients who relapsed did so after weaning off prednisone or when their prednisone dose was below 5 mg (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract). One patient relapsed on a dose of 9 mg/day. The mean ± SD prednisone dose at the time of relapse was 3 ± 3 mg/day. GCA relapses included PMR symptoms in all 7 patients and cranial signs or symptoms (e.g., headache) in 3 (43%) (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract). The remaining 3 primary endpoint failures did not have clinical signs or symptoms of a relapse, but their inflammatory markers were elevated at week 52. The ESRs among these patients were 41, 42, and 18, and the CRP levels were 36, 3, and 29. The alternative definition of prednisone-free remission at week 52, a sensitivity analysis in which the acute phase reactants were excluded, was therefore met by only 6 patients (46%).

Safety. A total of 51 adverse events occurred in 13 patients, and at least 1 adverse event was observed in 12 (92%) (Table 3). Of all the adverse events, 14 (28%) were considered to be related or possibly related to prednisone (e.g., insomnia, mood swings, ecchymosis, hyperglycemia, osteoporosis, weight gain, and infection). In contrast, only 2 adverse events (4%; both nonserious infections) were considered to be related or possibly related to UST. These included a case of pneumonia that occurred in a patient at week 4, on 20 mg of prednisone and after 1 UST injection, and a case of cystitis that occurred in a subject at week 52, after 7 UST doses and receiving no prednisone. One patient developed mild diverticulitis, which was classified as a serious adverse event because it required hospitalization. This complication, considered not to be related to UST or prednisone, occurred at week 20, after 4 UST injections and at a time when the patient was also taking 4 mg of prednisone.

DISCUSSION

Treatment-related side effects (6) and disease relapse (4,5) are the most common problems faced by GCA patients. In this prospective study, UST added to a 6-month prednisone taper was not associated with a significant rate of sustained, prednisone-free remission in GCA patients with either new-onset or relapsing disease. Our goal was to include 20 patients in a proof of concept trial, but we closed enrollment early after 7 of the first 10 patients recruited relapsed and failed the primary efficacy endpoint. Of the 13 patients who completed the trial, only 3 (23%) achieved the primary endpoint. The majority of patients, 7 of 13, relapsed with

Table 3. Safety*

	GCA patients (n = 13)
Patients with at least 1 AE	12 (92)
Patients with at least 1 SAE	1 (8)
Diverticulitis	
Infections	
Diverticulitis	1 (8)
Pneumonia	1 (8)
Urinary tract infection	1 (8)
Bronchitis	1 (8)
Gastroenteritis	1 (8)
AE causality†	
AEs related or possibly related to prednisone	14 (28)
AEs related or possibly related to UST	2 (4)

* Values are the number (%). AE = adverse event; GCA = giant cell arteritis; SAE = serious AE; UST = ustekinumab.
† The denominator for the AE causality proportions is 51 total adverse events.

characteristic signs or symptoms of GCA and increased inflammatory markers. The other 3 patients classified as treatment failures did not develop recurrent clinical manifestations of GCA but finished the treatment phase of the trial with elevated levels of ESR and/or CRP, raising concern on the part of the investigators about subclinical disease activity and impending flare.

The findings of our study stand in contrast to a prior report from Ireland, in which no disease relapses were observed within 1 year among 25 GCA patients who received UST 90 mg every 12 weeks (11). In that study, the authors reported that the median prednisolone dose of the participants decreased from 15 mg/day at baseline to 5 mg/day after 52 weeks of UST treatment, but only 24% of the patients were able to discontinue prednisolone completely (11). Differences in the designs of these 2 studies offer probable explanations for their differing results. Whereas the Irish study was limited to GCA patients with a history of disease relapse (11), we enrolled both new-onset (~40%) and relapsing (~60%) patients.

Prior research has shown that IL-12/23p40 is not significantly upregulated in temporal arteries in patients with new-onset GCA, but becomes highly expressed once the disease is well established, particularly in relapsing patients (12). This fact could explain the better outcomes observed in the prior study. Nevertheless, an analysis limited to patients with relapsing disease in our cohort did not show a trend for better outcomes (data not shown). The key difference between the 2 trials, however, is that 76% of the patients in the earlier study continued to receive glucocorticoids even at week 52 (11), but all of those in our study were required to attempt to taper off prednisone completely within 6 months of starting UST. Notably, the frequency of UST administration in our trial, every 8 weeks as opposed to every 12 in the other study, led to higher cumulative UST exposures in our patients (11). In sum, the findings of this trial are in agreement with prior preclinical research suggesting that IL-12/23p40 might not be a critical mediator in the inflammation associated with GCA (13). When prednisone was tapered, patients receiving UST had a high tendency to relapse, and they often did so even while still taking prednisone.

The demonstration that IL-6 signaling blockade with tocilizumab in combination with prednisone is an effective treatment strategy for GCA represented a substantial step forward in the management of this disease (7,8). Nevertheless, up to 35% of patients receiving tocilizumab fail treatment, either due to refractory disease, disease relapse, or treatment-related side-effects (7,9,10), highlighting the need for alternative treatment options. Unfortunately, several other immunomodulatory agents have demonstrated effects that are modest at best (e.g., methotrexate and abatacept) (14,15) or have shown little efficacy whatsoever compared to prednisone alone (e.g., tumor necrosis factor inhibitors) (16,17). Trials using the Janus kinase inhibitor upadacitinib (ClinicalTrials.gov NCT03725202) and the granulocyte/macrophage-colony stimulating factor receptor antagonist mavrilumumab (ClinicalTrials.gov NCT03827018) are currently ongoing. Additionally, IL-23p19, which is much more abundant in GCA temporal arteries than IL-12/23p40 (13), has been demonstrated to have IL-23 receptor–independent proinflammatory functions signaling through gp130 and signal transducer and activator of transcription 3 (18), representing a potential treatment target.

Our study has certain limitations, stemming primarily from the relatively small number of patients enrolled and its open-label design. These could have led to type 2 errors (failing to find an effect when in fact UST does have some efficacy in GCA) as well as to bias in the assessment of the outcomes. The goal of the study, however, was to generate preliminary data on the use of UST in GCA patients following a rigorous protocol that aligns with both a major unmet need in GCA (i.e., successful discontinuation of glucocorticoids) and the current standard for clinical trials in GCA (e.g., the GiACTA trial [9]), which emphasize glucocorticoid tapering to discontinuation. Given that tocilizumab treatment is associated with a relapse rate of ~30% within 1 year (9,10), the investigators felt that it was unethical to continue recruiting patients into this UST trial in the setting of such a high treatment failure rate.

In conclusion, our results demonstrated that UST was well tolerated but did not prevent disease relapse in a significant proportion of GCA patients once prednisone was discontinued or tapered to a low daily dose. Given the small sample size of our study, firm conclusions about the efficacy of the IL-12/23p40 blockade in GCA cannot be drawn until more robust or personalized data are available.

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. Unizony 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. Stone, Unizony.
Acquisition of data. Fernandes, Unizony.
Analysis and interpretation of data. Matza, Unizony.

ROLE OF THE STUDY SPONSOR

Janssen Pharmaceuticals 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 Janssen Pharmaceuticals.

REFERENCES

1. Buttgereit F, Dejaco C, Matteson EL, Dasgupta B. Polymyalgia rheumatica and giant cell arteritis: a systematic review. JAMA 2016;315:2442–58.
2. Soriano A, Muratore F, Pipitone N, Boiardi L, Cimino L, Salvarani C. Visual loss and other cranial ischaemic complications in giant cell arteritis. Nat Rev Rheumatol 2017;13:476–84.

3. Garcia-Martinez A, Arguis P, Prieto-Gonzalez S, Espigol-Frigole G, Alba MA, Butjosa M, et al. Prospective long term follow-up of a cohort of patients with giant cell arteritis screened for aortic structural damage (aneurysm or dilatation). Ann Rheum Dis 2014;73:1826–32.
4. Stone JH, Tuckwell K, Dimonaco S, Klearman M, Aringer M, Blockmans D, et al. Glucocorticoid dosages and acute-phase reactant levels at giant cell arteritis flare in a randomized trial of tocilizumab. Arthritis Rheumatol 2019;71:1329–38.
5. Labarca C, Koster MJ, Crowson CS, Makol A, Ytterberg SR, Matteson EL, et al. Predictors of relapse and treatment outcomes in biopsy-proven giant cell arteritis: a retrospective cohort study. Rheumatology (Oxford) 2016;55:347–56.
6. Proven A, Gabriel SE, Orces C, O'Fallon WM, Hunder GG. Glucocorticoid therapy in giant cell arteritis: duration and adverse outcomes. Arthritis Rheum 2003;49:703–8.
7. Stone JH, Tuckwell K, Dimonaco S, Klearman M, Aringer M, Blockmans D, et al. Trial of tocilizumab in giant-cell arteritis. N Engl J Med 2017;377:317–28.
8. Villiger PM, Adler S, Kuchen S, Wermelinger F, Dan D, Fiege V, et al. Tocilizumab for induction and maintenance of remission in giant cell arteritis: a phase 2, randomised, double-blind, placebo-controlled trial. Lancet 2016;387:1921–7.
9. Unizony S, McCulley T, Spiera R, Pei J, Sidiropoulos PN, Best JH, et al. Clinical outcomes of patients with giant cell arteritis treated with tocilizumab in real-world clinical practice: decreased incidence of new visual manifestations. Arthritis Res Ther 2021;23:8.
10. Unizony S, Bao M, Han J, Luder Y, Sidiropoulos PN, Pei J, et al. Risk factors for treatment failure in patients with giant cell arteritis treated with tocilizumab plus prednisone versus prednisone alone. Annual Scientific Meeting of the American College of Rheumatology, Atlanta, GA, November 2019.
11. Conway R, O'Neill L, Gallagher P, McCarthy GM, Murphy CC, Veale DJ, et al. Ustekinumab for refractory giant cell arteritis: a prospective 52-week trial. Semin Arthritis Rheum 2018;48:523–8.
12. Visvanathan S, Rahman MU, Hoffman GS, Xu S, Garcia-Martinez A, Segarra M, et al. Tissue and serum markers of inflammation during the follow-up of patients with giant-cell arteritis: a prospective longitudinal study. Rheumatology (Oxford) 2011;50:2061–70.
13. Espigol-Frigole G, Planas-Rigol E, Lozano E, Corbera-Bellalta M, Terrades-Garcia N, Prieto-Gonzalez S, et al. Expression and function of IL12/23 related cytokine subunits (p35, p40, and p19) in giant-cell arteritis lesions: contribution of p40 to Th1- and Th17-mediated inflammatory pathways. Front Immunol 2018;9:809.
14. Mahr AD, Jover JA, Spiera RF, Hernández-García C, Fernández-Gutiérrez B, LaValley MP, et al. Adjunctive methotrexate for treatment of giant cell arteritis: an individual patient data meta-analysis. Arthritis Rheum 2007;56:2789–97.
15. Langford CA, Cuthbertson D, Ytterberg SR, Khalidi N, Monach PA, Carette S, et al. A randomized, double-blind trial of abatacept (CTLA-4Ig) for the treatment of giant cell arteritis. Arthritis Rheumatol 2017;69:837–45.
16. Hoffman GS, Cid MC, Rendt-Zagar KE, Merkel PA, Weyand CM, Stone JH, et al. Infliximab for maintenance of glucocorticosteroid-induced remission of giant cell arteritis: a randomized trial. Ann Intern Med 2007;146:621–30.
17. Seror R, Baron G, Hachulla E, Debandt M, Larroche C, Puechal X, et al. Adalimumab for steroid sparing in patients with giant-cell arteritis: results of a multicentre randomised controlled trial. Ann Rheum Dis 2014;73:2074–81.
18. Espigol-Frigole G, Planas-Rigol E, Ohnuki H, Salvucci O, Kwak H, Ravichandran S, et al. Identification of IL-23p19 as an endothelial proinflammatory peptide that promotes gp130-STAT3 signaling. Sci Signal 2016;9:ra28.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 898–908
DOI 10.1002/acr.24201

AMERICAN COLLEGE of RHEUMATOLOGY
Empowering Rheumatology Professionals

Serious Infections in Patients With Gout in the US: A National Study of Incidence, Time Trends, and Outcomes

Jasvinder A. Singh[1] and John D. Cleveland[2]

Objective. To study the epidemiology of serious infections in patients hospitalized with gout.

Methods. We identified patients with gout hospitalized with a primary diagnosis of pneumonia, sepsis/bacteremia, urinary tract infection (UTI), skin and soft tissue infections (SSTIs), or opportunistic infections (OIs) in a US National Inpatient Sample from 1998 to 2016 and examined factors associated with utilization and mortality.

Results. We noted 1,140,085 hospitalizations of patients with serious infections and gout (11% of all hospitalizations of patients with gout; 1998–2000 [8.9%], 2015–2016 [14.5%]). Compared to patients without gout, patients with gout hospitalized with serious infections were older (median age 65 versus 74 years), more of them had a Charlson–Deyo comorbidity index score ≥2 (42% versus 65%), and fewer were female (53% versus 35%) or non-White (40% versus 35%), respectively. The most common infection was pneumonia (52%) in 1998–2000 and sepsis (52%) in 2015–2016. Median hospital charges and hospital stays were higher for patients with sepsis and OIs in 2015–2016 ($41,000–$42,000; 5.1–5.5 days) versus those with UTI, pneumonia, or SSTIs ($15,000–$17,000; 3.0–3.9 days). Compared to patients with sepsis, the multivariable-adjusted odds of health care utilization and in-hospital mortality were significantly lower for patients with UTI, SSTIs, and pneumonia, and non-home discharge or in-hospital mortality were lower in patients with OIs. Among patients hospitalized with infections, older age, Medicaid coverage, a higher Charlson–Deyo comorbidity index score, Black race, and Northeast and nonrural hospital location were associated with significantly higher health care utilization and mortality, while female sex, Medicare insurance, and lower income were associated with higher utilization.

Conclusion. Given an increasing rate of serious infections, especially sepsis and pneumonia, in individuals with gout, development of effective interventions targeting factors associated with health care utilization and mortality will improve outcomes and reduce burden.

INTRODUCTION

Gout is the most common inflammatory arthritis in adults. Gout-related utilization of the emergency department (1–3) or inpatient care (4,5) has a significant cost burden (6,7). Acute and/or chronic symptomatic gout is a significant contributor to its health care utilization burden (8). Gout-related hospitalization burden has increased over time and has surpassed the hospitalization burden associated with rheumatoid arthritis (9). Understanding the causes of hospitalizations of patients with gout and the associated outcomes is important.

