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. 5, May 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 • May 2021 • NO. 5

Cover image: The figure on the cover (from Seven et al, page 742, Supplementary Figure 1) shows the complete cartilaginous compartment in 9 MRI slices. The first slice (top left) was the most anterior slice where the joint space and at least the iliac bone was visible, and the ninth slice (bottom right) was the most posterior slice where the cartilaginous joint was visible.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 613–616
DOI 10.1002/acr.24198

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

Time to Stop the Fibromyalgia Criteria Wars and Refocus on Identifying and Treating Individuals With This Type of Pain Earlier in Their Illness

Daniel Clauw

In this issue of *Arthritis Care & Research,* Häuser et al compare the prevalence of fibromyalgia (FM), in a German population, using the 2016 modification of the 2010/2011 American College of Rheumatology criteria for FM (ACR FM 2016) to the prevalence of FM in the same population, using the new Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy (AAPT) criteria (1,2). The authors found that the prevalence of FM using the ACR FM 2016 criteria was 3.4%, whereas when using the AAPT criteria, 5.7% were diagnosed with FM. Not surprisingly, individuals who met the ACR FM 2016 criteria had higher average levels of all symptomatology than those that met the AAPT criteria.

This article by Häuser et al has an excellent and balanced discussion. However, because this was a data-driven study comparing the prevalence and clinical features of these 2 cohorts meeting different FM criteria, the authors were appropriately circumspect and failed to opine that one of these criteria is better than the other. Yet, there are a dizzying array of other FM criteria that have been developed. Some are classification criteria, some are purported to be diagnostic criteria, and others were developed as screening criteria. Each of these criteria identifies slightly different subsets of individuals, resulting in only modest agreement between the criteria when applied in practice or research. Similarly, each criteria set will give different estimates of the prevalence of FM in populations.

This lack of agreement between criteria sets is not good for the field of FM. In fact, some have even used the fact that different criteria identify different patients as evidence that FM is neither a legitimate diagnosis nor a concept (3,4). Even after many years of FM research, we still ask: Why are there so many FM criteria sets, and which one should be used in practice and research?

History of FM criteria

The first accepted criteria for FM were the result of diligent work published in 1990, and as such are known as the ACR 1990 criteria for the classification of FM (5). These criteria required that the individual have chronic widespread pain (defined as involving the upper and lower body, the right and left side of the body, and the axial skeleton), as well as the presence of 11 or more of possible 18 tender points (9 paired regions of the body that are painful when palpated with modest pressure). Another recommendation in the ACR 1990 criteria was to abandon the distinction between primary FM (where FM is seen without a concurrent autoimmune or inflammatory condition) and secondary FM (where there is such an identifiable underlying condition).

These criteria were essential to research in the field of FM, and to the broader field of pain research. The ACR 1990 criteria enabled investigators around the world to perform mechanistic and other research into the pathophysiology and treatment of FM. However, this research also began to identify many problems with the ACR 1990 criteria. First, it became clear that the tenderness in FM was widespread and not at all confined to the 18 areas of the body considered tender points (6,7). Second, studies showed that although many male individuals experienced chronic widespread pain, very few of them also had 11 or more tender points. Therefore, using the ACR 1990 criteria implied that FM was a disease almost exclusively affecting female individuals (~9:1 female:male ratio) (8). Finally, it was also clear that most practitioners had little experience in performing a tender point examination and therefore were reluctant to consider the diagnosis of FM in their patients.

For all these reasons, nearly 20 years after the first FM criteria were published, Dr. Wolfe and colleagues set out to develop new FM criteria that did not require performing a tender point count and that relied entirely on patient self-report of symptoms. The first of these criteria sets was the ACR 2010 preliminary criteria,

Daniel Clauw, MD: University of Michigan, Ann Arbor.
No potential conflicts of interest relevant to this article were reported.
Address correspondence to Daniel Clauw, MD, University of Michigan, 24 Frank Lloyd Wright Drive, PO Box 385, Ann Arbor, MI 48105. Email: dclauw@med.umich.edu.

Submitted for publication March 16, 2020; accepted in revised form March 24, 2020.

which were required to be administered by an examiner (9). These criteria identified widespread pain in a much different manner, that is, by having individuals identify up to 19 sites in the body in which they had experienced pain during the past week. With 19 possible sites assessed, the Widespread Pain Index (WPI) score can range from 0 to 19. These ACR 2010 criteria equally weighted the importance of nonpain symptoms and had the examiner query the individual about the presence and severity of symptoms such as fatigue, sleep, memory, and other somatic symptoms. The individual then subjectively scored that severity (using the Symptom Severity score [SSS], which ranges from 0 to 12). The examiner was given guidance that this score might approximate a 0–3 patient rating in fatigue, sleep, and memory problems, as well as a 0–3 subjective score of listed somatic symptoms; a point of emphasis in the article was a physician/examiner must make the rating after speaking with the patient.

Various cut points in the WPI and SSS scores in combination were then used to determine whether an individual met the diagnostic criteria for FM. This combined score was also proposed as an overall score of FM severity. Using these ACR 2010 criteria, the presence of FM was higher than when using the ACR 1990 criteria, primarily because many more male individuals could be included, as the ACR 2010 criteria did not require a threshold of tenderness (since women are inherently more pain and sensory sensitive then men). Using the ACR 2010 criteria or any of the newer symptom-based criteria will have closer to a 2:1 female-to-male ratio rather than 9:1 ratio. This sex ratio is very similar to those found with nearly all chronic pain conditions, which are typically 1.5–2-times more common in women than in men (10). However, this historical bias toward clinicians underdiagnosing FM in males continues to exist in clinical practice, as is seen in the current study by Häuser et al in this issue.

It became clear that the ACR 2010 criteria could be simply modified to be entirely self-report, by directly asking the person to rate the presence and severity of fatigue, sleep, and memory problems, as well as other related syndromes. This modification of the ACR 2010 criteria is termed the ACR 2011 Survey Criteria (11). In these criteria each symptom domain is scored as absent (0), mild (1), moderate (2), or severe (3) in severity (resulting in a component score ranging 0–9). Then individuals are queried regarding the presence of abdominal pain or cramps, headache, or depression and assigned 1 point for each positive response. Thus, this SSS score can range from 0 to 12. These 2011 ACR Survey criteria have similar cut points as the ACR 2010 criteria (i.e., requiring combinations of WPI and SSS scores to meet FM criteria) and can be used to calculate a summed score of the WPI and SSS to get a severity of "fibromyalgianess" score.

The 2016 modification of the ACR 2011 Survey Criteria are not new criteria; they use the ACR 2011 Survey Criteria questionnaire and simply propose a new scoring scheme to categorically diagnose FM (12). The need for the modification only became apparent as the measure became widely used. First, counting the number of bodily pain sites was not a good way to measure widespread pain, which is widely acknowledged to be the most important finding in determining the extent of central nervous system involvement in the pain. The problem with simply counting pain sites is that contiguous sites of pain that likely have a peripheral origin are counted as multifocal pain. For example, if an individual has a lumbar radiculopathy, they might have 3 or more sites of pain in the back and buttocks, which means something quite different from a pathophysiologic standpoint than having noncontiguous pain in 3 separate regions of the body. Thus, to meet diagnostic criteria for FM using these ACR 2016 criteria individuals must have pain in 4 of 5 body regions (left/right upper quadrants, left/right lower quadrants, axial skeleton). The new criteria also clarified that the diagnosis of FM was valid regardless of whether individuals had another disorder causing pain (i.e., formerly called secondary FM).

Around this same time the American Pain Society (APS) and the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks (ACTTION) Initiative on Methods, Measurement, and Pain Assessments in Clinical Trials (IMMPACT) undertook a broad effort to develop diagnostic criteria for a number of pain conditions using similar methodologies, resulting in the ACTTION-APS Pain Taxonomy (AAPT) effort (13). The AAPT process was somewhat more methodologically rigorous than the ACR criteria and examined many data sets to identify the optimal manner to judge widespread pain as well as the most discriminating nonpain symptoms. The AAPT FM criteria define FM as 1) 6 or more sites of pain from a total of 9 sites (head, left/right arm and leg, chest, abdomen, upper back and spine, lower back and spine); 2) moderate-to-severe sleep problems or fatigue, and 3) requirement that symptoms have been present for 3 months, as in the ACR 2016 criteria. Because the ACR criteria were often based on data from the National Databank for Rheumatic Diseases, the data used to derive the ACR criteria primarily queried individuals about musculoskeletal pain, whereas the AAPT process looked more broadly at pain. Accordingly, the AAPT criteria identify more individuals than the 2016 criteria, as these criteria define widespread pain using more body regions that are frequent sites of pain. For example, headache and visceral pain in the chest or abdomen are not counted in the ACR 2016 criteria, and the ACR criteria considers any axial pain to be 1 site, whereas the AAPT criteria count the upper and lower back as separate sites of pain. A number of other groups have come out with even more diagnostic or screening criteria for FM.

Given the history of FM criteria, what criteria should clinicians and researchers use to diagnose or measure FM?

To decide whether FM is either present or absent, either the ACR 2016 criteria or the AAPT criteria will suffice, even though the AAPT process was in many ways methodologically superior to

the ACR process. But in clinical practice and research, this author strongly prefers the 2011/2016 ACR Survey Criteria because it can be scored as a continuous quantitative measure. Dr. Wolfe was the first to suggest that the degree of "fibromyalgianess" is more important than whether an individual has enough symptoms to meet diagnostic criteria, and decades of excellent epidemiologic, mechanistic, or clinical trials suggest he was correct (14). Everyone has a little "fibromyalgianess," and some people have a lot. Many studies have shown that when FM is comorbid with knee or hip osteoarthritis, for example, each 1-point increase in the FM score makes individuals less responsive to surgery, whether the score increases from 2 to 4, or from 12 to 14 (the latter score moving the individual into an FM diagnosis) (15,16). This makes sense from a pathophysiologic standpoint; a higher FM score simply means that the central nervous system is more likely to be contributing to that individual's pain. Many studies have also suggested that higher FM scores are also indicative of opioid nonresponsiveness to pain (17). Again, this parallels mechanistic studies showing that the endogenous opioid system may play a role in the pathogenesis of this process, so much so that the opioid antagonist naltrexone can be helpful to treat individuals with FM (18,19). Another advantage of using the 2011/2016 2011/2016 ACR Survey Criteria is it easily identifies individuals with chronic pain who have issues with comorbid sleep, fatigue, or sleep problems, many of which have effective treatments. Finally, using a continuous measure to recognize subthreshold FM may identify individuals who will benefit from established treatments of FM. Practitioners should think of FM along the lines of rheumatoid arthritis or gout: if you wait until it is too advanced to diagnose and treat it, the damage may be done.

Decades of research into FM have now clearly shown objective abnormalities in central nervous pain and sensory processing, autonomic function, and even low grade inflammation (20). In fact, FM is now acknowledged as the prototypical condition where pain and other symptoms originate from the central nervous system and systemic factors rather than from ongoing inflammation (nociceptive pain) or nerve damage (neuropathic pain). This third mechanism of pain has been officially adopted by the International Association for the Study of Pain and is termed nociplastic pain (21). Nociplastic pain mechanisms are thought to play the primary role in conditions such as FM, tension headache, irritable bowel syndrome, and many chronic overlapping pain conditions, as well as often being superimposed upon nociceptive or neuropathic pain conditions such as autoimmune disorders.

We have come a long way in better understanding FM. In just 3 decades the condition has moved from being considered a wastebasket term to describe individuals without any underlying pathophysiologic processes to explain their pain and other symptoms, to now being thought to be the prototypical nociplastic pain condition. The criteria wars have brought us to a place where there are now several instruments that can and should be used to screen, diagnose, and treat individuals for FM and "fibromyalgianess" earlier, before psychological, behavioral, and other comorbidities make this condition difficult or impossible to reverse.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Häuser W, Brähler E, Ablin J, Wolfe F. Modified 2016 American College of Rheumatology fibromyalgia criteria, the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society Pain Taxonomy, and the prevalence of fibromyalgia. 2021;73:617–25.
2. Hauser W, Perrot S, Clauw DJ, Fitzcharles MA. Unravelling fibromyalgia-steps toward individualized management. J Pain 2018;19:125–34.
3. Walitt B, Katz RS, Bergman MJ, Wolfe F. Three-quarters of persons in the US population reporting a clinical diagnosis of fibromyalgia do not satisfy fibromyalgia criteria: the 2012 National Health Interview Survey. PLoS One 2016;11:e0157235.
4. Wolfe F, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RL, Mease PJ, et al. Comparison of physician-based and patient-based criteria for the diagnosis of fibromyalgia. Arthritis Care Res (Hoboken) 2016;68:652–9.
5. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: report of the Multicenter Criteria Committee. Arthritis Rheum 1990;33:160–72.
6. Petzke F, Clauw DJ, Ambrose K, Khine A, Gracely RH. Increased pain sensitivity in fibromyalgia: effects of stimulus type and mode of presentation. Pain 2003;105:403–13.
7. Petzke F, Gracely RH, Park KM, Ambrose K, Clauw DJ. What do tender points measure? Influence of distress on 4 measures of tenderness. J Rheumatol 2003;30:567–74.
8. Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence and characteristics of fibromyalgia in the general population. Arthritis Rheum 1995;38:19–28.
9. Wolfe F, Clauw DJ, FitzCharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10.
10. Vincent A, Lahr BD, Wolfe F, Clauw DJ, Whipple MO, Oh TH, et al. Prevalence of fibromyalgia: a population-based study in Olmsted County, Minnesota, utilizing the Rochester Epidemiology project. Arthritis Care Res (Hoboken) 2013;65:786–92.
11. Wolfe F, Clauw DJ, FitzCharles MA, Goldenberg DL, Hauser W, Katz RS, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR preliminary diagnostic criteria for fibromyalgia. J Rheumatol 2011;38:1113–22.
12. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RL, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–29.
13. Arnold LM, Bennett RM, Crofford LJ, Dean LE, Clauw DJ, Goldenberg DL, et al. AAPT Diagnostic Criteria for Fibromyalgia. J Pain 2019;20:611–28.
14. Wolfe F. Fibromyalgianess. Arthritis Rheum 2009;61:715–6.
15. Brummett CM, Janda AM, Schueller CM, Tsodikov A, Morris M, Williams DA, et al. Survey criteria for fibromyalgia independently predict increased postoperative opioid consumption after lower-extremity joint arthroplasty: a prospective, observational cohort study. Anesthesiology 2013;119:1434–43.
16. Brummett CM, Urquhart AG, Hassett AL, Tsodikov A, Hallstrom BR, Wood N, et al. Characteristics of fibromyalgia independently predict

poorer long-term analgesic outcomes following total knee and hip arthroplasty. Arthritis Rheumatol 2015;67:1386–94.

17. Janda AM, As-Sanie S, Rajala B, Tsodikov A, Moser SE, Clauw DJ, et al. Fibromyalgia survey criteria is associated with increased postoperative opioid consumption in women undergoing hysterectomy. Anesthesiology 2015;122:1103–11.
18. Schrepf A, Harper DE, Harte SE, Wang H, Ichesco E, Hampson JP, et al. Endogenous opioidergic dysregulation of pain in fibromyalgia: a PET and fMRI study. Pain 2016;157:2217–25.
19. Younger J, Noor N, McCue R, Mackey S. Low-dose naltrexone for the treatment of fibromyalgia: findings of a small, randomized, double-blind, placebo-controlled, counterbalanced, crossover trial assessing daily pain levels. Arthritis Rheum 2013;65:529–38.
20. Sluka KA, Clauw DJ. Neurobiology of fibromyalgia and chronic widespread pain. Neuroscience 2016;338:114–29.
21. Kosek E, Cohen M, Baron R, Gebhart GF, Mico JA, Rice AS, et al. Do we need a third mechanistic descriptor for chronic pain states? Pain 2016;157:1382–6.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 617–625
DOI 10.1002/acr.24202

American College of Rheumatology
Empowering Rheumatology Professionals

Modified 2016 American College of Rheumatology Fibromyalgia Criteria, the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society Pain Taxonomy, and the Prevalence of Fibromyalgia

Winfried Häuser,[1] Elmar Brähler,[2] Jacob Ablin,[3] and Frederick Wolfe[4]

Objective. To study the prevalence of fibromyalgia (FM) in the general population according to a 2016 modification of the American College of Rheumatology criteria (FM 2016) and the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy criteria (AAPT), and to compare diagnostic and clinical variables between the criteria sets.

Methods. We studied 2,531 randomly selected subjects from the German general population in 2019. Pain regions from the Michigan Body Map were fitted to the FM 2016 and the AAPT criteria, and criteria symptom items were derived from validated questionnaires assessing somatic and psychological symptom burden and disability. We determined FM criteria prevalence and criteria-related scales including widespread and multisite pain (MSP) and symptom scales, and measured symptom burden and disability.

Results. According to the FM 2016 criteria, the prevalence of FM was 3.4% (n = 75 subjects; 95% confidence interval [95% CI] 2.7, 4.3) compared with 5.7% (n = 130 subjects; 95% CI 4.8, 6.8) for the AAPT criteria; κ = 0.65. Compared with AAPT-positive subjects, FM 2016–positive subjects had higher MSP, Widespread Pain Index score, Polysymptomatic Distress Scale scores, Symptom Severity Scores, and psychological symptom burden. Physician-diagnosed FM was reported by 1.1% of the subjects. Of these, 44.0% met the FM 2016 criteria, and 47.5% met the AAPT criteria.

Conclusion. The prevalence of FM in the German general population is 73% greater using the AAPT criteria than the FM 2016 criteria. The AAPT criteria select individuals with less symptom severity and fewer pain sites. The FM 2016 criteria, but not the AAPT criteria, provide a general severity measure for FM.

INTRODUCTION

The first criteria for fibromyalgia (FM) endorsed by a medical association were the American College of Rheumatology (ACR) 1990 criteria (1). Although not designed as diagnostic criteria, they were used in clinical and epidemiology studies to identify cases of FM. The ACR 1990 criteria foundered and were ultimately abandoned on the issue of tender points, which involved an examination that was difficult to perform, biased, and rarely used in clinical practice, as well as on the issue of the uncertainty of the definition and meaning of the criterion of widespread pain (4-quadrant pain) (2).

New diagnostic criteria without tender point counts were published in 2010 as the ACR preliminary diagnostic criteria (3) and then modified for self-report in the next year (4). The criteria included the assessment of pain at 19 sites and a 4-item Symptom Severity Scale, from which an overall FM severity score, the Polysymptomatic Distress (PSD) Scale score, could be calculated. In 2016, a further modification of the American College of Rheumatology criteria added a widespread pain criterion and clarified scoring (FM 2016) (5). In 2018, the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society Pain Taxonomy group published simplified pain taxonomy criteria for FM (AAPT) (6). These

[1]Winfried Häuser, MD: Klinikum Saarbrücken, Saarbrücken, Germany, and Technische Universität München, Munich, Germany; [2]Elmar Brähler, PhD: University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany; [3]Jacob Ablin, MD: Institute of Rheumatology, Tel Aviv Sourasky Medical Center, Tel Aviv University, Tel Aviv, Israel; [4]Frederick Wolfe, MD: National Data Bank for Rheumatic Diseases, Wichita, Kansas.

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

Address correspondence to Frederick Wolfe, MD, National Data Bank for Rheumatic Diseases, 1035 North Emporia, Suite 288, Wichita, KS 67214. Email: fwolfe@arthritis-research.org.

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

SIGNIFICANCE & INNOVATIONS

- Although the criteria are similar, the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy (AAPT) criteria are used to diagnose 78% more individuals with fibromyalgia (FM) than a 2016 modification of the American College of Rheumatology criteria (FM 2016).
- The AAPT criteria select individuals with less symptom severity and fewer pain sites.
- The FM 2016, but not the AAPT criteria, provide a general FM severity measure.

criteria diagnosed FM based on the presence of pain at 6 of 9 pain sites (called multisite pain) and the presence of at least moderate fatigue or sleep problems.