Gout is associated with significant comorbidities (10–12), which lead to increased morbidity and mortality and contribute to higher health care utilization (13). In the recent years, cardiovascular and renal comorbidity in gout has been the focus of interest in gout-related comorbidity burden (10). On the other hand, the impact of hospitalizations of patients with serious infections and gout has received little attention.

Few studies have examined hospitalizations of patients with serious infections and gout. Pneumonia and cellulitis were 2 of the top 5 diagnoses for hospitalizations in patients with gout in a study of gout-related hospitalizations in the UK and New Zealand from

Supported by the Division of Rheumatology at the University of Alabama at Birmingham (research funds) and the VAMC, Birmingham, Alabama (resources and use of facilities).

[1]Jasvinder A. Singh, MBBS, MPH: VAMC, Birmingham, Alabama and University of Alabama at Birmingham; [2]John D. Cleveland, MS: University of Alabama at Birmingham.

Dr. Singh has received consulting fees and/or speaking fees from Medisys, Fidia, UBM, Medscape, WebMD, Clinical Care Options, ClearView Healthcare Partners, Putnam Associates, Spherix (less than $10,000 each), Crealta/Horizon, and Simply Speaking (more than $10,000 each) and owns stock or stock options in Amarin and Viking Therapeutics. No other disclosures relevant to this article were reported.

Address correspondence to Jasvinder A. Singh, MBBS, MPH, University of Alabama at Birmingham, Faculty Office Tower 805B, 510 20th Street S, Birmingham, AL 35294. Email: Jasvinder.md@gmail.com.

Submitted for publication November 24, 2019; accepted in revised form March 24, 2020.

SIGNIFICANCE & INNOVATIONS

- Hospitalizations of patients with serious infections constituted 11% of all hospitalizations of patients with a nonprimary diagnosis of gout in the US from 1998 to 2016, the most common of which were pneumonia from 1998 to 2000 and sepsis from 2015 to 2016.
- Sepsis hospitalization rates in patients with gout per 100,000 National Inpatient Sample claims increased 19.3 fold from 1998–2000 to 2015–2016, while hospitalization rates for opportunistic infections (OIs) increased to 2.6 fold, and pneumonia rates increased 2.2 fold, respectively.
- In-hospital mortality was highest for sepsis at 10.1%, followed by OI (4%) and pneumonia (3.1%), and lowest for skin and soft tissue infections (0.5%) and urinary tract infection (0.6%) in patients with gout hospitalized with serious infections.
- Older age, Medicare coverage, Charlson–Deyo comorbidity index score, female sex, Black race, lower income, and Northeast and nonrural hospital location were associated with higher health care utilization and/or higher in-hospital mortality in patients with gout hospitalized with serious infections.

1999 to 2009 (14). In a study from 2009 to 2014 in Australia and New Zealand, cardiovascular disease and infections were the top 2 primary disease categories for hospitalizations among patients with gout (15). A study of individuals in the UK with gout versus those without gout reported that the rate of pneumonia and urinary tract infection (UTI) or infection-related mortality did not differ by gout (16). While the important contribution of serious infections to hospitalizations of patients with gout has now been established in some settings, much remains unknown with regard to outcomes and the impact of hospitalizations of patients with infection.

To our knowledge, there are no published, comprehensive, national epidemiologic studies of infection-related hospitalizations of patients with gout in the US. Our study objectives were the following: 1) to compare the characteristics of patients hospitalized with serious infections with gout versus those without gout in the US; 2) to obtain the estimates of and study the time trends in hospitalizations of patients with serious infection and gout; 3) to examine the health care utilization of patients hospitalized with serious infections and gout; and 4) to assess factors associated with the health care utilization of patients hospitalized with serious infections and gout.

MATERIALS AND METHODS

Data source and study cohort. Our analysis used data from the US National Inpatient Sample (NIS) from 1998 to 2016. The NIS represents a 20% stratified sample of discharges in the US and is a component of the Healthcare Cost and Utilization Project (HCUP). The study cohort included patients admitted to the hospital with an International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) code 274.xx or an ICD-10-CM code of M10 for the diagnosis of gout in any position other than the primary diagnosis position (i.e., nonprimary position; the primary diagnosis of interest was infections) during the index hospitalization. Although there is a risk for misclassification bias, previous studies showed an ICD code–based algorithm for identifying individuals with gout to be a valid and practical approach for identifying those individuals for database studies with reasonable accuracy, a sensitivity of 90%, a specificity of 100%, and positive predictive values of 80% or higher (17–19), although some studies reported lower accuracy (20,21).

The University of Alabama at Birmingham's Institutional Review Board approved this database study and waived the need for informed consent. All investigations were conducted in conformity with ethical principles of research.

Hospitalizations of patients with serious infections of interest and covariates. We identified 5 common serious infections in patients hospitalized with the primary diagnosis position for hospitalization using ICD-9-CM codes based on overall NIS infection hospitalization frequencies. These codes for infections were valid in administrative data sets, with positive predictive values of 70–100% in patients with rheumatoid arthritis (22–24). Infections of interest included the following: 1) pneumonia (003.22, 481.0, 513.0, 480.xx, 482.xx, 483.xx, 485.xx, and 486.xx); 2) sepsis/bacteremia (referred to as sepsis from here onwards; 038.xx and 790.7); 3) UTI (590.xx); 4) skin and soft tissue infections (SSTIs; 040.0, 569.61, 681.xx, 682.xx, 785.4, 728.86, and 035.xx); and 5) opportunistic infections (OIs; 010.xx–018.xx, 031.xx, 078.5, 075.xx, 053.xx, 112.4, 112.5, 112.81, 112.83, 130.xx, 136.3, 117.5, 027.0, 039.xx, 117.3, 114.xx, 115.xx, 116.0), as previously reported (25,26). When multiple ICD-9-CM codes for the same infection category occurred in the same hospitalization, such as 480.xx and 482.xx in the same discharge, pneumonia would only be counted once. We also used the ICD-10-CM codes for infections for the 2015–2016 data because the coding system changed from the ICD-9-CM to ICD-10-CM in 2015 in the US (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24201/abstract). We defined a composite infection as any of the 5 serious infection categories during hospitalization; for frequencies, this equates to the sum of all 5 infections.

Covariates included demographics (age, sex, race/ethnicity, income), comorbidity, insurance type, and hospital characteristics. Age was categorized as <50, 50 to <65, 65 to <80 and ≥80 years, and race/ethnicity was categorized as White, Black, Hispanic, and other/missing in order to correspond to clinically relevant categories for patients with gout based on an a priori decision. This approach to age categorization is similar to that used in previously published studies using NIS data (27–29) and allows comparisons to studies that use Medicare data (mostly for individuals ≥65 years). Household income, based on the patient's zip code, was categorized from

the lowest (poorest) to the highest quartile (wealthiest) (4 quartiles). Thresholds for each quartile varied by year as provided by the NIS (30); e.g., the upper threshold for quartile 1 was $28,999 in 1998 and $39,999 in 2014. Comorbidity was assessed by the Charlson–Deyo comorbidity index, a valid measure consisting of 17 common medical comorbidities (myocardial infarction, congestive heart failure, cerebrovascular disease, dementia, renal disease, liver disease, chronic pulmonary disease, diabetes mellitus, etc.) based on the presence of ICD-9-CM codes at index admission (31). Charlson–Deyo comorbidity index score was categorized as none, 1 or 2, or >2, as previously reported (32–34). Insurance status was categorized as Medicaid, Medicare, private insurance, and self/other (35), as previously reported (36). Medicaid provides coverage for low-income and disabled Americans, and Medicare provides health care coverage for Americans age ≥65 years. Hospitals were categorized based on location, and teaching status was categorized as rural, urban, or urban teaching hospital. Hospital region and hospital bed size were other standard NIS variables included in the models.

Table 1. Characteristics of all patients hospitalized with gout and of patients hospitalized with serious infections without gout versus with gout*

Characteristic	All patients hospitalized with gout (n = 10,405,447)†	Patients hospitalized with serious infections without gout (n = 48,826,485)†	Patients hospitalized with serious infections with gout (n = 1,140,085)†
Age, mean ± SE; median years	70.4 ± 0.03; 71.6	59.5 ± 0.08; 64.7	72.1 ± 0.04; 73.7
Age category, years			
<50	786,867 (7.56)	14,004,211 (28.90)	78,460 (6.90)
50 to <65	2,406,433 (23.13)	9,769,839 (20.16)	226,172 (19.89)
65–79	4,272,086 (41.06)	12,846,021 (26.51)	441,964 (38.87)
≥80	2,939,428 (28.25)	11,835,309 (24.43)	390,326 (34.33)
Sex			
Male	6,989,356 (67.18)	22,718,878 (46.91)	740,313 (65.12)
Female	3,414,355 (32.82)	25,710,174 (53.09)	396,514 (34.88)
Race			
White	6,414,610 (61.65)	29,015,095 (59.86)	743,776 (65.42)
Black	1,662,952 (15.98)	5,193,503 (10.71)	149,575 (13.16)
Hispanic	401,389 (3.86)	4,174,347 (8.61)	47,529 (4.18)
Other/missing	1,926,178 (18.51)	10,090,131 (20.82)	196,063 (17.24)
Charlson–Deyo comorbidity index score			
0	1,651,015 (15.87)	15,510,561 (32.00)	173,267 (15.24)
1	1,930,307 (18.55)	12,715,215 (26.23)	222,160 (19.54)
≥2	6,824,126 (65.58)	20,252,048 (41.78)	741,552 (65.22)
Income category			
0–25th percentile	2,582,576 (25.32)	12,696,846 (26.83)	282,831 (25.36)
25–50th percentile	2,596,043 (25.46)	13,017,631 (27.51)	287,677 (25.79)
50–75th percentile	2,493,249 (24.45)	11,342,129 (23.97)	275,439 (24.70)
75–100th percentile	2,526,666 (24.77)	10,265,741 (21.69)	269,365 (24.15)
Insurance			
Private	2,082,338 (20.04)	10,772,518 (22.27)	184,373 (16.24)
Medicare	7,343,196 (70.69)	26,623,425 (55.04)	852,152 (75.06)
Medicaid	540,299 (5.20)	7,029,774 (14.53)	57,669 (5.08)
Other	227,919 (2.19)	1,480,898 (3.06)	21,613 (1.90)
Self	194,623 (1.87)	2,466,366 (5.10)	19,450 (1.71)
Hospital location/teaching			
Rural	1,271,228 (12.25)	6,881,175 (14.96)	151,521 (13.76)
Urban nonteaching	3,762,821 (36.27)	18,833,673 (40.93)	422,336 (38.36)
Urban teaching	5,341,651 (51.48)	20,294,696 (44.11)	527,111 (47.88)
Hospital charges, mean ± SE; median $	39,044 ± 223; 22,672	34,554 ± 167; 16,743	37,804 ± 237; 21,071
1998–2000	15,293 ± 232; 9,562	13,111 ± 264; 8,293	18,287 ± 340; 9,615
2015–2016	55,767 ± 542; 34,265	53,647 ± 433; 28,717	51,210 ± 546; 30,538
Total hospital charges > median, $	6,711,966 (64.50)	27,785,443 (57.32)	674,040 (59.28)
Length of hospital stay, mean ± SE; median days	5.4 ± 0.01; 3.4	6.0 ± 0.01; 3.7	6.1 ± 0.02; 4.1
Proportion with length of hospital stay > the median of 3 days	5,679,276 (54.58)	28,746,958 (59.30)	746,117 (65.62)
Discharge status			
Rehabilitation or nursing facility	2,450,676 (24.26)	11,357,982 (25.33)	309,575 (28.79)
Home	7,649,867 (75.74)	33,479,292 (74.67)	765,625 (71.21)
Died during hospitalization	233,511 (2.25)	3,024,281 (6.24)	54,788 (4.82)

* Values are the number (%) unless indicated otherwise.
† These are national estimates using the recommended weights from the National Inpatient Sample and are based on the 20% national sample of all hospitalizations.

Health care utilization outcomes and in-hospital mortality. The health care utilization outcomes of interest included the total hospital charges above the median (using the standard NIS variable TOTALCHG), the length of hospital stay (using the standard NIS variable LOS and categorized as >3 days versus ≤3 days based on the overall NIS median), and the proportion of patients discharged to a rehabilitation or nursing facility (i.e., a short-term hospital, skilled nursing facility, intermediate care facility, or another type; using the standard NIS variable DISPUNIFORM) rather than home. We also assessed in-hospital mortality.

Statistical analysis. We compared demographics for all infection-related hospitalizations in patients with gout versus those without gout in the nonprimary position. The rest of the analyses were limited to infection-related hospitalizations in patients with gout. Characteristics of patients hospitalized for each type of infection were compared to those of patients with gout. For time trends from 1998 to 2016, we examined the frequencies and rates of the 5 infections and analyzed each for trends over time using the Cochran-Armitage test, weighted by the number of hospitalizations in each year. We used adjusted logistic regression models for length of hospital stay >3 days (median), total hospital charges above the median (based on the median for the year), discharge status (home versus inpatient), and in-hospital mortality. Both the length of hospital stay and the total hospital charges had a left-skewed distribution (as expected, with a few very high values). This made it most appropriate for a logistic regression to avoid undue influence of extreme values on a linear regression. We calculated odds ratios and 95% confidence intervals.

RESULTS

Characteristics of patients hospitalized with infection: gout versus no gout. Using weighted data, we observed 48,826,485 serious infections in patients hospitalized without gout and 1,140,085 infections in patients hospitalized with gout. Hospitalizations of patients with serious infections constituted 11% of all hospitalizations of patients with a nonprimary diagnosis of gout (1.14 million of 10.4 million) (Table 1). The mean age of patients with gout hospitalized with a primary diagnosis of one of the serious infections was 72.1 years (median 73.8 years), two-thirds were male, 65% were White, the Charlson–Deyo comorbidity index score was ≥2 in 65%, and 75% had Medicare as the insurance payer (Table 1).

Compared to patients hospitalized with a serious infection without gout, patients with gout hospitalized with a serious infection were a decade older (median age 65 versus 74 years), less likely to be female (53% versus 35%) or non-White (40% versus 35%), more likely to be in the lowest income quartile (22% versus 24%), had a Charlson–Deyo comorbidity index score of ≥2 (42% versus 65%), had Medicare as the insurance payer (55% versus 75%), or were admitted to an urban teaching hospital (44% versus 48%), respectively (Table 1). Unadjusted total hospital charges, the length of hospital stay, and discharge to a non-home setting were slightly higher in patients with gout versus those without gout, and in-hospital mortality was slightly lower (Table 1).

Patients with gout hospitalized with a serious infection: time trends in rates overall and each serious infection. Hospitalizations of patients with serious infections and gout increased from 8.9% (67,995 of 762,576) of all hospitalizations in patients with gout in 1998–2000 to 14.5% (250,510 of 1,721,300) in 2015–2016 (Table 2 and Table 3). In 1998–2000, pneumonia was the most

Table 2. Frequency of hospitalizations of patients with primary diagnosis of various types of serious infection and a nonprimary diagnosis of gout over time*

Study period	OI	SSTI	UTI	Pneumonia	Sepsis	Composite infection
1998–2000†	1,220	19,905	1,756	35,434	9,679	67,995
2001–2002	762	16,721	1,360	26,322	7,715	52,880
2003–2004	925	21,176	1,775	31,783	9,915	65,574
2005–2006	1,265	27,345	2,069	36,812	13,801	81,292
2007–2008	1,409	32,067	2,300	41,299	22,083	99,157
2009–2010	1,844	40,619	2,847	48,414	43,098	136,822
2011–2012	2,164	46,950	3,529	56,896	70,816	180,356
2013–2014	2,220	46,610	3,360	55,970	97,340	205,500
2015–2016‡	2,160	47,860	18,085	53,600	128,805	250,510

* Values are the number. The primary position codes from the International Classification of Diseases, Ninth Revision (ICD-9) used to define each infection category were as follows: opportunistic infections (OIs) included tuberculosis (010–018), nontuberculous mycobacteria (031), cytomegalovirus (078.5), Epstein-Barr virus (075), herpes zoster (053), candidiasis (112.4, 112.5, 112.81, 112.83), toxoplasmosis (130), pneumocystosis (136.3), cryptococcosis (117.5), listeriosis (027.0), nocardiosis (039), aspergillosis (117.3), coccidioidomycosis (114), histoplasmosis (115), and blastomycosis (116.0); skin and soft tissue infections (SSTIs) (040.0, 569.61, 681, 682, 785.4, 728.86, and 035); urinary tract infection (UTI) (590); pneumonia (003.22, 481.0, 513.0, 480, 482, 483, 485, and 486); and sepsis/bacteremia (038 and 790.7). Composite infection = any of the 5 infections listed above.
† The first study period consisted of 3 years, and all subsequent periods were 2 years each.
‡ 2015 was the first year that ICD-10 Clinical Modification codes were used; therefore, some rates from this period may reflect the transition of the coding system rather than only time trends (i.e., these estimates may be a little unstable).

Table 3. Time trends in the rates of hospitalizations of patients with serious infections in the general National Inpatient Sample (NIS) cohort per 100,000 NIS claims and of hospitalizations of patients with serious infections among individuals with gout using 2 denominators (per 100,000 gout claims and per 100,000 NIS claims)*

	OI	SSTI	UTI	Pneumonia	Sepsis	Composite infection	Total claims
Hospitalized infection rates in the general NIS cohort per 100,000 NIS claims							
1998–2000†	159.5	974.9	313.5	3,398.4	989.5	5,835.8	103,665,051
2001–2002	139.4	1,097.6	335.2	3,303.2	919.3	5,794.8	72,617,381
2003–2004	138.9	1,269.2	352.3	3,260.2	1,088.6	6,109.3	74,571,583
2005–2006	143.9	1,490.7	350.9	3,236.0	1,451.3	6,672.7	75,919,595
2007–2008	138.6	1,531.7	337.0	2,891.9	1,834.9	6,734.1	76,366,797
2009–2010	138.6	1,627.2	351.0	2,849.1	2,223.6	7,189.5	75,086,597
2011–2012	130.5	1,688.9	343.2	2,791.8	2,949.0	7,903.3	73,447,261
2013–2014	121.4	1,665.6	324.3	2,578.2	3,939.6	8,629.1	70,956,610
2015–2016‡	110.9	1,667.9	882.1	2,401.0	5,109.1	10,171.1	71,445,363
Change from 1998–2000 to 2015–2016§	0.7	1.7	2.8	0.7	5.2	1.7	0.7
Hospitalized infection rates in the gout cohort per 100,000 gout claims							
1998–2000†	160.0	2,610.2	230.3	4,646.6	1,269.3	8,916.5	762,576
2001–2002	124.3	2,728.3	221.9	4,294.8	1,258.8	8,628.1	612,880
2003–2004	125.1	2,864.3	240.1	4,299.1	1,341.1	8,869.7	739,301
2005–2006	144.6	3,125.3	236.5	4,207.3	1,577.3	9,290.9	874,964
2007–2008	130.0	2,959.3	212.3	3,811.3	2,038.0	9,150.8	1,083,586
2009–2010	138.8	3,057.0	214.3	3,643.6	3,243.6	10,297.2	1,328,731
2011–2012	133.9	2,904.0	218.3	3,519.2	4,380.1	11,155.4	1,616,755
2013–2014	133.3	2,798.8	201.8	3,360.8	5,845.0	12,339.7	1,665,355
2015–2016‡	125.5	2,780.5	1,050.66	3,113.9	7,483.0	14,553.5	1,721,300
Change from 1998–2000 to 2015–2016§	0.8	1.1	4.6	0.7	5.9	1.6	2.3
Hospitalized infection rates in the gout cohort per 100,000 NIS claims							
1998–2000†	1.2	19.2	1.7	34.2	9.3	65.6	103,665,051
2001–2002	1.1	23.0	1.9	36.3	10.6	72.8	72,617,381
2003–2004	1.2	28.4	2.4	42.6	13.3	87.9	74,571,583
2005–2006	1.7	36.0	2.7	48.5	18.2	107.1	75,919,595
2007–2008	1.9	42.0	3.0	54.1	28.9	129.8	76,366,797
2009–2010	2.5	54.1	3.8	64.5	57.4	182.2	75,086,597
2011–2012	3.0	63.9	4.8	77.5	96.4	245.6	73,447,261
2013–2014	3.1	65.7	4.7	78.9	137.2	289.6	70,956,610
2015–2016‡	3.0	67.0	25.31	75.0	180.3	350.6	71,445,363
Change from 1998–2000 to 2015–2016§	2.6	3.5	14.9	2.2	19.3	5.3	0.7

* Values are the number. OI = opportunistic infection; SSTI = skin and soft tissue infection; UTI = urinary tract infection.
† The first study period consisted of 3 years, not 2 years, and all subsequent periods were 2 years each.
‡ 2015–2016 was the first year that International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification codes were used; therefore, some rates from this period may reflect the transition of the coding system rather than only time trends.
§ Change was calculated by dividing the number in 2015–2016 by the number from 1998–2000.

common infection (52%; 35,434 of 67,995) followed by SSTI (29%; 19,905 of 67,995) and sepsis (14%; 9,679 of 67,995). In 2015–2016, sepsis was the most common infection (52%;128,805 of 250,510), followed by pneumonia (21%; 53,600 of 250,510) and then SSTI (19%; 47,860 of 250,510; Table 2). Over the 20-year period, sepsis (36%), followed by pneumonia (34%), were the top 2 causes of hospitalizations of patients with serious infections and gout (Table 2).

The frequencies of all 5 types of serious infections in patients hospitalized with gout increased significantly over the study period (Table 2). Rates per 100,000 NIS claims in 2015–2016 were as follows: OI, 3; SSTI, 67; UTI, 25; pneumonia, 75; sepsis, 180; and composite infection, 351 (Table 3 and Figure 1). The increase in the rate of sepsis claims massively outpaced those for all other hospitalizations of patients with serious infections, increasing 5.9 fold from 1998–2000 to 2015–2016 (using total gout claims as the denominator); SSTI claims increased 1.1 fold; UTI claims, 4.6 fold; and OI and pneumonia claims decreased 0.7 and 0.8 fold, respectively (Table 3). Examining these claims per 100,000 NIS claims, the rates in 2015–2016 were 2.6, 3.5, 14.9, 2.2, and 19.3 fold higher, respectively. Starting in 2011–2012, sepsis outnumbered pneumonia as the primary infection diagnosis for patients hospitalized with serious infections and gout (Table 3 and Figure 1).

Characteristics of patients hospitalized for each type of serious infection and outcomes. Two-thirds of patients with gout hospitalized with serious infections were White, male, had a Charlson–Deyo comorbidity index score ≥2, and had Medicare as payer type (Table 4). Patients with gout hospitalized with

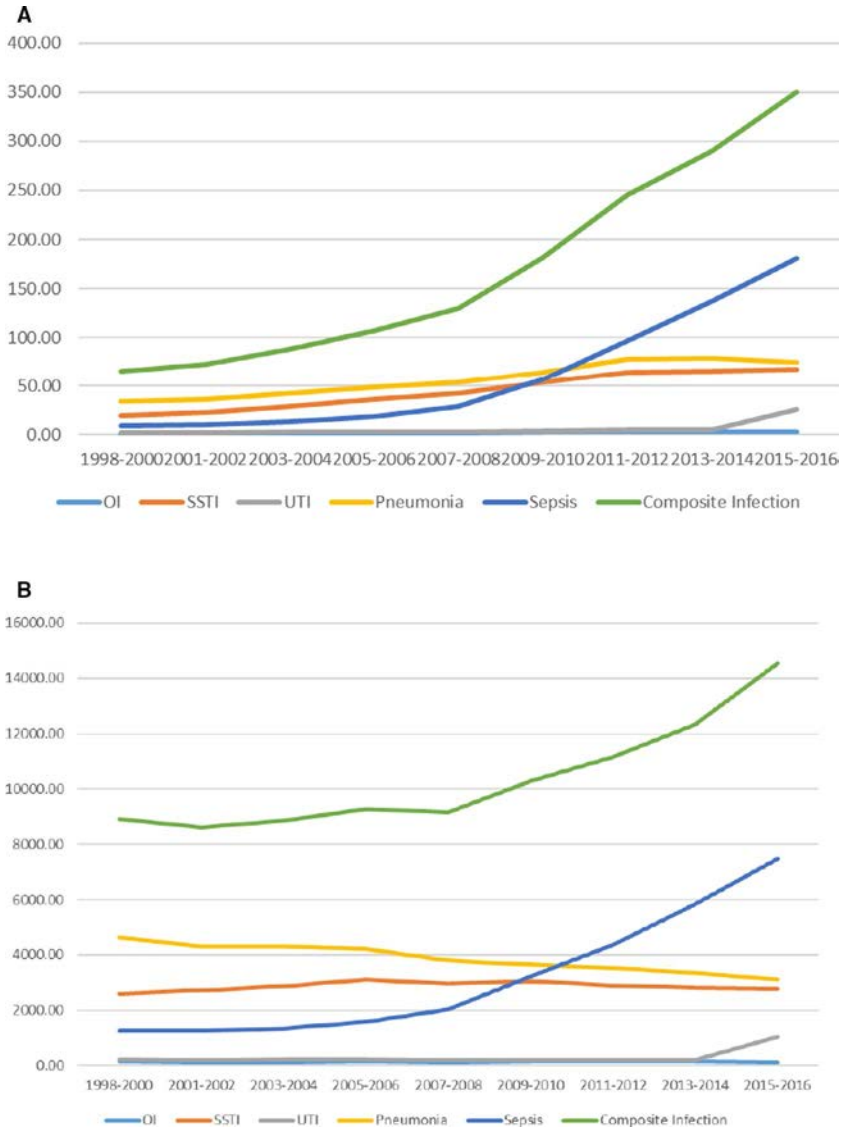

Figure 1. Rates of infection in patients hospitalized with gout per 100,000 total National Inpatient Sample claims (**A**) and per 100,000 overall gout claims (**B**). Infection rates are per 100,000 claims. The y-axis scales are different for the 2 panels. The x-axis shows the study time periods from 1998 to 2016. OI = opportunistic infection; SSTI = skin and soft tissue infection; UTI = urinary tract infection.

pneumonia or sepsis were 5 years older than those admitted with an SSTI or OI, at 74–76 years versus 69–70 years, respectively (Table 4). More than two-thirds of patients with gout hospitalized with each serious infection were male, except for UTI, 45% were male (Table 4).

The median length of hospital stay over the study period was highest in hospitalizations for OI and sepsis, at 5.5 and 5.1 days, respectively, and lowest for UTI, at 3 days (Table 4). Above-median length of hospital stay was highest, at 63–74%, in patients discharged for sepsis, pneumonia, and OI. The median hospital charges were highest for sepsis ($34,864) and OI ($28,612), followed by pneumonia ($17,511) (Table 4). We noted that 40% of patients with sepsis and gout versus 20–27% of patients with other serious infections were discharged to a non-home setting (Table 4). In-patient mortality for patients with gout hospitalized with a serious infection was highest for sepsis, at 10.1%, followed by 4% for OI, 3.1% for pneumonia, 0.6% for UTI, and 0.5% for SSTI (Table 4).

Time trends in health care utilization and mortality in patients with gout hospitalized with serious infection. Median hospital charges and hospital stay were much higher for sepsis and OI in 2015–2016 ($41,000–$42,000; 4.8–5.5 days) compared to those for UTI, pneumonia, or SSTI ($21,000–$26,000; 3.0–3.6 days) (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24201/abstract). The largest increase in median hospital charges was seen for sepsis, at 3.6 fold, followed by that for OI (3.5 fold), SSTI (3.2 fold), UTI (3.2 fold), and pneumonia (3 fold) (see Supplementary Table 2).