There is considerable concordance between the 2 diagnostic criteria in that a central symptom of FM is chronic widespread/multisite pain that requires a defined number of pain sites. But there are also differences: how many pain sites or regions are required to diagnose widespread/multisite pain, and on the importance of fatigue, sleep problems and additional somatic and psychological symptoms (Table 1). The criterion of widespread pain of the FM 2016 uses musculoskeletal pain sites, defining FM to be a musculoskeletal disorder according to the meaning of the term "fibromyalgia." In contrast to the FM 2016 criteria, the AAPT multisite pain criterion includes the head, chest, and abdomen as 3 of the 9 pain areas. Pain in these areas is usually not of musculoskeletal origin. Abdominal pain and headache are also minor symptoms in the Somatic Severity Scale of the FM 2016 criteria. Thus, visceral pain syndromes and any headaches might contribute to the diagnosis of FM according to both criteria.

Systematic reviews have found the average prevalence of FM in the general population to be 2% (7). Epidemiology studies mainly used the ACR 1990 and ACR 2011 preliminary criteria. To the best of our knowledge, the prevalence of FM syndrome, according to the FM 2016 and the AAPT criteria and their concordance, has not been assessed in a sample representative of the general population until now. In addition, we do not know if cases of FM identified by the FM 2016 and AAPT criteria differ according to demographic and clinical variables. Therefore, the aims of this study were the following: 1) to compare the prevalence of FM according to the FM 2016 and AAPT criteria and to assess the concordance of the FM 2016 and AAPT criteria; 2) to compare FM criteria and cases according to the FM 2016 and AAPT criteria with regards to demographic and clinical variables; and 3) to compare specific criteria items, including multisite pain, widespread pain, Widespread Pain Index (WPI) score, fatigue and sleep problems, and Somatic Severity Scale score, that contribute to criteria positivity and differences in prevalence rates between the 2 criteria sets.

SUBJECTS AND METHODS

Design and subjects. A random sample of German residents ages ≥14 years was selected for a cross-sectional survey that included several questionnaires on somatic and psychological symptoms (a health survey) as well as questionnaires on eating and political attitudes. The sampling design included 3 consecutive steps. First, a sample of 258 living areas was randomly selected from a nonoverlapping stratum of all area units:

Table 1. 2016 modification of the American College of Rheumatology fibromyalgia (FM) diagnostic criteria and the AAPT criteria*

Reference	Criteria	Collection of the data needed for diagnosis	Exclusion of other diseases
2016 revisions to the 2010/2011 FM diagnostic criteria (ref. 5)	1) Evaluate the presence of pain at 19 sites with self-report form; 2) evaluate 4-item symptom scale on same form; diagnosis requires: 3) widespread pain, defined as pain in at least 4 of 5 regions (except face, chest, and abdomen); 4) pain score (WPI) ≥7 and SSS ≥5 or WPI score 4–6 and SSS score ≥9; 5) symptoms have been present at a similar level for at least 3 months	1-page self-report form	A diagnosis of FM is valid irrespective of other diagnoses. A diagnosis of FM does not exclude the presence of other clinically important illnesses. A diagnosis of FM does not mean it is the patient's only diagnosis or even the most important diagnosis. It is only an acknowledgment that the patient has symptoms of FM and satisfies FM criteria.
AAPT diagnostic criteria (ref. 6)	1) Multisite pain defined as ≥6 pain sites from a total of 9 possible sites; 2) moderate-to-severe sleep problems or fatigue; 3) multisite pain plus fatigue or sleep problems must have been present for at least 3 months	Patient self-report (using established scales for the symptoms) or by clinician rating†	The presence of another pain disorder or related symptoms does not rule out a diagnosis of FM. However, a clinical assessment is recommended to evaluate for any condition that could fully account for the patient's symptoms or contribute to the severity of the symptoms.

* AAPT = Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy criteria; SSS = Symptom Severity Score; WPI = Widespread Pain Index.
† Arnold LM: personal communication.

210 areas were sampled from Western Germany, and 48 areas were sampled from Eastern Germany. The random selection of households was implemented in the second step. Finally, 1 individual matching the inclusion criteria was randomly selected from each household. The procedure was designed to yield a national sample representative in terms of age, sex, and education.

Sociodemographic data were collected by trained interviewers face-to-face. In addition, participants completed a battery of self-report questionnaires. The assistant waited until the participants answered all questionnaires and offered to help if participants did not understand the meaning of the questions.

Data collection took place between May 2019 and July 2019. A first attempt was made at 5,393 addresses, and 2,531 individuals (45.9%) participated fully. Reasons for nonparticipation included the following: 4 unsuccessful attempts to contact (13.6%) or meet the selected household member (3.0%); the household member declined to participate (22.9%); the member was on a holiday break (0.6%); the member was ill and not able to follow the interview (0.5%); or the member declined to interview (12.3%). All participants were informed about the study procedures and signed an informed consent form. The study was approved by the Institutional Ethics Review Board of the University of Leipzig (Az 145/19-ek).

Questionnaires and definition of FM diagnostic criteria. *Michigan Body Map (MBM).* Using a sociodemographic questionnaire, we assessed age, sex, marital status, educational status, current professional status, and family income. To facilitate the collection of survey-based criteria regarding widespread body pain, a German version of the MBM (8) was created. The MBM is a graphic mannequin with the 19 areas from the WPI superimposed on it in anatomically relevant locations. In addition, the MBM contains 16 additional pain sites.

Somatic Symptom Scale 8. The Somatic Symptom Scale 8 is the short form of the Patient Health Questionnaire 15 (PHQ-15) (9). It comprises 8 items (stomach or bowel problems; back pain; pain in the arms, legs, or joints; headaches; chest pain or shortness of breath; dizziness; feeling tired or having low energy; and trouble sleeping), with each symptom scored from 1 (not bothered at all) to 5 (bothered very strongly) within the last 7 days (10).

The PHQ-4 and Short Form 12 (SF-12) Health Survey items. The 4-item PHQ for depression and anxiety was used to assess psychological symptom burden. Respondents rated how often, on a scale from 0 (not at all) to 3 (nearly every day), they experienced feeling depressed, loss of interest, feeling nervous or anxious, and an inability to stop worrying over the past 2 weeks (11,12,13). The total score therefore ranges from 0 to 12. We used the items "general health" (0 = excellent; 1 = very good; 2 = good; 3 = fair; 4 = poor) and "pain interference" ("During the past 4 weeks, how much did pain interfere with your normal work including both work outside the home and housework?" [0 = not at all; 1 = a little bit; 2 = moderately; 3 = quite a bit; 4 = extremely]) to assess health-related quality of life.

The Self-Administered Comorbidity Questionnaire (SCQ). The SCQ assessed the presence, treatment, and functional limitations of 13 common diseases (heart disease; high blood pressure; lung disease; diabetes mellitus; ulcer or other stomach disease; kidney disease; liver disease; anemia or other blood disease; cancer; depression; rheumatic disorders; pancreatic disease and inflammatory bowel disease). For the last 2 diagnoses, we substituted pancreatic disease for osteoarthritis and inflammatory bowel disease for low back pain. We calculated the presence of the disease using the associated disability subscale (range 0–13) (14,15).

Bodily Distress Syndrome Checklist (BDS-25). The BDS-25 is a self-report instrument that may be used for case finding of BDS in clinical care and research (16). It includes 25 items covering cardiopulmonary, gastrointestinal, musculoskeletal, and general symptoms scored from 0 (not at all) to 4 (a lot).

FM 2016 variables. The WPI in this study was derived from the MBM. The FM 2016 criterion of widespread pain (5) is satisfied by the presence of pain in 4 or 5 musculoskeletal body regions. It is composed of 5 regions (axial, left upper, right upper, left lower, right lower) of 3 pain sites each (15 total sites). When pain in 4 or 5 regions is noted, the criterion of widespread pain of the FM 2016 is satisfied. Jaw, chest, and abdominal pain are not included in the assessment of pain region.

The WPI (3,17) provides a score of painful sites that ranges from 0 to 19. It was first defined in the ACR 2010 criteria and their 2011 self-report modification. It is a measure of extensiveness of pain. It includes the 15 sites of the widespread pain criterion plus jaws, chest, and abdomen. We used a WPI score of ≥7 as a marker of widespread pain severity in this study.

The Symptom Severity Score (SSS) is a measure of symptom severity (range 0–12) first defined in the ACR 2010 criteria and their 2011 self-report modification (3,17). It includes measures of fatigue (range 0–3), unrefreshed sleep (range 0–3), cognitive difficulties (range 0–3), headache (range 0–1), pain or cramps in the lower abdomen (range 0–1), and depression (range 0–1). It measures somatic and nonsomatic symptoms of FM. In general, an SSS score of ≥5 is required for a diagnosis of FM according to the FM 2016. For the fatigue and sleep items, we utilized the 7th and 8th items of the Somatic Symptom Scale 8. We created the items included in the SSS as follows: pain or cramps in the lower abdomen (item 8 of the BDS-25 [quite a bit or a lot]); and depression (item 2 of the PHQ-4 [several days; more than one-half of days; nearly every day] and item 23 of the BDS-25 [somewhat, quite a bit, a lot]). We used an SSS score of ≥5 as a marker for somatic symptom severity in this study and compared it to "fatigue or sleep problems moderate or greater" used in the AAPT criteria.

The PSD Scale score (also called the FM Severity Scale score) (range 0–31) is the sum of the WPI and SSS scores (3,17). The PSD Scale measures the magnitude and severity of fibromyalgia

symptoms ("fibromyalgianess"). By definition, FM criteria cannot be satisfied if the PSD Scale score is <12. The FM 2016 criteria require the following: 1) WPI score ≥7 and SSS score ≥5, or a WPI score 4–6 and an SSS score ≥9; 2) the presence of widespread pain as defined above; and 3) symptoms of at least 3 months in duration (5).

AAPT criteria. Participants satisfied the AAPT criteria if at least 6 of 9 of the following sites were endorsed for pain: head or face, left arm, right arm, chest, abdomen or pelvis, upper back and spine, lower back and spine including buttocks, left leg, and right leg (6). In addition, the subject must have fatigue or sleep problems of at least moderate severity. For the fatigue and sleep items, we utilized the 7th and 8th items of the Somatic Symptom Scale 8.

Statistical analyses. Data were analyzed by Stata, version 15.1, using survey methods that incorporated sampling weights (18). In instances where diagnostic groups were compared and subjects might be a member of >1 group at a time, we used seemingly unrelated estimation (SUEST) to characterize group differences. The SUEST procedure combines information from the 2 models and then tests the null hypothesis that the coefficients are equivalent across the 2 models (18).

The concordance between the different criteria was assessed using the kappa statistic. The kappa statistic was interpreted as none (κ = 0–0.20), minimal (κ = 0.21–0.39), weak (κ = 0.40–0.59), moderate (κ = 0.60–0.79), strong (κ = 0.80–0.90), and almost perfect (κ > 0.90) (19).

RESULTS

The mean age of study participants was 48.7 years (95% confidence interval [95% CI] 47.6, 47.9), and 51.0% (95% CI 48.8, 53.1) were women. Among individuals satisfying the FM 2016 or AAPT criteria, 67% were women. The FM 2016 and AAPT criteria were met by 3.4% (n = 75; 95 CI 2.7, 4.3) and 5.7% (n = 130; 95 CI 4.8, 6.8) of patients, respectively (Table 2). The prevalence of FM according to the FM 2016 and the AAPT increased with age (Figure 1 and Table 2). For the FM 2016 and AAPT, respectively, the prevalence of FM was 1.0% and 2.6% (ages <40 years), 2.6%

Table 2. Fibromyalgia (FM) criteria–related variables for the FM 2016 and AAPT criteria*

	All (n = 2,531)	FM 2016 criteria	AAPT criteria	FM 2016–AAPT difference	*P*
Age, years	48.7 (47.6, 47.9)	64.4 (60.3, 68.6)	60.9 (57.3, 64.3)	3.5 (0.5, 6.6)	0.023
Female sex, % (95% CI)	51.0 (48.8, 53.1)	66.8 (55.0, 76.9)	66.6 (57.6, 73.2)	1.0 (0.7, 1.5)†	0.984
FM criteria					
FM criteria, no. positive		75	130		
FM criteria (% +), all ages		3.4 (2.7, 4.3)	5.7 (4.8, 6.8)	2.3 (1.6, 3.0)	0.000
FM criteria (% +), ages <40 years		1.0 (0.5, 1.9)	2.6 (1.7, 3.9)		
FM criteria (% +), ages 40–65		2.6 (1.8, 3.8)	4.6 (3.4, 6.0)		
FM criteria (% +), ages ≥65		8.0 (5.8, 11.0)	11.9 (9.2, 15.3)		
Criteria-related variables					
FM 2016 WSP criterion (% +)	6.9 (5.8, 8.1)	100.0	75.6 (67.3, 82.4)	Determined	
ACR 1990 WSP criterion (% +)	11.2 (9.9, 12.6)	98.9 (92.7, 99.8)	94.9 (88.4, 97.0)	4.4 (0.5, 35.7)†	0.170
Multisite pain score ≥6 (% +)	7.3 (6.3, 8.6)	90.0 (82.1, 97.2)	100.0	Determined	
Multisite pain score (range 0–9)	1.1 (1.0, 1.1)	7.0 (6.7, 7.2)	6.9 (6.7, 7.1)	0.1 (–0.1, 0.3)	0.343
WPI score ≥7 criterion, % (95% CI)	5.6 (4.7, 6.7)	89.9 (79.9, 95.2)	66.8 (57.8, 74.8)	3.2 (1.4, 7.1)†	0.005
WPI score (range 0–19)	1.0 (0.9, 1.1)	9.8 (8.9, 10.6)	8.4 (7.7, 9.1)	1.4 (0.8, 1.9)	0.000
SSS score ≥5 criterion, % (95% CI)	14.3 (12.8, 15.8)	100.0	80.3 (72.1, 86.5)	Determined	
SSS score (range 0–12)	1.8 (1.7, 1.9)	7.9 (7.4, 8.4)	7.0 (6.5, 7.5)	Determined	
PSD score (range 0–31)	2.7 (2.6, 2.9)	17.7 (16.7, 18.7)	15.4 (14.5, 16.3)	2.3 (1.5, 3.1)	0.000
PSD score ≥12 criterion, % (95% CI)	6.5 (5.5, 7.7)	100.0	77.5 (68.9, 84.2)	Determined	
Somatic Symptom Scale 8 sleep or fatigue ≥ moderate, % (95% CI)	22.9 (21.1, 24.7)	100.0	100.0	Determined	
Non-criteria variables					
Somatic Symptom Scale 8 (range 0–32)	3.9 (3.7, 4.1)	16.4 (15.3, 17.5)	15.5 (14.6, 16.4)	1.0 (0.1, 1.9)	0.034
SCQ comorbidity score (range 0–13)	0.6 (0.5, 0.7)	3.0 (2.4, 3.6)	2.6 (2.1, 3.0)	0.4 (0.1, 0.8)	0.022
SF-12 general health (range 0–4)	1.6 (1.5, 1.6)	3.1 (3.0, 3.3)	3.0 (2.9, 3.1)	2.1 (0.7, 3.5)	0.001
SF-12 pain interference (range 0–4)	1.6 (1.6, 1.7)	2.8 (2.6, 3.0)	2.6 (2.4, 2.8)	0.2 (0.1, 0.4)	0.010
SF-12 pain interference ≥ moderate, % (95% CI)	0.6 (0.5, 0.6)	93.1 (84.0, 97.2)	87.7 (81.9, 93.5)	1.7 (0.8, 3.6)†	0.187
PHQ-4 total score (range 0–12)	2.2 (2.1, 2.3)	7.6 (6.7, 8.5)	6.9 (6.1, 7.7)	0.7 (0.1, 1.4)	0.028
Physician diagnosis of FM, % (95% CI)	1.1 (0.9, 1.6)	14.6 (8.4, 24.2)	9.4 (5.5, 15.5)	1.6 (1.0, 2.6)†	0.267

* Values are the mean (95% confidence interval [95% CI]) unless indicated otherwise. % + = percent positive for the condition under study; AAPT = Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy criteria; ACR 1990 = American College of Rheumatology (ACR) 1990 criteria for FM; Determined = difference values not calculated because one value is determined by criterion requirement; FM 2016 = modification of the ACR criteria for FM; OR = odds ratio; PHQ-4 = Patient Health Questionnaire 4; PSD = Polysymptomatic Distress Scale; SCQ = Self-Administered Comorbidity Questionnaire; SF-12 = Short Form 12 Health Survey; SSS = Symptom Severity Score; WPI = Widespread Pain Index; WSP = widespread pain.
† Values are the OR (95% CI).

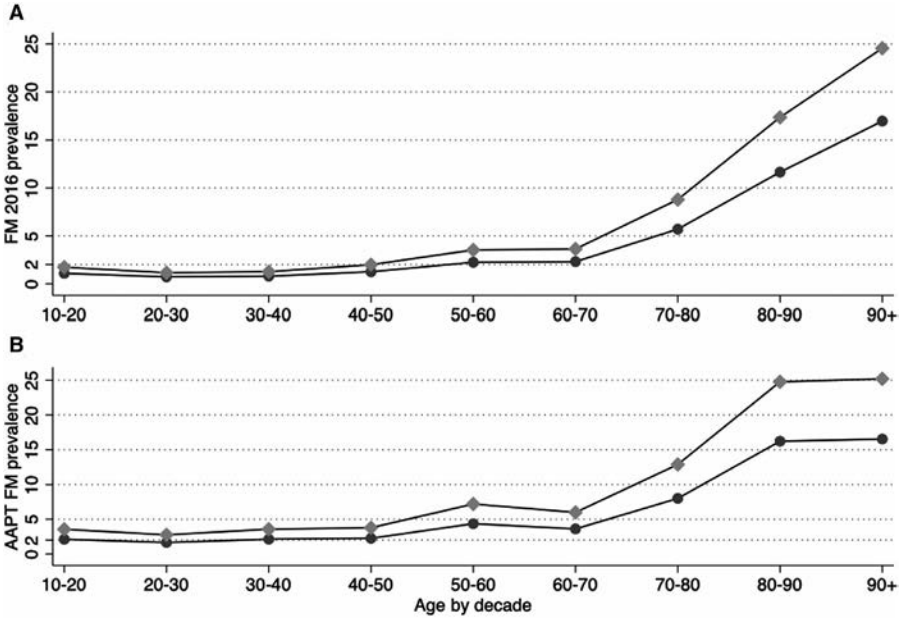


Figure 1. The prevalence of fibromyalgia (FM) by a 2016 modification of the American College of Rheumatology criteria for FM (FM 2016) (**A**) and the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy criteria (**B**). Circles represent male participants. Diamonds represent female participants.

and 4.6% (ages 40–65 years), and 8.0% and 11.9% (ages ≥65). There was no significant difference in prevalence by sex. Of the 137 subjects who satisfied either the FM 2016 or AAPT criteria, 62 (45.3%) satisfied only the AAPT criteria, 68 (49.6%) satisfied the FM 2016 and AAPT criteria, and 7 (5.4%) satisfied only the FM 2016 criteria (Table 3). The kappa statistic for the FM 2016 and the AAPT was 0.650.

As shown in Table 2, the FM 2016 widespread pain criterion was satisfied by 100% of patients using the FM 2016 and by 76% using the AAPT criteria. When the widespread pain criterion of pain site extent was used (WPI score ≥7), it was satisfied by 89.9% of subjects positive by the FM 2016 and 66.8% of subjects positive by the AAPT criteria. The widespread pain criterion of the AAPT (multisite pain score ≥6) was noted in 90.0% of subjects positive by the FM 2016 and 100% of subjects positive by the AAPT criteria. Of the subjects positive for FM by the FM 2016 and the AAPT criteria, 100% met the criterion of having moderate or greater fatigue or sleep problems. The FM 2016 symptom criterion (SSS score ≥5) was met by 100% of subjects positive by the FM 2016 and 80.3% of subjects positive by the AAPT criteria. The PSD Scale score was greater in the FM 2016 than in the AAPT criteria (17.7 versus 15.4), and more individuals satisfied the criterion of PSD Scale score ≥12 in the FM 2016 compared with the AAPT criteria (100.0% versus 77.5%) (Table 2).

We also examined non-criteria severity items in Table 2. There were significant differences between the criteria for the Somatic Symptom Scale 8, the Somatic Symptom Scale 6, the SF-12 general health subscale, the SF-12 pain interference subscale, PHQ-4 total scores, and SCQ score.