Table 4. Characteristics of patients with gout hospitalized with various types of serious infections*

Characteristic	OI (n = 13,941)	SSTI (n = 298,394)	UTI (n = 36,978)	Pneumonia (n = 384,723)	Sepsis (n = 402,944)	Composite infection (n = 1,136,980)
Age, mean ± SE; median years	69.0 ± 0.28; 70.1	68.2 ± 0.07; 69.3	72.3 ± 0.17; 74.2	74.6 ± 0.05; 76.4	72.6 ± 0.05; 73.9	72.1 ± 0.04; 73.7
Age category, years						
<50	1,327 (9.5)	34,980 (11.7)	2,752 (7.4)	17,744 (4.6)	21,657 (5.4)	78,460 (6.9)
50 to <65	3,639 (26.1)	78,772 (26.4)	6,951 (18.8)	58,826 (15.3)	77,984 (19.4)	226,172 (19.9)
65 to <80	5,322 (38.2)	107,823 (36.1)	13,952 (37.7)	150,159 (39.0)	164,708 (40.9)	441,964 (38.9)
≥80	3,652 (26.2)	76,811 (25.7)	13,318 (36.0)	157,974 (41.1)	138,570 (34.4)	390,326 (34.3)
Sex						
Male	9,621 (69.0)	196,499 (65.9)	16,693 (45.2)	252,348 (65.6)	265,152 (65.8)	740,313 (65.1)
Female	4,319 (31.0)	101,854 (34.1)	20,275 (54.8)	132,310 (34.4)	137,756 (34.2)	396,514 (34.9)
Race						
White	7,694 (55.2)	199,528 (66.9)	23,774 (64.3)	249,825 (64.9)	262,954 (65.3)	743,776 (65.4)
Black	2,182 (15.6)	33,310 (11.2)	5,615 (15.2)	47,756 (12.4)	60,713 (15.1)	149,575 (13.2)
Hispanic	768 (5.5)	13,439 (4.5)	1,853 (5.0)	12,774 (3.3)	18,695 (4.6)	47,529 (4.2)
Other/missing	3,297 (23.6)	52,113 (17.5)	5,736 (15.5)	74,341 (19.3)	60,577 (15.0)	196,063 (17.2)
Charlson–Deyo comorbidity index score						
0	2,926 (21.0)	69,603 (23.3)	6,504 (17.6)	46,415 (12.1)	47,818 (11.9)	173,267 (15.2)
1	2,735 (19.6)	68,632 (23.0)	7,536 (20.4)	82,853 (21.5)	60,404 (15.0)	222,160 (19.5)
≥2	8,279 (59.4)	160,159 (53.7)	22,938 (62.0)	255,455 (66.4)	294,721 (73.1)	741,552 (65.2)
Income category						
0–25th percentile	3,275 (24.1)	73,982 (25.3)	10,208 (28.2)	94,339 (25.0)	101,028 (25.5)	282,831 (25.4)
25–50th percentile	3,290 (24.2)	75,000 (25.7)	9,609 (26.5)	101,230 (26.8)	98,547 (24.9)	287,677 (25.8)
50–75th percentile	3,206 (23.6)	71,431 (24.5)	8,479 (23.4)	92,713 (24.6)	99,609 (25.2)	275,439 (24.7)
75–100th percentile	3,802 (28.0)	71,652 (24.5)	7,920 (21.9)	89,254 (23.6)	96,737 (24.4)	269,365 (24.2)
Insurance						
Private	2,849 (20.5)	63,725 (21.4)	6,185 (16.7)	51,321 (13.4)	60,293 (15.0)	184,373 (16.2)
Medicare	9,556 (68.6)	197,738 (66.4)	27,492 (74.5)	307,242 (80.0)	310,125 (77.0)	852,152 (75.1)
Medicaid	944 (6.8)	19,606 (6.6)	2,075 (5.6)	15,271 (4.0)	19,772 (4.9)	57,669 (5.1)
Other	276 (2.0)	7,754 (2.6)	586 (1.6)	5,751 (1.5)	7,247 (1.8)	21,613 (1.9)
Self	296 (2.1)	9,016 (3.0)	587 (1.6)	4,483 (1.2)	5,068 (1.3)	19,450 (1.7)
Hospital region						
Northeast	2,826 (20.2)	67,095 (22.4)	6,149 (16.6)	76,938 (19.9)	68,959 (17.1)	221,967 (19.5)
Midwest	3,296 (23.6)	76,279 (25.5)	9,028 (24.3)	99,753 (25.8)	94,730 (23.5)	283,086 (24.8)
South	4,602 (32.9)	107,134 (35.8)	15,078 (40.7)	137,706 (35.6)	135,853 (33.7)	400,372 (35.1)
West	3,246 (23.2)	48,745 (16.3)	6,827 (18.4)	72,134 (18.7)	103,709 (25.7)	234,660 (20.6)
Hospital location/teaching						
Rural	1,015 (7.4)	38,445 (13.4)	5,174 (14.4)	63,580 (17.3)	43,307 (10.9)	151,521 (13.8)
Urban nonteaching	4,492 (32.8)	118,148 (41.1)	12,611 (35.2)	145,290 (39.6)	141,795 (35.7)	422,336 (38.4)
Urban teaching	8,191 (59.8)	131,028 (45.6)	18,077 (50.4)	158,200 (43.1)	211,614 (53.3)	527,111 (47.9)
Hospital bed size						
Small	1,594 (11.4)	51,469 (17.2)	7,337 (19.8)	69,823 (18.1)	62,993 (15.7)	193,216 (17.0)
Medium	3,227 (23.1)	81,605 (27.3)	10,355 (28.0)	105,725 (27.4)	112,939 (28.1)	313,852 (27.6)
Large	9,129 (65.4)	165,549 (55.4)	19,285 (52.2)	210,053 (54.5)	226,442 (56.3)	630,458 (55.4)
Total hospital charges, mean ± SE; median $	53,673 ± 1,557; 28,612	23,125 ± 183; 15,013	25,532 ± 430; 17,344	28,564 ± 212; 17,511	58,766 ± 459; 34,864	37,804 ± 237; 21,071
Total hospital charges > median, $	9,853 (70.7)	140,853 (47.2)	16,455 (44.5)	221,933 (57.7)	284,946 (70.7)	674,040 (59.3)
Length of hospital stay, mean ± SE; median days	8.9 ± 0.21; 5.5	5.1 ± 0.02; 3.6	4.5 ± 0.05; 3.0	5.5 ± 0.02; 3.9	7.4 ± 0.03; 5.1	6.1 ± 0.02; 4.1
Proportion with length of hospital stay > the median of 3 days	10,234 (73.4)	177,219 (59.4)	18,707 (50.6)	243,234 (63.2)	296,723 (73.6)	746,117 (65.6)
Discharge status						
Rehabilitation or nursing facility	3,105 (23.3)	57,927 (19.7)	9,762 (26.7)	94,072 (25.4)	144,708 (40.2)	309,575 (28.8)
Home	10,203 (76.7)	236,853 (80.3)	26,799 (73.3)	276,586 (74.6)	215,184 (59.8)	765,625 (71.2)
Died during hospitalization	558 (4.0)	1,389 (0.5)	216 (0.6)	11,916 (3.1)	40,708 (10.1)	54,788 (4.8)

* Values are the number (%) unless indicated otherwise. OI = opportunistic infection; SSTI = skin and soft tissue infection; UTI = urinary tract infection.

The length of hospital stay decreased over time for all patients hospitalized with infections (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24201/abstract). In-hospital mortality in patients with gout decreased for those hospitalized with sepsis (10.6% to 8.7%), pneumonia (4% to 2.8%), and an SSTI (0.7% to 0.4%) and was stable for those with an OI (5.1% to 5.3%) and a UTI (0.6% to 0.6%) (see Supplementary Table 2).

Factors associated with health care utilization and mortality in patients hospitalized with serious infections and gout. Multivariable-adjusted analyses showed no significant difference in the odds of above-median hospital charges or above-median hospital stay between sepsis and OI, but significantly lower odds of discharge to a care facility and in-hospital mortality were seen in OI versus sepsis (Table 5). Adjusted odds of health care utilization and in-hospital mortality were lower in UTI, SSTI, and pneumonia compared to sepsis (Table 5). Compared to sepsis, other serious infections were associated with odds ratios between 0.05 and 0.42 for in-hospital mortality, depending on the type of serious infection.

Older age, Medicaid insurance, higher Charlson–Deyo comorbidity index score, Black race, and Northeast and nonrural hospital location were associated with higher health care utilization and higher in-hospital mortality (except discharge to a care facility, which was higher for the Midwest); female sex and Medicare insurance (reference = private insurance) were associated with higher health care utilization; and lower income was associated with higher odds of discharge to a care facility in patients with gout hospitalized with serious infections (Table 5).

DISCUSSION

We studied the epidemiology of serious infections in patients hospitalized with gout using national sample data over 2 decades. We examined time trends, patient characteristics, and outcomes of 5 common serious infection hospitalizations in patients with gout, as well as factors associated with health care utilization and in-hospital mortality for patients with serious infections. Several findings merit further discussion.

Compared to patients without gout, patients with gout hospitalized with a serious infection were a decade older, more likely to have Charlson–Deyo comorbidity index score ≥2, more likely to have Medicare as the insurance payer, and less likely to be female. These differences might partially explain higher unadjusted hospital charges and a longer stay in patients with gout hospitalized with serious infections compared to those without gout.

Our study is among the first to describe the epidemiology of serious infections in patients hospitalized with gout in the US. In unadjusted analyses, hospitalizations of patients with serious infections accounted for 11% of all hospitalizations in patients with a nonprimary diagnosis of gout. Over time, the proportion of patients hospitalized with serious infections increased from 8.9% in 1998–2000 to 14.5% in 2015–2016. The increase in the rate of hospitalizations of patients with serious infections over time during the study period replicates the findings of a recent study from the UK and New Zealand for a US population (14) and extends the findings to a more contemporary period.

The overall rate of hospitalizations of patients with serious infections in our study substantiates the findings from the earlier reports from New Zealand and Australia that infections were among the top 5 reasons for hospitalizations in patients with gout (14,15). Our findings support the increasing relative and absolute contribution of serious infections to overall hospitalizations of patients with gout in the US given the increasing rate of hospitalizations for gout in the US (9).

The frequencies of all 5 types of infections in patients hospitalized with gout increased significantly over the observation period, faster for some serious infections (sepsis, SSTI) than others. Specifically, the hospitalization rate for patients with serious infections per 100,000 NIS claims in the general population from 1998 to 2000 versus 2015 to 2016 (and increase/decrease) varied by the type of infection: OI, 159 versus 110 (0.7 fold); SSTI, 975 versus 1,667 (1.7 fold); UTI, 313 versus 882 (2.8 fold); pneumonia, 3,398 versus 2,401 (0.7 fold); and sepsis, 989 versus 5,109 (5.2 fold). The corresponding respective rates in the gout cohort per 100,000 NIS claims were: OI, 1.2 versus 3.0 (2.6 fold); SSTI, 19.2 versus 67 (3.5 fold); UTI, 1.7 versus 25.3 (14.9 fold); pneumonia, 34.2 versus 75 (2.2 fold); and sepsis, 9.3 versus 180.3 (19.3 fold). Pneumonia was the most common serious infection in patients hospitalized with gout in 1998–2000. Rates of hospitalizations for sepsis increased steeply and crossed over to a higher rate than for pneumonia in 2011–2012, making it the most common serious infection in patients hospitalized in 2015–2016.

Patient and comorbidity characteristics of individuals hospitalized with serious infections and gout matched closely with those of all hospitalizations in people with gout, as expected. Not surprisingly, compared to patients without gout, unadjusted hospital charges, the length of hospital stay, and discharge to a non-home setting were slightly higher in patients with gout hospitalized with a serious infection.

Over time, health care utilization for all patients hospitalized with serious infections increased, and in-hospital mortality decreased, for patients with gout. Median hospital charges increased in sepsis 3.6 fold, in OIs 3.5 fold, in SSTIs 3.2 fold, in UTI 3.2 fold, and in pneumonia 3 fold (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24201/abstract). Thus, hospital charges for all hospitalizations of patients with serious infections and gout increased 3.0–3.6 fold, which parallels the increase in hospital charges for hospitalizations in the general population over time in the US overall. In-hospital mortality for patients with gout was highest in sepsis, at 10.1%, and lowest in SSTI, at 0.5%. In-hospital mortality decreased from 1998–2000 to 2015–2016 for all infections, except OI (which increased slightly from 5.1% to

Table 5. Multivariable-adjusted correlates of health care utilization and mortality in patients hospitalized with gout*

	Hospital charges > median	Length of hospital stay >3 days	Discharge to a care facility	In-hospital mortality
Age category, years				
<50	Ref.	Ref.	Ref.	Ref.
50 to <65	1.07 (1.03–1.12)†	1.13 (1.09–1.18)†	1.74 (1.63–1.85)†	1.55 (1.34–1.79)†
65 to <80	1.04 (0.99–1.08)	1.16 (1.11–1.21)†	2.57 (2.41–2.75)†	2.37 (2.06–2.74)†
≥80	0.95 (0.91–1.00)	1.26 (1.20–1.31)†	5.26 (4.93–5.63)†	3.85 (3.33–4.45)†
Sex				
Male	Ref.	Ref.	Ref.	Ref.
Female	1.04 (1.02–1.06)†	1.15 (1.13–1.17)†	1.30 (1.27–1.32)†	0.97 (0.93–1.01)
Race/ethnicity				
White	Ref.	Ref.	Ref.	Ref.
Black	1.11 (1.08–1.14)†	1.11 (1.08–1.15)†	1.10 (1.07–1.14)†	1.11 (1.05–1.18)†
Hispanic	1.51 (1.44–1.59)†	1.05 (1.00–1.10)	0.76 (0.72–0.81)†	0.97 (0.87–1.07)
Other/missing	1.02 (0.99–1.04)	1.03 (1.00–1.06)	0.84 (0.82–0.87)†	1.03 (0.97–1.09)
Charlson–Deyo comorbidity index score				
0	Ref.	Ref.	Ref.	Ref.
1	1.26 (1.22–1.30)†	1.23 (1.19–1.27)†	1.25 (1.20–1.29)†	1.27 (1.15–1.39)†
≥2	1.50 (1.46–1.54)†	1.52 (1.48–1.56)†	1.56 (1.51–1.61)†	1.86 (1.72–2.01)†
Income category				
0–25th percentile	0.98 (0.95–1.01)	1.01 (0.98–1.04)	1.08 (1.05–1.11)†	0.99 (0.93–1.06)
25–50th percentile	0.94 (0.92–0.96)†	1.02 (0.99–1.05)	1.06 (1.03–1.09)†	0.98 (0.92–1.04)
50–75th percentile	0.94 (0.91–0.96)†	1.00 (0.97–1.02)	1.04 (1.01–1.07)†	0.94 (0.89–0.99)
75–100th percentile	Ref.	Ref.	Ref.	Ref.
Primary infection diagnosis				
Sepsis	Ref.	Ref.	Ref.	Ref.
OI	0.98 (0.90–1.07)	1.05 (0.96–1.15)	0.52 (0.47–0.57)†	0.42 (0.34–0.51)†
SSTI	0.39 (0.38–0.40)†	0.55 (0.54–0.57)†	0.40 (0.38–0.41)†	0.05 (0.04–0.06)†
UTI	0.35 (0.34–0.37)†	0.36 (0.34–0.38)†	0.49 (0.46–0.52)†	0.05 (0.04–0.07)†
Pneumonia	0.62 (0.61–0.64)†	0.61 (0.60–0.62)†	0.42 (0.41–0.43)†	0.28 (0.27–0.29)†
Insurance payer				
Medicare	1.12 (1.09–1.15)†	1.17 (1.13–1.20)†	1.63 (1.57–1.70)†	1.00 (0.93–1.08)
Medicaid	1.30 (1.25–1.37)†	1.23 (1.17–1.29)†	1.44 (1.35–1.53)†	1.14 (1.01–1.29)†
Other	1.12 (1.05–1.20)†	0.99 (0.93–1.06)	1.23 (1.12–1.34)†	1.54 (1.32–1.79)†
Private	Ref.	Ref.	Ref.	Ref.
Self	1.14 (1.06–1.23)†	1.02 (0.95–1.09)	0.66 (0.58–0.75)†	1.44 (1.17–1.77)†
Hospital region				
Northeast	Ref.	Ref.	Ref.	Ref.
Midwest	0.71 (0.69–0.73)†	0.78 (0.76–0.80)†	1.06 (1.03–1.09)†	0.82 (0.77–0.87)†
South	0.79 (0.77–0.81)†	0.86 (0.84–0.88)†	0.81 (0.78–0.83)†	0.94 (0.89–1.00)
West	0.89 (0.86–0.92)†	0.62 (0.60–0.64)†	0.74 (0.72–0.77)†	0.96 (0.90–1.02)
Hospital location/teaching				
Rural	Ref.	Ref.	Ref.	Ref.
Urban nonteaching	2.43 (2.36–2.50)†	1.30 (1.27–1.34)†	0.97 (0.94–1.00)	1.20 (1.11–1.29)†
Urban teaching	2.19 (2.12–2.25)†	1.22 (1.19–1.26)†	0.87 (0.85–0.90)†	1.25 (1.17–1.35)†
Hospital bed size				
Small	Ref.	Ref.	Ref.	Ref.
Medium	1.35 (1.31–1.39)†	1.14 (1.11–1.18)†	0.98 (0.95–1.01)	1.12 (1.05–1.20)†
Large	1.90 (1.86–1.95)†	1.34 (1.30–1.37)†	0.90 (0.88–0.93)†	1.29 (1.22–1.37)†