Table 3. Comparison of fibromyalgia (FM) criteria variables among FM-positive groups*

Variables	FM 2016 (–) AAPT (+)	FM 2016 (+) AAPT (+)	FM 2016 (+) AAPT (–)
Subjects in group, no. (%)	62 (45.3)	68 (49.6)	7 (5.4)
Age, years	58.1 (53.1, 63.1)	63.4 (58.9, 67.8)	74.3 (68.0, 80.5)
Female, %	63.8 (50.4, 75.3)	69.0 (56.6, 79.2)	47.3 (16.2, 80.6)
WPI score (range 0–19)	6.6 (5.9, 7.4)	10.0 (9.0, 10.9)	8.0 (6.7, 9.3)
SSS score (0–12)	5.7 (5.1, 6.3)	8.1 (7.6, 8.6)	6.5 (5.2, 7.8)
PSD score (0–31)	12.4 (11.5, 13.2)	18.0 (17.0, 19.1)	14.5 (13.0, 16.1)
AAPT pain site score (0–9)	6.5 (6.3, 6.7)	7.2 (7.0, 7.5)	5.0 (NC)

* Values are the mean (95% confidence interval) unless indicated otherwise. AAPT = Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy criteria; FM 2016 = modification of the American College of Rheumatology criteria for FM; NC = not calculable; PSD = Polysymptomatic Distress Scale; SSS = Symptom Severity Score; WPI = Widespread Pain Index.

Patient-reported physician diagnosis of FM was 1.1% (95% CI 0.9, 1.6). There were no statistically significant differences between the proportion of physician-diagnosed cases of FM for the FM 2016 and the AAPT criteria (14.6 % versus 9.4%) (Table 2).

To further illuminate differences between the FM 2016 and AAPT criteria, in Table 3 we compared 3 criteria groups: 62 subjects who satisfied the AAPT criteria but not the FM 2016 criteria, 68 subjects who satisfied both criteria, and 7 subjects who were FM 2016 positive and AAPT negative. The WPI score was substantially greater for the FM 2016/APPT–positive group compared with the AAPT group alone. This was also true of the AAPT pain site score and the SSS score. The overall FM severity score (using the PSD Scale score) was greatly elevated in the combined group compared to the AAPT solo group (18.0 versus 12.4). Figures 2A, B, and C show the distribution of WPI score, SSS score, and PSD Scale scores for the FM 2016 group separately from the AAPT group. It demonstrates that AAPT criteria shift the distribution curves to the left, indicating less FM criteria variable severity in AAPT subjects. Figure 2D shows the effect of examining AAPT (+) FM 2016 (-) separately from AAPT (-) FM 2016 (+) distributions. In the 62 AAPT-only subjects, the distribution curve is shifted much more to the left, indicating that solo AAPT criteria subjects have less severe FM symptoms.

We examined factors that might be responsible for the difference in prevalence between the 2 sets of criteria under investigation (Table 4). One such factor was the decision by the AAPT group to mandate as the somatic symptom variable the presence of either moderate or greater fatigue or moderate or greater sleep problems. This factor was present in 22.9% of the population, and all individuals meeting the AAPT or FM 2016 criteria satisfied the requirement (Table 2). By contrast, a more rigorous requirement for somatic symptoms variables was the SSS score ≥5 of the FM 2016 criteria. It was found in 14.3% of the general population but only 80.3% of those meeting AAPT criteria. In Table 4, we recalculated the prevalence of positivity by the AAPT criteria using the SSS definition. We found that the prevalence would have been 3.2%, less than the prevalence by the FM 2016 criteria. To be certain that this observation could not be the result of the questionnaire we used, we noted that 41.8% of those who satisfied the fatigue or sleep variable also satisfied the condition of an SSS score ≥5. Then we returned to the 2013 study of the German population that used the full set of ACR 2011 questions (which are the same in the FM 2016) (20) and found that 34.2% of those with moderate or greater fatigue or sleep problems satisfied the SSS criterion. Thus, difference in questionnaires could not be the result of the difference in the prevalence of the somatic symptom condition.

In a similar vein, we explored the effect of not scoring pain in the AAPT pain sites in the 3 nonmusculoskeletal regions: head and face, chest, and abdomen/pelvis. On that assumption, the prevalence by AAPT criteria would be 3.5% (Table 4).

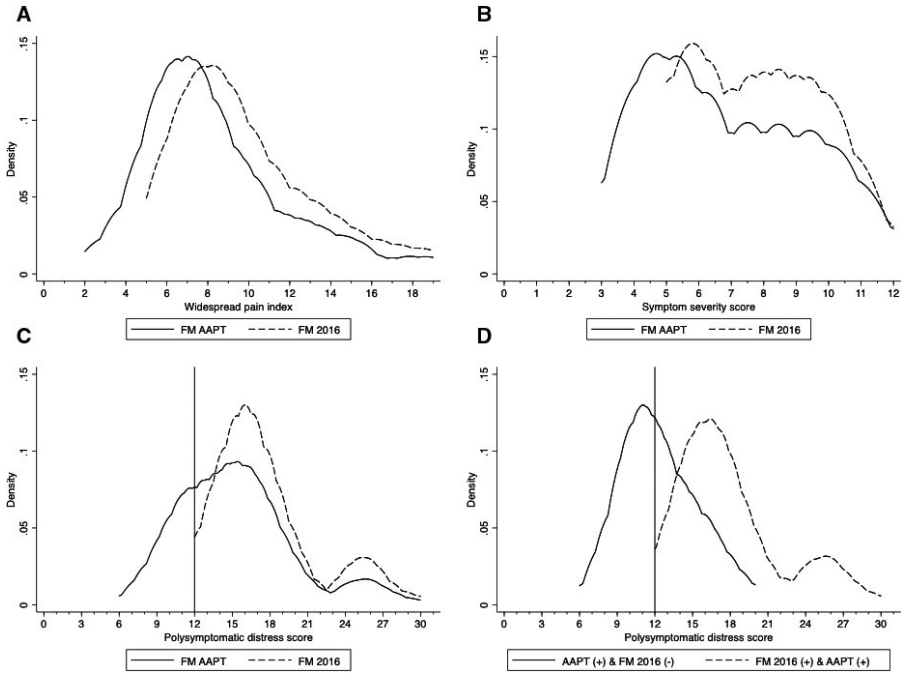


Figure 2. Distribution of Widespread Pain Index (**A**), Symptom Severity Score (**B**), and Polysymptomatic Distress Scale (**C**) scores for the fibromyalgia (FM) 2016 group separately from the AAPT group and the effect of examining AAPT positive–FM 2016 negative separately from AAPT negative–FM 2016 positive distributions (**D**). AAPT = the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy criteria; FM 2016 = modification of the American College of Rheumatology criteria for FM.

DISCUSSION

This study shows that the prevalence of FM in the general population is 73% greater when diagnosed with the AAPT criteria than when diagnosed with the FM 2016 criteria. This occurs because it easier to meet the multisite pain criterion and single variable symptom criterion required for diagnosis than it is to meet the widespread pain, WPI, and symptom criteria of the FM 2016 criteria. It follows, therefore, that other scales that measure somatic and psychological symptom burden will be more abnormal in the FM 2016 criteria positive subjects.

The definition of FM and the various criteria used to diagnose it have changed dramatically over the >40 years of the disorder's recognized existence (21). So, it would be expected that different individuals would be identified as having FM at different prevalence levels by different criteria sets. In addition, we know that FM exists as a continuum of symptoms (20). The cutoff that separates fibromyalgia from non-fibromyalgia is not a statistical or disease measure; it represents the decisions of those who make criteria. The AAPT group picked a multisite pain cutoff of 6, but they could have picked 5 or 7, just as Wolfe et al (5) could have chosen other values for the FM 2016 criteria.

It may be useful to consider other pain cutoff points. The criteria used by Yunus et al (22, 23) required only pain or aching at ≥3 anatomic sites and a miscellany of non-pain symptoms. The ACR 1990 criteria (1) required 4 (rarely 3) pain sites compared to the 6 in the AAPT criteria, and conditionally more in the FM 2016 criteria. The ACR 1990 criteria did not require non-pain symptoms, but the mandatory tender point count was strongly influenced by psychological distress (24,25). Although there was no standardized way to measure the extent of pain involvement and non-pain symptoms prior to the introduction of the WPI, the SSS, and the PSD Scale, the distribution curves of Figure 2 provide information as to the effect of different criteria. Easier criteria move the key PSD Scale distribution curve to the left and increase apparent FM prevalence. Similarly, criteria that require more pain and more symptoms move the PSD Scale, the WPI, and the SSS curves to the right.

The data of the current study show that the AAPT criteria are easier to satisfy and that they identify individuals with FM with less pain and fewer symptoms than the FM 2016 criteria. But such data cannot tell us which criteria set is better, more accurate, or truer. The validity of the criteria examined here have to come from external sources, e.g., the comparison of the 2 criteria sets for patients in clinical settings. However, there are important differences between the criteria. The existence of the PSD Scale in the FM 2016 as a measure of severity, and the FM symptom continuum and its absence in the AAPT scale, are important. With the AAPT criteria, which did not provide a useful measure of severity, it can appear as though there are only 2 states: FM positive and FM negative. With the PSD Scale from the FM 2016 criteria, it is easier to understand the spectrum of symptoms and how one can go into and out of a diagnosis (26). Similarly, FM severity in the AAPT criteria is represented by the pain site score and ranges from 6 to 9 (4 units). By contrast, the PSD Scale score for those with FM according to the FM 2016 criteria ranges from 12 to 31 (20 units).

A second advantage of the FM 2016 criteria for a diagnosis of FM is that the AAPT, with 3 of 9 pain sites that are primarily nonmusculoskeletal (head and face, chest, and abdomen and pelvis), may be capturing symptoms of more than a musculoskeletal disorder. Rather, it may be identifying symptoms related to a broader somatic symptom disorder (27,28). Data from the current study show that if head and face, chest, abdomen and pelvis are excluded, AAPT prevalence is approximately the same as prevalence according to the FM 2016 (3.4% [95% CI 2.7, 4.3]).