* Values are the adjusted odds ratio (95% confidence interval). OI = opportunistic infection; Ref. = reference; SSTI = skin and soft tissue infection; UTI = urinary tract infection.
† Significant.

5.3%), and was stable for UTI (0.6%). These novel estimates of the disease burden of serious infections can inform patients hospitalized with gout and their providers of the disease prognosis.

Compared to sepsis, multivariable-adjusted analyses showed significantly lower odds of discharge to a care facility and in-hospital mortality in patients hospitalized with OI but no significant difference in the odds of higher hospital charges or a longer hospital stay. Compared to sepsis, the odds of health care utilization and in-hospital mortality were also lower for UTI, SSTI, and pneumonia. We identified sepsis to be associated with the worst health care utilization and in-hospital mortality outcomes compared to all other types of serious infections in patients with gout.

We found that older age, Medicaid insurance, higher Charlson–Deyo comorbidity index score, Black race, and Northeast and nonrural hospital location were each independently associated with higher health care utilization and higher in-hospital mortality in

patients with gout hospitalized with infections. Female sex, lower income, and Medicare insurance payer type were associated with higher health care utilization. These patient characteristics indicate frailty, poor socioeconomic status and access to health care, and/or more severe hospitalizations of patients with serious infections, which are all poor prognostic factors for serious infection outcomes (37–39). While some of the associations that we noted are intuitive (older age, lower income, higher comorbidity), other associations (female sex, Black race, Medicare insurance, hospital characteristics) are new. Our study results can help increase awareness of outcomes and prognosis for at-risk populations with these characteristics who are admitted to hospitals with serious infections. If our findings are replicated, these findings might allow the development of a prognostic score for predicting outcomes in patients with gout hospitalized with these serious infections.

Our findings must be interpreted while considering the study's limitations and strengths. Our study is at risk of misclassification bias because we used the ICD-9-CM codes to identify patients with gout. While several studies have supported high positive predictive values of these codes (17–19), others have reported a lower accuracy (20,21). Validation within the NIS is not possible given the lack of medical records. The NIS does not include data from military or Veterans Affairs hospitals, thus impacting the generalizability of findings to patients admitted to these hospitals in the US. The NIS does not have data on medications, laboratory tests, or disease severity, which limits the examination of these important variables. Study strengths include the use of national data, the inclusion of several potential confounders, and the occurrence of sufficient events for the condition(s) of interest.

In conclusion, we examined the epidemiology and time trends of 5 common types of serious infections in patients hospitalized with gout in a national US study. We analyzed health care utilization and mortality associated with common serious infections in patients hospitalized with gout. A better understanding of the predictors of these outcomes will not only allow a better understanding of the prognosis but can also allow the development of individual-level or systems-level interventions to improve outcomes of serious infections in patients hospitalized with gout.

AUTHOR CONTRIBUTIONS

Both authors were involved in drafting the article or revising it critically for important intellectual content, and both authors approved the final version to be submitted for publication. Mr. Cleveland 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. Singh.
Acquisition of data. Cleveland.
Analysis and interpretation of data. Singh, Cleveland.

REFERENCES

1. Singh JA, Sarkin A, Shieh M, Khanna D, Terkeltaub R, Lee SJ, et al. Health care utilization in patients with gout. Semin Arthritis Rheum 2011;40:501–11.
2. Garg R, Sayles HR, Yu F, Michaud K, Singh J, Saag KG, et al. Gout-related health care utilization in US emergency departments, 2006 through 2008. Arthritis Care Res (Hoboken) 2013;65:571–7.
3. Singh JA, Yu S. Time trends, predictors, and outcome of emergency department use for gout: a nationwide US study. J Rheumatol 2016;43:1581–8.
4. Singh JA, Yu S. Gout-related inpatient utilization: a study of predictors of outcomes and time trends. Arthritis Res Ther 2016;18:57.
5. Singh JA, Bharat A, Khanna D, Aquino-Beaton C, Persselin JE, Duffy E, et al. Health care utilization in patients with gout: a prospective multicenter cohort study. BMC Musculoskelet Disord 2017;18:233.
6. Saseen JJ, Agashivala N, Allen RR, Ghushchyan V, Yadao AM, Nair KV. Comparison of patient characteristics and gout-related health-care resource utilization and costs in patients with frequent versus infrequent gouty arthritis attacks. Rheumatology (Oxford) 2012;51:2004–12.
7. Park H, Rascati KL, Prasla K, McBayne T. Evaluation of health care costs and utilization patterns for patients with gout. Clin Ther 2012;34:640–52.
8. Lee G, Roberts L. Healthcare burden of in-hospital gout. Intern Med J 2012;42:1261–3.
9. Lim SY, Lu N, Oza A, Fisher M, Rai SK, Menendez ME, et al. Trends in gout and rheumatoid arthritis hospitalizations in the United States, 1993–2011. JAMA 2016;315:2345–7.
10. Zhu Y, Pandya BJ, Choi HK. Comorbidities of gout and hyperuricemia in the US general population: NHANES 2007–2008. Am J Med 2012;125:679–87e1.
11. Sattui SE, Singh JA, Gaffo AL. Comorbidities in patients with crystal diseases and hyperuricemia. Rheum Dis Clin North Am 2014;40:251–78.
12. Elfishawi MM, Zleik N, Kvrgic Z, Michet CJ Jr, Crowson CS, Matteson EL, et al. The rising incidence of gout and the increasing burden of comorbidities: a population-based study over 20 years. J Rheumatol 2018;45:574–9.
13. Singh JA, Strand V. Gout is associated with more comorbidities, poorer health-related quality of life and higher healthcare utilisation in US veterans. Ann Rheum Dis 2008;67:1310–6.
14. Robinson PC, Merriman TR, Herbison P, Highton J. Hospital admissions associated with gout and their comorbidities in New Zealand and England 1999–2009. Rheumatology (Oxford) 2013;52:118–26.
15. Robinson PC, Kempe S, Tebbutt I, Roberts L. Epidemiology of inpatient gout in Australia and New Zealand: temporal trends, comorbidities and gout flare site. Int J Rheum Dis 2017;20:779–84.
16. Spaetgens B, de Vries F, Driessen JH, Leufkens HG, Souverein PC, Boonen A, et al. Risk of infections in patients with gout: a population-based cohort study. Sci Rep 2017;7:1429.
17. Singh JA, Hodges JS, Toscano JP, Asch SM. Quality of care for gout in the US needs improvement. Arthritis Rheum 2007;57:822–9.
18. Dehlin M, Stasinopoulou K, Jacobsson L. Validity of gout diagnosis in Swedish primary and secondary care: a validation study. BMC Musculoskelet Disord 2015;16:149.
19. Singh JA. Veterans Affairs databases are accurate for gout-related health care utilization: a validation study. Arthritis Res Ther 2013;15:R224.
20. Harrold LR, Saag KG, Yood RA, Mikuls TR, Andrade SE, Fouayzi H, et al. Validity of gout diagnoses in administrative data. Arthritis Rheum 2007;57:103–8.
21. Malik A, Dinnella JE, Kwoh CK, Schumacher HR. Poor validation of medical record ICD-9 diagnoses of gout in a veterans affairs database. J Rheumatol 2009;36:1283–6.
22. Schneeweiss S, Robicsek A, Scranton R, Zuckerman D, Solomon DH. Veteran's affairs hospital discharge databases coded serious bacterial infections accurately. J Clin Epidemiol 2007;60:397–409.
23. Grijalva CG, Chung CP, Stein CM, Gideon PS, Dyer SM, Mitchel EF Jr, et al. Computerized definitions showed high positive predictive

values for identifying hospitalizations for congestive heart failure and selected infections in Medicaid enrollees with rheumatoid arthritis. Pharmacoepidemiol Drug Saf 2008;17:890–5.

24. Patkar NM, Curtis JR, Teng GG, Allison JJ, Saag M, Martin C, et al. Administrative codes combined with medical records based criteria accurately identified bacterial infections among rheumatoid arthritis patients. J Clin Epidemiol 2009;62:321–7, 7e1–7.
25. Jinno S, Lu N, Jafarzadeh SR, Dubreuil M. Trends in hospitalizations for serious infections in patients with rheumatoid arthritis in the US Between 1993 and 2013. Arthritis Care Res (Hoboken) 2018;70:652–8.
26. Tektonidou MG, Wang Z, Dasgupta A, Ward MM. Burden of serious infections in adults with systemic lupus erythematosus: a national population-based study, 1996–2011. Arthritis Care Res (Hoboken) 2015;67:1078–85.
27. Admon LK, Bart G, Kozhimannil KB, Richardson CR, Dalton VK, Winkelman TN. Amphetamine- and opioid-affected births: incidence, outcomes, and costs, United States, 2004–2015. Am J Public Health 2018:e1–7.
28. Jain NB, Ayers GD, Peterson EN, Harris MB, Morse L, O'Connor KC, et al. Traumatic spinal cord injury in the United States, 1993–2012. JAMA 2015;313:2236–43.
29. Blecker S, Paul M, Taksler G, Ogedegbe G, Katz S. Heart failure-associated hospitalizations in the United States. J Am Coll Cardiol 2013;61:1259–67.
30. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP). NIS description of data elements. ZIPINC_QRTL - median household income for patient's ZIP code (based on current year). September 2008. URL: www.hcup-us.ahrq.gov/db/vars/zipinc_qrtl/nisnote.jsp.
31. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 1992;45:613–9.
32. Wysocki JD, Srivastav S, Winstead NS. A nationwide analysis of risk factors for mortality and time to endoscopy in upper gastrointestinal haemorrhage. Aliment Pharmacol Ther 2012;36:30–6.
33. Navaneethan U, Parasa S, Venkatesh PG, Ganapathi TT, Kiran RP, Shen B. Impact of inflammatory bowel disease on post-cholecystectomy complications and hospitalization costs: a Nationwide Inpatient Sample study. J Crohns Colitis 2013;7:e164–70.
34. Sundaram V, Jalan R, Ahn JC, Charlton MR, Goldberg DS, Karvellas CJ, et al. Class III obesity is a risk factor for the development of acute-on-chronic liver failure in patients with decompensated cirrhosis. J Hepatol 2018;69:617–25.
35. Centers for Medicare and Medicaid Services. Medicare eligibility: sho may enroll in Medicare. 2017. URL: https://www.ehealthmedicare.com/about-medicare/eligibility/.
36. Sabesan VJ, Petersen-Fitts G, Lombardo D, Briggs D, Whaley J. Medicaid payer status is linked to increased rates of complications after treatment of proximal humerus fractures. J Shoulder Elbow Surg 2017;26:948–53.
37. Hogg RS, Strathdee SA, Craib KJ, O'Shaughnessy MV, Montaner JS, Schechter MT. Lower socioeconomic status and shorter survival following HIV infection. Lancet 1994;344:1120–4.
38. Xu HF, White RS, Sastow DL, Andreae MH, Gaber-Baylis LK, Turnbull ZA. Medicaid insurance as primary payer predicts increased mortality after total hip replacement in the state inpatient databases of California, Florida and New York. J Clin Anesth 2017;43:24–32.
39. Manoso MW, Cizik AM, Bransford RJ, Bellabarba C, Chapman J, Lee MJ. Medicaid status is associated with higher surgical site infection rates after spine surgery. Spine (Phila Pa 1976) 2014;39:1707–13.

Arthritis Care & Research
Vol. 73, No. 6, June 2021, pp 909–917
DOI 10.1002/acr.24489

American College of Rheumatology
Empowering Rheumatology Professionals

Modifications in Systemic Rheumatic Disease Medications: Patients' Perspectives During the Height of the COVID-19 Pandemic in New York City

Carol A. Mancuso,[1] Roland Duculan,[2] Deanna Jannat-Khah,[1] Medha Barbhaiya,[1] Anne R. Bass,[1] Lisa A. Mandl,[1] and Bella Mehta[1]

Objective. Concerns about severe acute respiratory syndrome coronavirus 2 (SARS–CoV-2) infection may have led to changes or discontinuation of immunosuppressive medications among patients with systemic rheumatic disease. Our goal was to assess patients' perspectives regarding medication modifications and deviations from planned uses during the height of the pandemic.