The FM 2016 diagnostic criteria are closer to the definition of FM according to the International Statistical Classification of Diseases and Related Health Problems, Eleventh Revision: "Fibromyalgia syndrome (FMS) is a form of chronic widespread pain, which is defined as pain in at least 4 of 5 body regions (in at least 3 or 4 body quadrants), and is associated with sleep disorders, cognitive dysfunction, and somatic symptoms. The symptoms have been present at a similar level for at least 3 months and are not better accounted for by another diagnosis" (29).

Finally, the AAPT definition of chronic widespread pain is unusual and different from all previous definitions (30,31). What the results of this study may be showing is that the definition of FM is inexact given the many sets of previous criteria and the discordance between the FM 2016 and AAPT criteria. There are no clear boundaries between what is and what is not FM

Table 4. Possible effect of altering AAPT criteria definitions*

Category	Alteration	AAPT, unaltered	AAPT, altered	FM 2016, unaltered
Prevalence	None	5.7 (4.8, 6.8)		3.4 (2.7, 4.3)
Somatic symptoms	Use FM 2016 SSS score ≥5, not AAPT ≥ moderate fatigue or sleep problems	5.7 (4.7, 6.8)	3.2 (2.7, 3.8)	
Pain sites	Do not include nonmusculoskeletal regions in AAPT site score†	5.7 (4.7, 6.8)	3.5 (2.5, 5.4)	

* Values are the % (95% confidence interval). AAPT = Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks–American Pain Society pain taxonomy criteria; FM 2016 = modification of the American College of Rheumatology criteria for FM; SSS = Symptom Severity Score.
† Head/face, chest, or abdomen/pelvic pain.

(32). A number of studies have found substantial discordance between patient-reported diagnosis of FM and assessment using FM criteria (33–35). While there are many potential reasons for such discordance, uncertainty related to criteria may be one contributing factor. Diagnosis of FM in clinical practice is associated with sex and expectation bias, and is susceptible to social pressures (35,36), including a tendency to push the bounds of criteria. The simultaneous use of diagnostic criteria and the PSD Scale, however, even in broad categories (37), can ameliorate problems caused by dichotomization and clinical pressures by documenting FM severity.

There are a number of limitations to our study. We followed the directions that the AAPT group outlined in their pain map, but we used the MBM for the AAPT and the FM 2016 criteria. We used similar variables for the FM 2016 symptom scales but not the wording used in the 2016 criteria study (5). These differences could have influenced the prevalence estimates. Given the difficulty of the various and variously applied criteria (38,39), it seems clear that epidemiologic estimates of FM will always be uncertain. The strength of the current study comes from the high-quality effort in data-gathering and the use of standard definitions.

In conclusion, we have shown that the prevalence of FM is 78% greater using the AAPT criteria compared with using the FM 2016 criteria. Additional studies are needed to replicate these results in clinical settings, where an increased prevalence of severe symptoms will be observed. We believe that a medical report with the diagnosis of FM should include a description of the criteria used for diagnosis. The simultaneous use of a PSD Scale measure, which can be part of the criteria questionnaire, will help to interpret the criteria results. A system of PSD Scale categories gives legitimacy to patient's symptoms while keeping the diagnosis intact. Such a system would work with the AAPT and FM 2016 criteria. For example, if there were categories of none (0–3), mild (4–7), moderate (8–11), severe (12–19), and very severe (20–31), then differences in criteria results, which might be the result of different criteria sensitivities, could be understood in terms of the common PSD Scale level. Doing this obviates borderline diagnosis error, allows for improvement and flares to be recognized, provides a way to conceptualize subsyndromal FM, and brings more science into diagnosis and management.

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. Wolfe 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. Häuser, Brähler, Wolfe.
Acquisition of data. Häuser, Brähler, Wolfe.
Analysis and interpretation of data. Häuser, Ablin, Wolfe.

REFERENCES

1. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: report of the multicenter criteria committee. Arthritis Rheum 1990;33:160–72.
2. Butler S, Landmark T, Glette M, Borchgrevink P, Woodhouse A. Chronic widespread pain: the need for a standard definition. Pain 2016;157:541–3.
3. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10.
4. Wolfe F, Clauw D, Fitzcharles MA, Goldenberg D, Häuser W, Katz RS, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR preliminary diagnostic criteria for fibromyalgia. J Rheumatol 2011;38:1113–22.
5. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RL, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–29.
6. Arnold LM, Bennett RM, Crofford LJ, Dean LE, Clauw DJ, Goldenberg DL, et al. AAPT diagnostic criteria for fibromyalgia. J Pain 2019;20:611–28.
7. Heidari F, Afshari M, Moosazadeh M. Prevalence of fibromyalgia in general population and patients, a systematic review and meta-analysis. Rheumatol Int 2017;37:1527–39.
8. Brummett CM, Bakshi RR, Goesling J, Leung D, Moser SE, Zollars JW, et al. Preliminary validation of the Michigan Body Map (MBM). Pain 2016;157:1205.
9. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64:258–66.
10. Gierk B, Kohlmann S, Kroenke K, Spangenberg L, Zenger M, Brähler E, et al. The Somatic Symptom Scale–8 (SSS-8): a brief measure of somatic symptom burden. JAMA Intern Med 2014:399–407.
11. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ–4. Psychosomatics 2009;50:613–21.
12. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. J Clin Epidemiol 1998;51:1171–8.
13. Bullinger M, Morfeld M, Kohlmann T, Nantke J, Van den Bussche H, Dodt B, et al. Der SF-36 in der Rehabilitationswissenschaftlichen Forschung-Ergebnisse aus dem Norddeutschen Verbund für Rehabilitationsforschung (NVRF) im Förderschwerpunkt Rehabilitationswissenschaften. Die Rehabilitation 2003;42:218–25.
14. 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.
15. Streibelt M, Schmidt C, Brünger M, Spyra K. Comorbidity from the patient perspective-does it work? Validity of a questionnaire on self-estimation of comorbidity (SCQ-D). Der Orthopade 2012;41:303–10. In German.
16. Budtz-Lilly A, Fink P, Ørnbøl E, Vestergaard M, Moth G, Christensen KS, et al. A new questionnaire to identify bodily distress in primary care: the 'BDS checklist'. J Psychosom Res 2015;78:536–45.
17. Wolfe F, Michaud K. The National Data Bank for rheumatic diseases: a multi-registry rheumatic disease data bank. Rheumatology (Oxford) 2011;50:16–24.
18. Stata. Stata statistical software: release 15.0. College Station (TX): StataCorp; 2018.

19. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22:276–82.
20. Wolfe F, Brähler E, Hinz A, Häuser W. Fibromyalgia prevalence, somatic symptom reporting, and the dimensionality of polysymptomatic distress: results from a survey of the general population. Arthritis Care Res (Hoboken) 2013;65:777–85.
21. Wolfe F, Walitt B. Culture, science and the changing nature of fibromyalgia. Nat Rev Rheumatol 2013;9:751.
22. Klaver-Krol E, Rasker JJ, Klaver M, Ten Klooster PM, Zwarts M. Fibromyalgia: increased reactivity of the muscle membrane and a role of central regulation. Clin Neurophysiol 2019;130:12–9.
23. Yunus M, Masi AT, Calabro JJ, Miller KA, Feigenbaum SL. Primary fibromyalgia (fibrositis): clinical study of 50 patients with matched normal controls. Semin Arthritis Rheum 1981;11:151–71.
24. Gracely RH, Grant MA, Giesecke T. Evoked pain measures in fibromyalgia. Best Pract Res Clin Rheumatol 2003;17:593–609.
25. Wolfe F. The relation between tender points and fibromyalgia symptom variables: evidence that fibromyalgia is not a discrete disorder in the clinic. Ann Rheum Dis 1997;56:268–71.
26. Walitt B, Fitzcharles MA, Hassett AL, Katz RS, Häuser W, Wolfe F. The longitudinal outcome of fibromyalgia: a study of 1555 patients. J Rheumatol 2011;38:2238–46.
27. Nimnuan C, Rabe-Hesketh S, Wessely S, Hotopf M. How many functional somatic syndromes? J Psychosom Res 2001;51:549–57.
28. Henningsen P, Zipfel S, Sattel H, Creed F. Management of functional somatic syndromes and bodily distress. Psychother Psychosom 2018;87:12–31.
29. Nicholas M, Vlaeyen JW, Rief W, Barke A, Aziz Q, Benoliel R, et al. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019;160:28–37.
30. McBeth J, Jones K. Epidemiology of chronic musculoskeletal pain. Best Pract Res Clin Rheumatol 2007;21:403–25.
31. Macfarlane GJ. Generalized pain, fibromyalgia and regional pain: an epidemiological view. Baillieres Best Pract Res Clin Rheumatol 1999;13:403–14.
32. Wolfe F. Criteria for fibromyalgia? What is fibromyalgia? Limitations to current concepts about fibromyalgia and fibromyalgia criteria [editorial]. Clin Exp Rheumatol 2017;35 Suppl 105:3–5.
33. Srinivasan S, Maloney E, Wright B, Kennedy M, Kallail KJ, Rasker JJ, et al. The problematic nature of fibromyalgia diagnosis in the community. ACR Open Rheumatol 2019;1:43–51.
34. Wolfe F, Schmukler J, Jamal S, Castrejon I, Gibson KA, Srinivasan S, et al. Diagnosis of fibromyalgia: disagreement between fibromyalgia criteria and clinician-based fibromyalgia diagnosis in a university clinic. Arthritis Care Res (Hoboken) 2019;71:343–51.
35. Walitt B, Katz RS, Bergman MJ, Wolfe F. Three-quarters of individuals in the US population reporting a clinical diagnosis of fibromyalgia do not satisfy fibromyalgia criteria: the 2012 National Health Interview Survey. PLoS One 2016;11:e0157235.
36. Wolfe F, Walitt B, Perrot S, Rasker JJ, Häuser W. Fibromyalgia diagnosis and biased assessment: sex, prevalence and bias. PLoS One 2018;13:e0203755.
37. Wolfe F, Walitt BT, Rasker JJ, Katz RS, Häuser W. The use of polysymptomatic distress categories in the evaluation of fibromyalgia (FM) and FM severity. J Rheumatol 2015;42:1494–501.
38. Queiroz LP. Worldwide epidemiology of fibromyalgia. Curr Pain Headache Rep 2013;17:356.
39. Mansfield KE, Sim J, Jordan JL, Jordan KP. A systematic review and meta-analysis of the prevalence of chronic widespread pain in the general population. Pain 2016;157:55.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 626–632
DOI 10.1002/acr.24170

Defining Pain That Does Not Interfere With Activities Among Rheumatoid Arthritis Patients

Yvonne C. Lee,[1] Patricia Katz,[2] Amanda Quebe,[3] Luna Sun,[3] Himanshu Patel,[3] Carol L. Gaich,[3] and Kaleb Michaud[4]

Objective. The objectives of this study were to: 1) characterize the distribution of noninterfering pain (defined as the pain intensity level at which individuals can function without interference) across different aspects of life among patients with rheumatoid arthritis (RA), and 2) identify clinical characteristics associated with differing levels of noninterfering pain.

Methods. Patients with RA in FORWARD, The National Databank for Rheumatic Diseases completed 8 items from the Patient-Reported Outcomes Measurement Information System (PROMIS) pain interference item bank that asked about interference with activities. If subjects reported pain interference, they were asked, "At what level would pain no longer interfere with this activity?" on a scale of 0 to 10. Subjects were also asked, "At what level of pain would you be able to do everything you want to do?" Multiple linear regression analyses examined associations between clinical characteristics and noninterfering pain.

Results. A total of 3,949 patients with RA completed the questionnaires. Pain interference was most common for daily activities and least common for ability to concentrate. The mean ± SD level at which pain no longer interfered with activities ranged from 2.7 ± 2.1 for ability to fall/stay asleep to 3.1 ± 2.0 for social activities. Overall, the mean ± SD threshold for noninterfering pain was 2.8 ± 1.9. The mean ± SD level of pain at which patients could do everything they wanted to do was 2.3 ± 1.9. More severe pain intensity was associated with higher noninterfering pain.

Conclusion. The mean pain level that did not interfere with activities was 3. High pain intensity was associated with high self-reported noninterfering pain.

INTRODUCTION

Patients with rheumatoid arthritis (RA) identify pain as one of the most important factors that affect their quality of life (1–3). Despite advances in RA treatment that have significantly improved disease control at a population level, individual patients often continue to experience significant pain (4–6). This pain differs from person to person and is multidimensional, encompassing sensory, emotional, and functional aspects (7,8).

Pain interference is defined as the effects of pain on engagement with the physical, cognitive, emotional, and social aspects of life (9). Pain intensity and pain interference are strongly correlated but not synonymous (10). Factors such as age, stress, and pain catastrophizing may all impact pain interference, independent of pain intensity (11–13). These observations suggest that there is heterogeneity in the amount of pain that individuals can endure without it interfering with daily activities, recreational activities, and mental health. While some patients appear quite stoic, continuing to manage household chores, work, and engage in social activities with a pain intensity of 5 of 10, others are incapacitated. In addition, there may be variability in the amount of pain that can be endured with respect to different aspects of life. For example, while a pain intensity of 5 of 10 may impact enjoyment of life, it may not affect an individual's ability to perform day-to-day activities.

While the ultimate goal may be to eliminate pain, experts have indicated that this end point may not be realistic (14). Pain in RA is multifactorial in nature, stemming from the inflammatory component of the disease, but is further associated with damage, peripheral and central sensitization, and psychosocial factors

Supported by Eli Lilly and Company.
[1]Yvonne C. Lee, MD, MMSc: Northwestern University Feinberg School of Medicine, Chicago, Illinois; [2]Patricia Katz, PhD: University of California, San Francisco; [3]Amanda Quebe, PharmD, Luna Sun, PhD, Himanshu Patel, DO, Carol L. Gaich, PharmD: Eli Lilly and Company, Indianapolis, Indiana; [4]Kaleb Michaud, PhD: FORWARD, The National Databank for Rheumatic Diseases, Wichita, Kansas, and University of Nebraska Medical Center, Omaha.

Dr. Lee has received research support from Pfizer. Drs. Quebe, Sun, Patel, and Gaich own stock or stock options in Eli Lilly and Company. Dr. Michaud has received research support from Pfizer. No other disclosures relevant to this article were reported.

Address correspondence to Yvonne C. Lee, MD, MMSc, 633 North St. Clair Street, 18-093, Chicago, IL 60611. Email: yvonne.lee@northwestern.edu.

Submitted for publication June 19, 2019; accepted in revised form February 11, 2020.

SIGNIFICANCE & INNOVATIONS

- This is the first study to assess the level of pain that patients with rheumatoid arthritis (RA) can tolerate without this pain interfering with specific activities.
- Results indicate that pain levels ≥3 of 10 are associated with interference with activities, regardless of the specific type of activity.
- RA patients who reported higher pain intensity also reported that they could work through higher levels of pain.

(15–19). Thus, control of the disease process, which is primarily aimed at control of inflammation, may not always result in resolution of pain (18,19). Reduction of pain to a level at which it does not interfere may be more achievable. For this reason, it is important to define the level of pain that is noninterfering and to identify characteristics that may be associated with a shift in pain threshold, enabling individuals to maximize function and quality of life in spite of pain. Therefore, the objectives of this study were to: 1) characterize the distribution of noninterfering pain (defined as the pain intensity level at which individuals can function without interference) across different aspects of life among RA patients in a large US-wide registry, and 2) identify clinical characteristics associated with differing levels of noninterfering pain.

PATIENTS AND METHODS

Sample. Data were derived from patients with rheumatologist-confirmed RA in FORWARD, The National Databank for Rheumatic Diseases (15,16). Recruitment for this US-based, observational, longitudinal study was mainly through rheumatologists. Self-reported data were collected every 6 months, using standardized questionnaires in the form of paper questionnaires sent through the mail, electronically via a secure online portal, or via a telephone interview. All study participants provided written informed consent. The Via Christi Institutional Review Board approved all study procedures. This analysis included 3,949 participants with RA who had completed the January 2018 comprehensive questionnaires by June 2018. To approximate the response rate, we used data from individuals who responded to the previous (July 2017) questionnaire as the denominator, and we calculated the response rate as the proportion of that group who also responded to the January 2018 questionnaire. All clinical variables were assessed at the same time as the data on noninterfering pain.

Measures. *Demographic characteristics and pain interference and noninterfering pain levels.* Age, sex, and RA duration were obtained via self-report. Subjects completed 8 items from the Patient-Reported Outcomes Measurement Information System (PROMIS) pain interference item bank that asked the degree to which pain interfered with activities (from "not at all" to "very much") (17). These 8 items included the 4 items in the static short form, as well as 4 additional items chosen to represent a range of activities (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24170/abstract). If subjects reported any pain interference, they were asked, "At what level would pain no longer interfere with this activity?" To assess participants' overall sense of noninterfering pain, participants were also asked, "At what level of pain would you be able to do everything you want to do?" Responses were collected using a 0–10 Numerical Rating Scale (NRS; 10 = severe pain).

Other patient-reported outcomes. Several other patient-reported outcomes were examined as potential characteristics associated with noninterfering pain. These included pain intensity, fatigue, and sleep problems, which were all measured using a single-item NRS with scores ranging from 0 (no pain/fatigue/sleep disturbance) to 10 (extreme pain/fatigue/sleep disturbance). Physical function was assessed with the Health Assessment Questionnaire II (HAQ-II); higher scores indicate worse function (18). Comorbidities were tallied using the Rheumatic Disease Comorbidity Index (19), and depression was evaluated using the Patient Health Questionnaire 8 (20). The Polysymptomatic Distress Scale, which includes the Widespread Pain Index and Symptom Severity Score from the American College of Rheumatology 2010 fibromyalgia criteria, was used to measure fibromyalgia symptoms as a proxy for centralized pain (21,22).

Statistical analysis. Descriptive statistics (e.g., mean ± SD and number [%]) were calculated for the total study sample. Histograms were plotted to visually display the distribution of variables. The percentage of patients who endorsed pain interference with each activity was calculated by dividing the number of patients who reported "a little bit," "somewhat," "quite a bit," or "very much" pain interference by the total number of patients who answered that question. Mean ± SDs and medians (interquartile ranges) were calculated to summarize the distribution of answers to each question about the pain level that would not interfere with each activity.

To determine if an overall mean of noninterfering pain could be calculated, internal consistency was assessed using Cronbach's alpha for all items and with each item deleted. Item-total correlations were also examined. For participants who answered at least 6 of 8 questions about noninterfering pain, the mean of all answered items was calculated to provide a summary measure of noninterfering pain.

Univariate and multiple linear regression analyses were used to examine associations between noninterfering pain and the following characteristics: age, sex, disease duration, body mass index, Rheumatic Disease Comorbidity Index, disability (using the HAQ-II), pain intensity NRS, fatigue NRS, self-reported

Table 1. Baseline cohort characteristics*

Clinical characteristics†	Value
Age, years	65.4 ± 11.9
Female sex, no. (%)	3,237 (83)
Obese, no. (%)	1,408 (36)
RA duration, years	21.7 ± 12.6
Physical function score, HAQ-II (range 0–3)	0.9 ± 0.7
Pain intensity score, NRS (range 0–10)	3.4 ± 2.6
Fatigue score, NRS (range 0–10)	3.8 ± 2.9
Sleep problems score, NRS (range 0–10)	3.6 ± 3.0
PSD score (range 0–31)	8.2 ± 6.8
Depression score, PHQ-8 (range 0–24)	4.9 ± 4.9

* Values are the mean ± SD unless indicated otherwise. HAQ-II = Health Assessment Questionnaire II; NRS = Numerical Rating Scale; PHQ-8 = Patient Health Questionnaire 8; PSD = Polysymptomatic Distress Scale; RA = rheumatoid arthritis.
† The total number of patients with data on each question was as follows: 3,949 for age; 3,877 for female sex; 3,875 for obese; 3,723 for RA duration; 3,805 for physical function; 3,934 for pain intensity; 3,908 for fatigue; 3,456 for sleep problems; 2,126 for PSD score; and 3,864 for depression.

current depression, sleep problems NRS, and fibromyalgia symptom severity (using the Polysymptomatic Distress Scale). Two adjusted models were evaluated: 1) age- and sex-adjusted models, and 2) fully adjusted models including all of the above variables. Prior to inclusion in the fully adjusted models, collinearity diagnostics were run to ensure that all tolerance estimates were above 0.1, and all variance inflation factors were below 3. All analyses used SAS, version 9.4.