Methods. Adult patients of 13 rheumatologists at an academic center with physician-diagnosed rheumatic disease and prescribed disease-modifying medications were interviewed by telephone and asked open-ended questions about the impact of SARS–CoV-2 on their medications. Responses were analyzed using content and thematic analyses to generate categories that described patterns of medication modification.

Results. A total of 112 patients (mean age 50 years, 86% women, 34% non-White race or Latino ethnicity) with diverse diagnoses (30% lupus, 26% rheumatoid arthritis, 44% other) who were taking various medications were enrolled. Patients reported clinically relevant issues that were iteratively reviewed to generate unique categories of medication modification: medications and increased or decreased risk of SARS–CoV-2 infection; role of hydroxychloroquine; maintaining medication status quo; role of glucocorticoids; increasing or decreasing existing medications in relation to clinical disease activity; postponing infusions; and medication plan if infected by SARS–CoV-2. Some modifications were suboptimal for disease control but were made to mitigate infection risk and to minimize potential harm when patients were unable to obtain laboratory tests and physical examinations due to cessation of in-person office visits.

Conclusion. During the height of the pandemic, substantial medication modifications were made that, in some cases, were temporizing measures and deviations from planned regimens. Future studies will assess short- and long-term sequelae of these medication modifications.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS–CoV-2) has profound infectious and inflammatory effects and is a particular threat to patients with preexisting comorbidity (1). In general, patients with systemic rheumatic disease (i.e., autoimmune, chronic, inflammatory conditions) are especially at increased risk of serious infection and worse clinical outcomes due to underlying immune dysfunction, superimposed immunosuppressive medications, and disease-related comorbidity (2,3).

Multiple classes of medications exist for rheumatic disorders with both general and highly specific effects, including anticytokine effects (4–6). While stopping immunosuppressive medications may decrease infection risks, given the "cytokine storm" associated with SARS–CoV-2, some immunomodulatory agents may mitigate certain inflammatory components of the virally induced coronavirus disease 2019 (COVID-19) syndrome (6).

Despite possible infection, standard rheumatologic guidelines recommend maintaining an established medication regimen when rheumatic conditions are stable (2,7). Challenges arise, however, when flares develop during an at-risk period, such as during the COVID-19 pandemic. Such periodic fluctuations in symptoms may be common under the best of circumstances but are more prevalent and severe during times of increased physical and psychological stress (8). Although potentially mitigated by telehealth visits for those with access, social distancing and

[1]Carol A. Mancuso, MD, Deanna Jannat-Khah, DrPH, MSPH, Medha Barbhaiya, MD, MPH, Anne R. Bass, MD, Lisa A. Mandl, MD, MPH, Bella Mehta, MBBS, MS: Hospital for Special Surgery and Weill Cornell Medical College, New York, New York; [2]Roland Duculan, MD: Hospital for Special Surgery, New York, New York.

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

Address correspondence to Carol A. Mancuso, MD, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021. Email: mancusoc@hss.edu.

Submitted for publication June 1, 2020; accepted in revised form October 13, 2020.

SIGNIFICANCE & INNOVATIONS

- The coronavirus disease 2019 pandemic influenced modifications in medications for systemic rheumatic diseases with potential short- and long-term sequelae for disease control.
- Modifications were not systematic but instead, in some cases, resulted in more or additional medications, particularly glucocorticoids, and in other cases resulted in fewer medications, such as delays in infusions and in response to flares.
- Guidelines for regulating medications for systemic rheumatic diseases during times of uncertainty should address immunosuppressive and immunomodulatory effects of existing regimens, response to increased or decreased rheumatic symptoms, and alternative methods to monitor disease activity in the absence of traditional physical examination and laboratory tests.

suspension of in-person office visits contribute to challenges as physicians are unable to gain information from physical examinations and laboratory tests to inform their decision-making.

To date, there is limited information about the impact of COVID-19 on patients' experiences with their medications for systemic rheumatic disease. Preliminary reports indicate that altering medications may not be uncommon (2,9). The effects of decreased or suboptimal medication dosing in other scenarios, such as intentional cessation and nonadherence, generally have shown unfavorable outcomes (10). We hypothesized that patients had medication modifications due to COVID-19 that deviated from established regimens.

Our goal was to obtain detailed information about patients' experiences with their medications during the COVID-19 pandemic. We sought this information in real-time during the height of the pandemic in New York City.

PATIENTS AND METHODS

Recruitment. This study was approved by the institutional review board at Hospital for Special Surgery (HSS), and all patients provided verbal consent. Enrollment started on April 2, 2020, when incidence and death rates from COVID-19 in New York City showed continuous increase, and ended on April 21, 2020, when these rates showed several consecutive days of decrease (11) (Figure 1). Patients were recruited from 13 rheumatology practices at HSS, selected because they had high volumes of patients with diverse diagnoses, serve patients of different socioeconomic statuses, and participate in clinical research. Patients were eligible if they were English speaking and were taking at least 1 immunosuppressive medication, including glucocorticoids and disease-modifying antirheumatic drugs (DMARDs) (12). All patients had a rheumatologist-diagnosed systemic rheumatic disease. Patients were identified by direct referral from their rheumatologists as having had a recent visit or communication with their rheumatologist. Patients also were identified by reviewing daily telehealth appointment schedules during the past week; patients were then recruited after obtaining approval from their rheumatologists.

Data collection. This study used interpretive qualitative methods and grounded theory, which are appropriate when explanations of new phenomena are sought. Ground theory relies on acquisition of actual data from which new theories are generated (13–16).

Patients were contacted by telephone and, if they agreed to participate, they were either interviewed at that time or at an alternative preferred date. Patients were asked the following open-ended questions: "What do you know about medications for rheumatic conditions and COVID-19? If you altered your medications because of COVID-19, why, in what ways? How do you think this might benefit/harm you? What did you and your rheumatologist talk about regarding COVID-19 and your medications?" These questions were modeled after a qualitative study of RA patients' perspectives about tapering medications (17).

Two nonrheumatologist physicians experienced in qualitative data collection and analysis (CAM and RD) interviewed patients. One investigator (CAM) conducted the interview, and both investigators wrote down responses in independent field notes. The second investigator's (RD) notes were detailed with verbatim key responses that in some instances were phrases but most often were short sentences. As needed, patients were

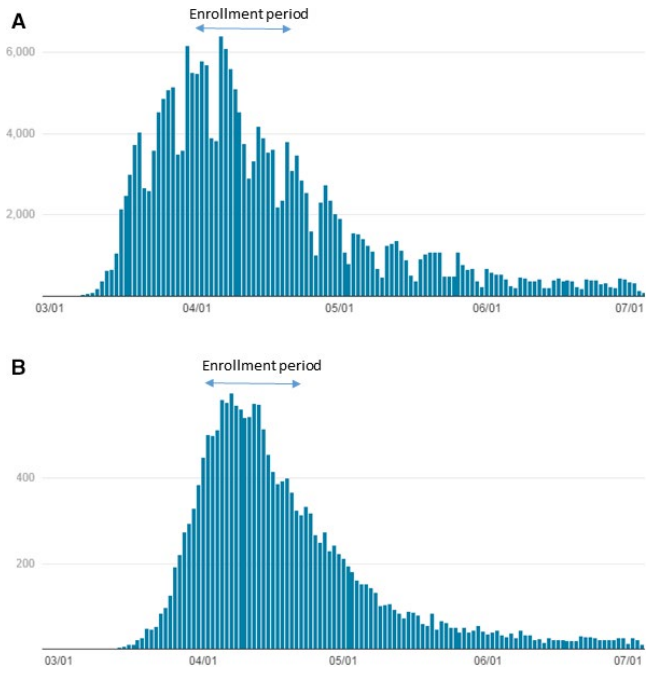

Figure 1. Number of cases (**A**) and deaths (**B**) due to coronavirus 2019 in New York City in 2020 according to date.

asked to repeat their responses for clarification and emphasis, and responses were not paraphrased. Responses also were repeated back to patients for content validation and to ensure accuracy in recording the patients' own words (14–16). Patients were aware that there were 2 physicians participating. All interviews addressed the same questions, and patients were encouraged to cite personal experiences to support their perspectives. Immediately at the conclusion of each interview, the investigators conferred to ensure comprehensive notation of responses and to create a single composite account of the conversation. As the interviews progressed, incoming responses were similar to previous responses, and the collective set of responses comprehensively addressed the research questions (18). Electronic medical records were reviewed to collect information about demographic characteristics, currently prescribed medications, clinical history, and any laboratory-confirmed diagnosis of SARS–CoV-2.

Qualitative analyses. The investigators' field notes were transcribed and, using open coding, were reviewed line-by-line to identify unique concepts based on an inductive approach (13,14). Concepts were then aggregated into categories according to the common phenomena they represented. Based on constant repeated review and a comparative analytic strategy, categories were iteratively examined and refined to ensure that they addressed unique features (13,19,20). Categories were then named to describe the main topics that they represented. The investigators who participated in the telephone calls did the initial coding and assembled the categories. Two other investigators (DJK and BM), one a rheumatologist, and the other a biostatistician with qualitative experience, subsequently and independently reviewed the transcripts and confirmed that they agreed with (i.e., corroborated) the categories (13,21).

RESULTS

In total, 112 patients participated and completed the interview. They were enrolled as follows: 117 were contacted by telephone; 105 participated in the interview at the time of the initial telephone call. Nine patients requested an arranged time for the interview; of these, 7 were interviewed within 3 days, and 2 could not be interviewed during the enrollment period. Three women refused to participate due to an inconvenient time or unwillingness to discuss personal situations.

Most participants were women (86%), the mean ± SD age was 50 ± 15 years, and 34% were self-described as non-White or Latino (i.e., 19% non-White race, another 15% Latino ethnicity) (Table 1). Patients had various diagnoses (i.e., 30% systemic lupus erythematosus, 26% RA, 44% other) for a mean ± SD duration of 11 ± 10 years; range 6 months to 57 years). All patients were taking at least 1 medication, and 63% were taking >1 medication (Table 1). Two patients had a confirmed positive diagnosis of SARS–CoV-2 within the preceding 4 weeks; one had been hospitalized, and the other was treated as an outpatient.

Table 1. Demographic and clinical characteristics (n = 112)*

Characteristics	Value
Women	96 (86)
Age, median (interquartile range) years	49 (36–60)
Race	
White	91 (81)
Black	12 (11)
Asian	9 (8)
Latino ethnicity	17 (15)
Diagnosis	
SLE	34 (30)
RA	30 (26)
UCTD	8 (7)
Psoriatic arthritis	8 (7)
Sjögren's syndrome	4 (3)
Mixed connective tissue disease	3 (3)
Spondyloarthritis	2 (2)
SLE–UCTD overlap	2 (2)
Sjögren's syndrome–RA overlap	2 (2)
Polymyalgia rheumatica	2 (2)
Granulomatosis with polyangiitis	2 (2)
Antiphospholipid syndrome–SLE overlap	2 (2)
Ankylosing spondylitis	2 (2)
Other†	11 (10)
Duration of disease, median (interquartile range) years	8 (4–14)
Medications	
Small molecules	
Hydroxychloroquine	58 (52)
Glucocorticoids	58 (52)
Methotrexate	17 (15)
Azathioprine	10 (9)
Mycophenolate mofetil	7 (6)
Biologics	
Tumor necrosis factor inhibitor	19 (17)
Adalimumab	8 (7)
Certolizumab	3 (3)
Etanercept	4 (3)
Infliximab	4 (3)
IL-1 inhibitor (canakinumab)	1 (1)
IL-5 inhibitor (mepolizumab)	2 (2)
IL-6 inhibitor (tocilizumab)	6 (5)
IL-17 inhibitor (secukinumab)	2 (2)
IL-12/23 inhibitor (ustekinumab)	1 (1)
B cell activating factor inhibitor (belimumab)	13 (12)
T cell costimulatory signal inhibitor (abatacept)	4 (3)
Anti–CTLA-4 monoclonal antibody (ipilimumab)	1 (1)
B cell CD20 monoclonal antibody (rituximab)	11 (10
JAK inhibitor (tofacitinib)	3 (3)

* Values are the number (%) unless indicated otherwise. IL = interleukin; RA = rheumatoid arthritis; SLE = systemic lupus erythematosus; UCTD = undifferentiated connective tissue disease.
† Small vessel vasculitis, SLE–Sjögren's syndrome overlap, scleroderma, SLE–scleroderma overlap, SLE–RA overlap, RA–polymyalgia rheumatica overlap, inflammatory polyarthralgia, eosinophilic granulomatosis with polyangiitis, atypical polyarthritis nodosa, adult-onset Still's disease.

Based on qualitative analyses, the following categories were identified regarding potential medication risks and alterations in medication regimens. Generic names have been substituted for brand names in quotations.

Table 2. Qualitative categories and quotations addressing the role of medications and risks of infection*

Categories	Patients' quotations
Rheumatic disease medications and risk of SARS–CoV-2	"I am not sure if the lupus or the medications makes me more vulnerable. Belimumab and mycophenolate definitely make me more at risk. But it is confusing. Am I more at risk of getting the virus or more at risk of what happens if I get it?" (woman with lupus, age 35 years). "I have interstitial lung disease, I am older, and I am on immunosuppressant medications methotrexate and prednisone. I am in big trouble if I get it. Medicines definitely make me more at risk" (man with Sjögren's syndrome, age 67 years). "I am potentially immunocompromised due to medications and therefore more vulnerable" (woman with RA, age 33 years). "I am on adalimumab; I am at higher risk, meaning it would take me longer to recover" (woman with atypical polyarthritis nodosa, age 30 years). "We take drugs that compromise the immune system. We may not be able to fight as others do if we get the illness" (woman with RA, age 52 years). "I stopped going to Christmas parties and didn't get the flu anymore; so if I avoid people I don't get sick. Or maybe the medications protect me? I took etanercept in the past and now certolizumab for 8–10 years. I don't know if the medications make more risk or maybe they make less risk" (man with RA, age 64 years).
Role of hydroxychloroquine	"I have taken hydroxychloroquine for 15 years, it works wonders for me. I have no flare ups. I have more or less a normal life. I have no organ problems. I do not want to change it. The pharmacy didn't have it and put me on a waiting list, but then I got it because lupus has priority" (woman with lupus, age 74 years). "I have no difficulty getting hydroxychloroquine. The pharmacist knows me well for 25 years. I got a 3-month supply. He said never worry. Five other patients got it too" (woman with RA, age 58 years). "It was difficult to get hydroxychloroquine; the pharmacy would not give me a 3-month supply. I know they are trying to use if off label for COVID and that is making it hard for patients with rheumatic conditions to get it" (woman with UCTD, age 52 years). "My doctor said if I had trouble getting hydroxychloroquine I could take half the dose because I have built up enough protection in my system. But I am concerned about the potency of the hydroxychloroquine and its ability to prevent flares at lower doses" (man with inflammatory polyarthralgia, age 50 years). "I have a prescription for 2 pills a day, but my doctor told me that I should be taking only once a day, so I am able to keep a reserve. I also got some on online so I now have over a year's supply" (woman with lupus, age 22 years). "Hydroxychloroquine has been suggested as a possible modality for the virus, Vitamin C and antibiotics too. But I don't know how this would work; nobody knows. We have to sit tight until scientists and doctors come up with a treatment, a vaccine, and understand transmission" (woman with RA, age 53 years).
Maintain medication status quo	"I spoke briefly with my doctor before the real epidemic started. He told to be careful, do the same thing as I was doing with my medications, make no changes to my medications. He said my medications don't increase my risk, but my preexisting condition does" (woman with RA/lupus, age 45 years). "My doctor is afraid if I don't get the rituximab infusion I will get a flare and thus be back-pedaling to get where I should be. I postponed it a couple of weeks and got it the end of March. Then I got another infusion 2 days ago for April" (woman with RA/Sjögren's syndrome, age 39 years). "Early on I discussed this with my doctor. He said for the time being it is more dangerous to alter medications, so stay on them" (woman with RA, age 33 years).
Role of glucocorticoids	"My mother passed away a couple of weeks ago. Then my lupus went berserk and I needed more prednisone. My doctor told me to continue at this higher dose for now and then taper it" (woman with lupus, age 54 years). "I take adalimumab and celecoxib and have a flare now. My RA is active, I have pain and swelling in my joints. My doctor started me on prednisone. I also take ibuprofen 800 mg 4 times a day. I am better but not there yet" (woman with RA, age 60 years). "I am having a flare now, I cannot walk on my left foot, and my left elbow is swollen. I decided to accept the flare instead of increasing medications. I am not responding to hydroxychloroquine and am not willing to advance to other drugs like methotrexate. My doctor supplemented the hydroxychloroquine with steroids, but I titrated it down now because I was going out of my mind" (woman with RA, age 53 years).
Plan if contracted SARS–CoV-2	"There have been no alterations in my medications unless I get a fever. Then I stop my medications" (woman with systemic sclerosis, age 57 years). "My doctor told me not to alter medications, but if I get sick to tell him right away and then maybe he would lower the medications" (man with small vessel vasculitis, age 31 years). "I had an appointment with my doctor the week before it all happened. She said take precautions and if it feels like I am becoming sick to start to lower the methotrexate" (woman with UCTD, age 44 years). "My doctor told me to stop methotrexate and adalimumab if I get infected, otherwise it is better not to have a flare" (woman with RA, age 33 years). "I emailed my doctor and asked what should I do if I get sick. She said if my temperature was greater than 100 and I got a sore throat, I should increase my hydroxychloroquine and vitamin C for 1 day, then continue my regular dose and call her right away" (woman with lupus, age 34 years).

* Generic names have been substituted for brand names in quotations. RA = rheumatoid arthritis; SARS–CoV-2 = severe acute respiratory syndrome coronavirus 2; UCTD = undifferentiated connective tissue disease.

Rheumatic disease medications and SARS–CoV-2 risk. Patients had different opinions about medications, but many thought that their medications increased the risk of contracting SARS–CoV-2 (Table 2). In some cases, patients considered this risk greater than the risk from their rheumatic condition. They also believed they would have worse outcomes because of their medications if they contracted the virus. "I think the risk from medications is definitely greater than the risk of RA. My medications make a bigger impact" (woman with RA, age 33 years). "We take drugs that compromise the immune system and may not be able to fight as others do if we get the virus" (woman with RA, age 52 years). "If I get the virus it would be bad news, a whole different ball game" (man with psoriatic arthritis, age 42 years).

These concerns, however, were countered by concerns for potential consequences of not taking medications, specifically worsening symptoms and triggering flares. "If I were to stop methotrexate my risk would go down, but which way is better? Balance the risk of a flare with the risk of COVID?" (woman with RA, age 50 years) "Medications are a double-edged sword; damned if you do and damned if you don't" (woman with lupus, age 48 years).

In contrast, some patients believed medications decreased the risk of infection. They believed better control of their condition enhanced their ability to be healthy and withstand infection. "I don't think medications increase my risk, I think I am at less risk because of the medications. I can fight better if I am healthier" (woman with Still's disease, age 27 years).

Role of hydroxychloroquine. Our study occurred during the initial controversy surrounding potential therapeutic effects of hydroxychloroquine for COVID-19 (21). Fifty-two percent of patients were taking hydroxychloroquine and attested to its importance in controlling their condition (Table 2). Of these, 61% reported no difficulty obtaining hydroxychloroquine; the rest, however, needed their rheumatologist's help to obtain the medication, mainly by communicating with pharmacists and insurance companies to verify diagnosis and prescription, or to arrange mail order delivery.

Some patients were markedly anxious about the availability of hydroxychloroquine and sought to ensure that they had a surplus. Some patients also reported that their rheumatologist advised decreasing the dose or switching from daily to alternate-day dosing if their supply was low; however, only 1 patient did so, and the rest reported that they did not miss any doses. Some patients also were surprised that hydroxychloroquine, a medication that they believed potentially weakens the immune system, could be beneficial in fighting infection. Most acknowledged that the role of hydroxychloroquine was unresolved and needed further study. "Initially I was very worried. But now I am not so sure. My nephrologist says hydroxychloroquine could possibly be good because it is a modulator for inflammation" (woman with lupus, age 34 years). "Some say hydroxychloroquine is good for inflammation. But I don't know too much about it. I think it is still experimental" (woman with lupus, age 48 years).

Maintain medication status quo. Most patients reported conversations with their rheumatologists within the past month about medications and susceptibility to the virus. Most reported that they were advised not to alter medications (Table 2). The main reason was to not induce change in the immune system that might affect the ability to resist infection. "I am well now on these medications. No flares. My doctor said stay the course." (man with psoriatic arthritis, age 61 years). "My doctor said the immune system should be where it should be, not over-active and not under-active" (woman with undifferentiated connective tissue disease, age 66 years).

In some cases, maintaining medication regimens was accomplished by self-injection instead of hospital-based infusions when a subcutaneous option existed. "I give myself injections at home now. I learned to do this online with remote learning sessions. My doctor walked me through it. I was a little nervous but he did not want me to go to the office and he did not want me to miss any doses" (woman with eosinophilic granulomatosis with polyangiitis, age 65 years).

Role of glucocorticoids. At the start of the pandemic, some patients were recuperating from a flare and were tapering prednisone. If patients were on multiple medications and were improving, the decision often was made to proceed with the predetermined tapering plan (Table 2). "I take hydroxychloroquine, methotrexate and abatacept; the prednisone is being tapered now. Two days ago I discussed this with my doctor. We decided to continue the taper because I am feeling better" (woman with RA, age 33 years).

For some patients, glucocorticoids increased during the study period, specifically in response to flares. Glucocorticoids were preferentially used rather than adding or modifying a DMARD. Patients commented that these were attempts to temporize and quickly reverse symptoms in lieu of more definitive treatments that require monitoring with laboratory tests. "I spoke to my doctor yesterday, I have puffy fingers. She increased my prednisone from 5 to 7.5 [mg]" (man with RA, age 66 years). "I am having a flare. My doctor would have prescribed a biologic but doesn't want to start a new drug now with the virus. It is hard to get blood work, there isn't the usual lab access. So we are using a band-aid instead, prednisone" (woman with RA, age 47 years).

Decrease existing medications. In some instances, patients reported that rheumatologists decreased medications to diminish SARS–CoV-2 infection risk. Patients believed that this was based on their rheumatologists' knowledge of their prior bouts of infection or attempts to streamline a multi-medication regimen (Table 3). More often, however, when medications were altered because of SARS–CoV-2, it was done by apprehensive patients. After changes in medications, some patients remained stable, but others did not and needed to resume medications. "I curtailed my methotrexate myself. I panicked early on and stopped

Table 3. Qualitative categories and quotations addressing increasing and decreasing medications*

Categories	Patients' quotations
Decrease existing medications	"I was sick in January, went to the doctor twice, I was pretty sick. Then I got another bout that was milder. I am extremely concerned because of my age and my medications. My doctor told me to stop rituximab but to continue the hydroxychloroquine at the maximum dose. I am happy to be off rituximab; the less medicine the better" (woman with RA/Sjögren's syndrome, age 57 years). "My doctor and I decided to come off the etanercept because of the virus. But then my right hand swelled up and I started it again, and we added prednisone too. It is better now" (woman with RA, age 73 years). "I stopped my medications, but not at my doctor's advice. I think if I get the virus my body has more of a chance to fight it if off medications. But I got more symptoms; all the old problems came back. Two days ago I told my doctor about the swelling and achiness. He said to go back on the medications right away or else my body might develop antibodies to the drugs and I might need new medicines, which are stronger" (man with RA, age 70 years). "In the beginning I was extremely concerned. My doctor said if you think it will help, go ahead and stop the tofacitinib, but he didn't want me to. I decided to stop it on March 16 because I was frightened that my immune system was not going to withstand the virus. I stopped for 4 weeks and then started not to feel great. I was feeling a little creaky; I thought maybe because I was not getting enough activity. My doctor said to resume the medicine, so I am back on it as of April 9" (woman with granulomatosis with polyangiitis, age 66 years).
Concern about decreasing medications	"I am worried if I stop medications my symptoms will come back. I had a lot of knee pain, I am afraid that pain would come back" (woman with RA, age 53 years). "I know the role of my medications, I know I need them. I will get sick if I don't have them. I take azathioprine, prednisone and hydroxychloroquine" (woman with lupus, age 41 years). "I am concerned if I am not taking my medications. In 1 week the arthritis would come roaring back with morning stiffness and swelling of my hands and feet. I also have lesions on my skin and scalp, they would come back too" (man with psoriatic arthritis, age 42 years).
No decrease in medications despite clinical improvement	"I had a visit with my doctor 3 weeks ago, she said I could probably decrease my steroids, but due to COVID maybe it is better not to change anything" (woman with mixed connective tissue disease, age 55 years). "I know I am suppressed because of adalimumab. I was in a flare and consulted my doctor when my symptoms were better. But he said not to decrease the adalimumab even though the flare was better; he otherwise would recommend decreasing it. That was because of COVID" (woman with RA/lupus, age 30 years).
No increase in medications despite a flare or lack of improvement	"My doctor wants me to keep the hydroxychloroquine as is. I am in-between. I know I need more medications because my hand is swelling, it is a flare. But we will hold for now" (woman with RA, age 50 years). "I had a flare and was started on belimumab as a result of blood work. The pain increased and the belimumab was stopped and I was started on prednisone 40 mg a day. My doctor did not want to try another biologic due to the virus" (woman with lupus, age 28 years). "I am taking hydroxychloroquine and prednisone, the goal was to move me off prednisone and onto methotrexate. But my doctor does not want 2 standing immunosuppressive drugs at the same time right now, so the start of methotrexate is on hold and we are slowing down the prednisone taper" (woman with UCTD, age 55 years). "When I get a cold I get a systemic response. I get very sick. I had symptoms starting in early March and tested positive for COVID on March 29. I was supposed to start certolizumab, but we decided not to. I still have a cough and maybe my heart and my lungs are still involved" (woman with spondyloarthritis, age 52 years).
Postponing infusions	"My next infusion of rituximab is for June. If I am feeling well my doctor said we will see if I can push the infusion until August" (woman with granulomatosis with polyangiitis, age 66 years). "I have not had an infusion since February 13. I am supposed to do it every 4–5 weeks. So far I feel OK. My doctor said do not come to the hospital for an infusion. If I deteriorate then we will start prednisone. Then we will go back to the infusion plan when I can come back to the hospital" (man with RA/polymyalgia rheumatica, age 77 years). "I had my infusion of infliximab 6 weeks ago. I am due this week, but it has been postponed until May. My doctor said let's see what happens in May. He gave me a prescription for prednisone which I would start if I cannot get the infliximab" (man with psoriatic arthritis, age 71 years).

* Generic names have been substituted for brand names in quotations. RA = rheumatoid arthritis; UCTD = undifferentiated connective tissue disease.

it. I am off it now and the results have been good 6 weeks later. My doctor said so far, so good, and didn't insist I restart it" (man with RA, age 41 years). "I stopped ustekinumab myself because should I be exposed to the virus I do not know what would happen. My reaction could be less, or maybe more. As a result I am in a bad flare with my psoriasis. It is everywhere, big patches even on my scalp, and it is very bad" (woman with undifferentiated connective tissue disease, age 59 years).

Concern about decreasing medications. Most patients were concerned that decreasing medications would result in worse symptoms and flares. Some were also concerned about the availability of medical care if they were to have a flare. "I take so many medications, I am afraid if I stopped them my pain would be worse" (woman with lupus, age 49 years). "When I stop my medications my skin gets tight and my joints hurt" (woman with systemic sclerosis, age 57 years). "I take methotrexate for RA

and would get sick if I didn't. I know methotrexate weakens my immune system and makes me more vulnerable. But I worry about a flare and if I can't get to the hospital" (woman with RA, age 50 years). "My fear is going to the hospital more than getting the virus. It is chaotic there" (woman with lupus, age 76 years).

No decrease in medications despite clinical improvement. In some situations, medications were not decreased when there was clinical improvement and there had been a plan to taper medications based on symptoms. According to patients, this was driven by the rheumatologist's desire to avoid possible flares and subsequent increased susceptibility to infection. "My doctor was lowering my prednisone dose but is holding for now so that I do not get sick" (woman with polymyalgia rheumatica, age 75 years).

In other situations, medications were not decreased due to lack of laboratory tests to support that there was improvement. "I was started on injections every week, then blood tests were fine and I went to every other week. I would normally have a blood test this week and based on that I could go to every 3 weeks. But I can't come for the test. My doctor said stay on the current regimen because we can't tell how the inflammation is now" (woman with polymyalgia rheumatica, age 74 years).

No increase or change in medications despite a flare or lack of improvement. In some cases, medications were not increased when patients and rheumatologists had previously discussed a plan and a timeline to increase doses if there was only a partial response and the medication was well tolerated (Table 3). "My doctor was supposed to start methotrexate but didn't though it was indicated because of the virus" (woman with undifferentiated connective tissue disease, age 66 years).

There also were plans to switch to other medications if there was no improvement and no clinical or laboratory test side effects. These changes often were not made because of the absence of physical examination and laboratory tests. "I had a bad flare for 2 days, there were no changes to my medications. My doctor wants blood work first before changing medications" (woman with lupus, age 51 years).

As above, some patients had flares and would otherwise have increased doses of medications. "I currently have a flare; my elbows, knees and toes are swollen. I have had RA for more than 18 years; I know a flare. My doctor said he would switch my medications from etanercept to adalimumab, but he has not done it yet. He otherwise would have" (woman with RA, age 52 years). "I am not on hydroxychloroquine now, but was supposed to start it. I am on prednisone. My [general practitioner] said to get off the prednisone immediately because it decreases my immunity and gives me hallucinations. I am having hallucinations now" (woman with Sjögren's syndrome, age 45 years).

Postponing infusions. Obtaining periodic medication infusions also posed challenges as risks of missing doses were weighed against risks of being exposed to clinical settings. In general, patients reported that rheumatologists encouraged continuing infusions if possible. "I did not get my infusion on March 30. My doctor told me to stay away. I already see a difference; I have less energy. I am due for the next infusion on April 27, I hope I can get it" (woman with RA, age 74 years).

Some patients, however, preferred to avoid infusions even if it meant worse symptoms. "I missed my infusion on March 17 and am planning to miss my infusion on April 22. I have a flare now, body pain all over, my arm and ankle are swollen, my legs are stiff, and walking is difficult. I am taking celecoxib; it bothers my stomach so I am taking famotidine. I didn't speak to my doctor about this flare" (woman with lupus, age 37 years).

Plan if contracted SARS–CoV-2. All patients reported that their rheumatologists emphasized safe practices, such as staying home and monitoring symptoms. Some patients reported that their rheumatologists also proposed a plan for what they should do with medications if viral symptoms developed (Table 2). "My doctor told me don't change anything during COVID. But if I get the virus I should leave the prednisone as is and go off the mycophenolate" (woman with eosinophilic granulomatosis with polyangiitis, age 65 years).

Instructions were more detailed for patients who were at particular risk of infection, for example, those who had direct exposure to someone who had the virus. "My doctor said the moment I develop any symptoms whatsoever, I should stop belimumab completely, increase hydroxychloroquine, and call him on his cellphone immediately" (woman with lupus, age 49 years).

DISCUSSION

Our qualitative study showed that patients and their rheumatologists faced challenges regulating medications for systemic rheumatic diseases during the height of the COVID-19 pandemic in New York City. Based on interviews with patients, we discerned that patients were aware of the increased risks simultaneously posed by their underlying immunocompromised state and medications used for treatment. They also identified challenges posed by the virus, including switching medications, periodic infusions, and managing flares.

The major finding of our study is that the COVID-19 pandemic influenced modifications in systemic rheumatic disease medications with potential short- and long-term sequelae for disease control. Modifications were not systematic but instead, in some cases, resulted in more or additional medications, particularly glucocorticoids, and in other cases resulted in fewer medications, such as delays in infusions and in responses to flares. The implications of our study are that guidelines for regulating medications for

systemic rheumatic diseases during times of uncertainty should address immunosuppressive and immunomodulatory effects of existing regimens, response to increased or decreased rheumatic symptoms, and alternative methods to monitor disease activity in the absence of traditional physical examinations and laboratory tests. (7).

Most patients in our study were aware of potential consequences of deviating from usual medications and dosing regimens. We found that the main deviation was that some medications were not dosed as they otherwise would have been. This was done both to minimize alterations to the immune system that might impede response to the virus and because of the inability to monitor physical examinations and laboratory tests. In some cases, medications were not decreased as planned when there was clinical improvement; in this scenario, patients were exposed to more medications than necessary. In other cases, medications were not increased or initiated to replace ineffective ones; in this scenario, patients were exposed to medications that were no longer indicated while their disease remained active.

Although our study took place during a brief period, some patients already reported worsening symptoms and flares that they attributed to suboptimal medications and to emotional and physical upheaval caused by the pandemic (22). Glucocorticoids were often used in response because they reversed symptoms quickly and were familiar to patients. This strategy, however, introduced new side effects and other management decisions associated with tapering.

The focus on hydroxychloroquine during the study period resulted in decreased supplies in some local pharmacies and increased concern among patients about future availability (23). Some rheumatologists had to intervene to ensure that their patients had the requisite supply. Some patients were curious to learn how the virus could invoke an inflammatory response that was similar to their rheumatic disease and thus potentially was treatable with similar medications.

With respect to prior studies, most reports to date are from physicians' perspectives (6,24), with limited information about patients' perspectives regarding rheumatic disease medications and COVID-19. One registry-based study addressing patients' point of view was conducted at the end of March 2020 and emailed patients across the US to ask about viral symptoms, medications, and health care access during the previous 2 weeks (9). In this registry study, ~11–14% of respondents reported self-imposed or physician-directed changes to medications, and these were attributed to concerns about the virus or recent viral symptoms. No information about alternative medication regimens, flare management, and suboptimal increases or decreases in medications was obtained.

Most information about elective medication alteration comes from reports of decreasing or stopping therapies under the guidance of physicians and in response to patients' requests for lower doses and drug holidays (25). Current guidelines support tapering treatment after remission, but the optimal approach is not known (26,27). Physicians' concerns include managing relapses, recapturing control with retreatment, and monitoring laboratory test progression (27,28). In a qualitative study assessing patients' perspectives about decreasing medications, major concerns included recurrence of symptoms, disease progression, quality of life, and prompt access to health care for flares (17).

There are several limitations to our study. First, our participants were patients at 1 tertiary care center in New York City, and their perspectives may differ from patients in other settings where the pandemic was less prevalent. Also, although rheumatic diseases generally are more common in women, men were underrepresented in our study. Second, physicians were all specialty-trained academic rheumatologists whose management practices may differ from those of other health care providers. Third, in order to capture current perspectives and practices, we sampled from patients who recently were in contact with their rheumatologists. This may represent a cohort with greater flux in their condition or greater likelihood of seeking and accepting alterations in medications. Fourth, we did not audiotape interviews in order to streamline enrollment and maximize participation during a challenging recruitment period. As such, some nuances and details may have been missed in field notes.

Our study provides a comprehensive view of patients' attitudes and medication practices during the height of the COVID-19 pandemic in New York City. Notable strengths of this study are that patients were enrolled from different rheumatology practices, had diverse diagnoses, and were taking a variety of rheumatic disease medications. Additionally, our study was conducted in real-time and therefore was not subject to recall error. It also was conducted within a narrow timeframe during the active period of the pandemic and thus was less subject to shifting perspectives based on subsequent information. We learned that during this most uncertain time, patients and their rheumatologists made substantial modifications to essential medications that, in some cases, were temporizing measures and deviations from planned regimens. These modifications occurred to minimize both infection risk and potential harm from not being able to monitor clinical status with physical examinations and laboratory tests. The near- and long-term sequelae of these modifications and deviations will be assessed in longitudinal follow-ups.

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. Mancuso 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. Mancuso, Duculan, Jannat-Khah, Barbhaiya, Bass, Mandl, Mehta.

Acquisition of data. Mancuso, Duculan.

Analysis and interpretation of data. Mancuso, Duculan, Jannat-Khah, Barbhaiya, Bass, Mandl, Mehta.

REFERENCES

1. Zunyou W, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72,314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020;323:1239–42.
2. Venerito V, Lopalco G, Iannone R. COVID-19, rheumatic diseases and immunosuppressive drugs: an appeal for medication adherence. Rheumatol Int 2020;40:827–8.
3. Arango M, Shoenfeld Y, Cervera R, Anaya JM, et al. Infection and autoimmune diseases. In: Anaya JM, Shoenfeld Y, Rojas-Villarraga A, et al, editors. Autoimmunity: from bench to bedside. Bogota: El Rosario University Press; 2013. Chapter 19. URL: https://www.ncbi.nlm.nih.gov/books/NBK459437/.
4. Ferro F, Elefante E, Baldini C, Bartoloni E, Puxeddu I, Talarico R, et al. COVID-19: the new challenge of rheumatologists. Clin Exp Rheumatol 2020;38:175–80.
5. Cron RQ, Chatham WW. The rheumatologist's role in COVID-19. J Rheumatol 2020;47:639–42.
6. Haberman R, Axelrad J, Chen A, Castillo R, Yan D, Izmirly P, et al. Covid-19 in immune-mediated inflammatory diseases: case series from New York. N Engl J Med 2020;383:85–8.
7. Mikuls TR, Johnson SR, Fraenkel L, Arasaratnam RJ, Baden LR, Bermas BL, et al. American College of Rheumatology guidance for the management of rheumatic disease in adult patients during the COVID-19 pandemic. Version 1. Arthritis Rheumatol 2020;72:1241–51.
8. De Brouwer SJ, Kraaimaat FW, Sweep FC, Creemers MC, Radstake RA, van Laarhoven AI, et al. Experimental stress in inflammatory rheumatic disease: a review of psychophysiological stress responses. Arthritis Res Ther 2010;12:R89.
9. Kaleb M, Wipfler K, Shaw Y, Simon TA, Cornish A, England BR, et al. Experiences of patients with rheumatic diseases in the United States during early days of the COVID-19 pandemic. ACR Open Rheumatol 2020;2:335–43.
10. Edwards CJ, Fautrel B, Schulze-Koops H, Hiuzinga TW, Kruger K. Dosing down with biologic therapies: a systematic review and clinicians' perspective. Rheumatology (Oxford) 2017;56:1847–56.
11. NYC Health. COVID-19: data. URL: https://www1.nyc.gov/site/doh/covid/covid-19-data.page.
12. Onecia B, Bansl P, Goyal A, Lappin SL. Disease modifying antirheumatic drugs (DMARD). Treasure Island (FL): StatPearls Publishing; 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK507863/.
13. Pawluch D, Neiterman E. What is grounded theory and where does it come from? In: Bourgeault I, Dingwall R, De Vries R, editors. The SAGE handbook of qualitative methods in health research. London: SAGE Publications; 2010. Chapter 9.
14. Morse JM. Critical analysis of strategies for determining rigor in qualitative inquiry. Qual Health Res 2015;25:1212–22.
15. Creswell JW, Poth CN. Qualitative inquiry and research design: choosing among five approaches. 4th ed. Thousand Oaks (CA): Sage Publications; 2018.
16. Hadi MA, Closs SJ. Ensuring rigour and trustworthiness of qualitative research in clinical pharmacology. Int J Clin Pharm 2016;38:641–6.
17. Chan SJ, Stamp LK, Liebergreen N, Ndukwe H, Marra C, Treharne GJ. Tapering biologic therapy for rheumatoid arthritis: a qualitative study of patient perspectives. Patient 2020;13:225–34.
18. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020; 15:e0232076.
19. Strauss AL, Corbin JM. Basics of qualitative research: techniques and procedures for developing grounded theory research. 2nd ed. Thousand Oaks (CA): Sage Publications; 1998.
20. Berkowitz M, Inui TS. Making use of qualitative research techniques. J Gen Intern Med 1998;13:195–9.
21. Campbell JL, Quincy C, Osserman J, Pedersen OK. Coding in-depth semistructured interviews: problems of unitization and intercoder reliability and agreement. Sociolog Meth Res 2013; 42:294–320.
22. Mancuso CA, Duculan R, Jannat-Khah D, Barbhaiya M, Bass AR, Mehta B. Rheumatic disease-related symptoms during the height of the COVID-19 pandemic. HSS J 2020;16 Suppl 1:36–44.
23. Mehta B, Salmon J, Ibrahim S. Potential shortages of hydroxychloroquine for patients with lupus during the coronavirus disease 2019 pandemic. JAMA Health Forum. April 2020. URL: https://jamanetwork.com/channels/health-forum/fullarticle/2764607.
24. Gianfrancesco MA, Hvrich KL, Gossec L, Strangfeld A, Carmona L, Mateus EF, et al. Rheumatic disease and COVID-10: initial data from the COVID-19 global rheumatology alliance provider registries. Lancet Rheumatol 2020;2:e250–3.
25. Edwards CJ, Fautrel B, Schulze-Koops H, Hiuzinga TW, Kruger K. Dosing down with biologic therapies: a systematic review and clinicians' perspective. Rheumatology (Oxford) 2017;56:1847–56.
26. Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guidelines for the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:1–25.
27. Lenert A, Lenert P. Tapering biologics in rheumatoid arthritis: a pragmatic approach for clinical practice. Clin Rheumatol 2017;36:1–8.
28. Stamp LK, Chan SJ, Marra C, Helme C, Treharne GJ. Tapering biologic therapy for people with rheumatoid arthritis in remission: a review of patient perspective and associated clinical evidence. Musculoskeletal Care 2019;17:161–9.

ARP Announcements

Association of Rheumatology Professionals
2200 Lake Boulevard NE, Atlanta, Georgia 30319
www.rheumatology.org

Download the New ACR Publications Mobile App

The brand-new ACR Publications app can be downloaded for free from the Apple store or Google Play. ACR members can log in for full-text access to all articles in A*rthritis Care & Research* and *Arthritis & Rheumatology*. Nonmembers can access abstracts of all *AC&R* and *A&R* articles, the full text of articles published more than one year ago, and select open-access articles published recently, as well as the full text of all articles from *ACR Open Rheumatology* and *The Rheumatologist*.

ARP Membership

The Association of Rheumatology Professionals (ARP), a division of the American College of Rheumatology, appreciates your continued membership and looks forward to serving you another year. Membership costs range from $30 to $140. ARP welcomes nurse practitioners, nurses, physician assistants, office staff, researchers, physical therapists, occupational therapists, assistants, and students. Student membership is complimentary; the Annual Meeting registration fee is waived for students who submit the required student verification letter. For information, go to www.rheumatology.org and select "Membership" or call 404-633-3777 and ask for an ARP staff member.

New ACR Journal Twitter Account (@ACR_Journals) and Social Media Editor

The ACR journals are heightening our focus on social media, to benefit authors and readers. Among our first activities is the introduction of an official ACR Journals Twitter account: @ACR_Journals. Followers will enjoy special features and the opportunity to engage with authors and other fellow professionals about studies published in *Arthritis Care & Research, Arthritis & Rheumatology,* and *ACR Open Rheumatology*. Authors of published articles will have the opportunity to use @ACR_Journals to share their work and engage in dialogue with others interested in the research. The journals welcome Dr. Paul Sufka of Minneapolis as our first Social Media Editor.