RESULTS

As of June 2018, 3,949 RA patients had completed the questionnaires on pain interference and noninterfering pain (Table 1). The approximate response rate was 82%. The average age was mean ± SD 65.4 ± 11.9 years, and 3,237 patients (83.5%) were female. The average RA duration was mean ± SD 21.7 ± 12.6 years.

PROMIS pain interference scores covered the entire potential score range (from 1 "not at all" to 5 "very much") for each item. Distributions were right-skewed, with the largest proportion of patients indicating that pain interfered "not at all" or "a little bit" with each specific activity. Pain interference was most common for daily activities (76.2%) and least common for ability to concentrate (51.3%) (Figure 1 and Table 2).

Noninterfering pain scores covered the entire potential score range (from 0 to 10) for each item. Participants indicated that the level of pain at which they would be able to do everything that they wanted to do was mean ± SD 2.3 ± 1.9. The mean ± SD level at which pain no longer interfered with specific activities ranged from 2.7 ± 2.1 for the ability to fall/stay asleep to 3.1 ± 2.0 for social activities. Distributions were right-skewed (Figure 2). The percentage of patients who reported that pain level had to be 0 in order for it to not interfere with specific activities ranged from 6.0% for work around the house up to 10.8% for the ability to fall asleep/stay asleep. On the opposite side of the spectrum, the percentage of patients who reported that pain level would have to be ≥8 before interfering with activities ranged from 3.4% for the ability to participate in social activities to 5.5% for ability to work.

Internal consistency across items for noninterfering pain for specific activities was high (Cronbach's α = 0.97). Estimates for Cronbach's alpha with each item deleted were all <0.97, indicating that no item should be dropped from the overall calculation. The correlation of each item with the total score was between 0.82 and 0.92, indicating that the items were measuring the same concept. Overall, the mean ± SD threshold for noninterfering pain, averaged across all 8 activities, was 2.8 ± 1.9.

Multiple linear regression models were constructed to assess the association between clinical characteristics and noninterfering pain. Collinearity diagnostics indicated no collinearity. All tolerance estimates were above 0.1, and all variance inflation factors were below 3. Of the 11 potential predictors, only more intense pain

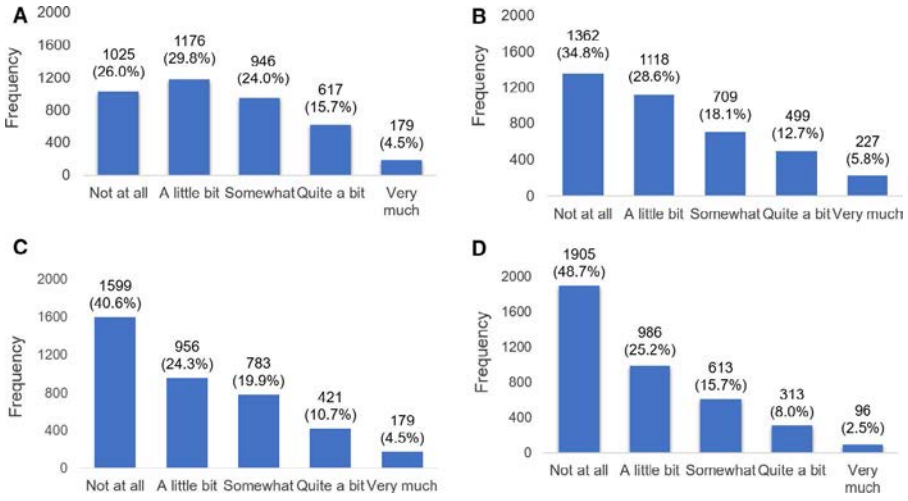


Figure 1. Distributions of responses for selected Patient-Reported Outcomes Measurement Information System (PROMIS) items on the amount of pain interference for work around the home (**A**), fall sleep/stay asleep (**B**), social activities (**C**), and concentrate (**D**).

Table 2. The proportion of patients with rheumatoid arthritis who experience pain interference and the pain level at which pain does not interfere with activities*

Activity†	% who experienced pain interference‡	Pain level that would not interfere with activity	
		Mean ± SD	Median (IQR)
Daily activities	76.2	3.0 ± 2.0	3 (2–4)
Work around the house	74.0	3.0 ± 2.0	3 (2–4)
Household chores	71.1	3.0 ± 2.1	2 (2–4)
Fall asleep/stay asleep	65.2	2.7 ± 2.1	2 (1–4)
Enjoyment of life	63.2	2.9 ± 2.2	2 (1–4)
Ability to work	61.7	3.0 ± 2.3	3 (1–4)
Social activities	59.4	3.1 ± 2.0	3 (2–4)
Ability to concentrate	51.3	2.8 ± 2.1	2 (1–4)

* IQR = interquartile range; PROMIS = Patient-Reported Outcomes Measurement Information System.
† The total number of patients who responded to each question was as follows: 3,943 for work around the house; 3,940 for daily activities; 3,938 for household chores and social activities; 3,915 for fall asleep/stay asleep; 3,911 for enjoyment of life; 3,913 for ability to concentrate; and 1,913 for ability to work. A total of 1,982 subjects did not answer the question regarding ability to work because they were retired or otherwise not working.
‡ According to PROMIS pain interference questions. The percentage of patients who endorsed pain interference with each activity was calculated by dividing the number of patients who reported "a little bit," "somewhat," "quite a bit," or "very much" pain interference by the total number of patients who answered that question.

was significantly associated with higher noninterfering pain levels ($\beta = 0.20$, $P < 0.0001$) (Table 3). In other words, when all other characteristics were held the same, patients with 1-unit higher pain intensity reported 0.2 units greater noninterfering pain level. A larger proportion of patients with high pain intensity reported higher noninterfering pain levels (Figure 3). Other characteristics tested were not statistically significant at $\alpha = 0.05$.

DISCUSSION

Our results show that pain has a significant impact on patients' lives with more than one-half of RA patients (51.3%) reporting that pain interferes at least a little bit with their ability to concentrate, and 76.2% reporting that pain interferes at least a little bit with daily activities. Participants reported an average pain intensity of 3.4 on a 0–10 scale, while also reporting that average pain intensity would have to be <3 in order for pain not to interfere with activities, and closer to 2 in order to enable them to do everything that they wanted to do.

Several studies have reported that pain and function are the 2 most important priorities to RA patients (23–25). In a study examining the feasibility of using PROMIS domains as patient-identified treatment targets (1), 37% selected pain as the domain of highest priority. Patients who chose pain as their highest priority were most interested in tracking their ability to complete activities of daily living, such as laundry or grocery shopping. Similarly, our study identified daily activities, work around the house, and household chores as the activities most commonly affected by pain. Together, these results emphasize the strong effect that pain exerts on common activities

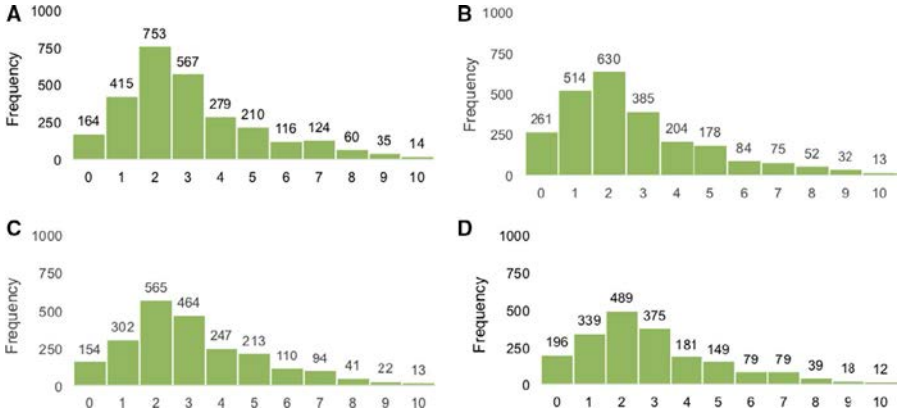


Figure 2. Distributions of responses to the question, "If pain interferes at all, at what level would pain no longer interfere with this activity?" for work around the home (**A**), fall sleep/stay asleep (**B**), social activities (**C**), and concentrate (**D**).

Table 3. Predictors of noninterfering pain level expressed as β coefficients and *P* values*

Clinical characteristic	Unadjusted		Age and sex adjusted		Fully adjusted	
	β	*P*	β	*P*	β	*P*
Age, years	−0.007	0.01	−0.006	0.03	−0.002	0.56
Female sex	0.05	0.59	0.02	0.80	−0.06	0.62
Obese (BMI ≥30 kg/m^2)	0.30	<0.0001	0.27	0.0001	−0.02	0.81
RA duration, years	0.004	0.11	0.006	0.03	0.004	0.23
Physical function score, HAQ-II (range 0–3)	0.67	<0.0001	0.70	<0.0001	0.03	0.72
Pain intensity score, NRS (range 0–10)	0.26	<0.0001	0.26	<0.0001	0.20	<0.0001
Fatigue score, NRS (range 0–10)	0.18	<0.0001	0.18	<0.0001	0.02	0.48
Sleep problems score, NRS (range 0–10)	0.13	<0.0001	0.13	<0.0001	0.01	0.60
PSD score (range 0–31)	0.07	<0.0001	0.06	<0.0001	0.01	0.27
Depression score, PHQ-8 (range 0–24)	0.08	<0.0001	0.08	<0.0001	−0.006	0.65
Rheumatic Disease Comorbidity Index (range 0–9)	0.13	<0.0001	0.14	<0.0001	0.02	0.40

* Higher scores on all scales indicate worse outcomes. BMI = body mass index; HAQ-II = Health Assessment Questionnaire II; NRS = Numerical Rating Scale; PHQ-8 = Patient Health Questionnaire 8; PSD = Polysymptomatic Distress Scale; RA = rheumatoid arthritis.

necessary for daily living and highlight the importance of managing pain to improve patients' functionality as a part of their normal, everyday routine.

While pain most frequently interfered with daily activities, falling asleep/staying asleep was the activity associated with the lowest noninterfering pain level. In other words, RA patients perceived falling asleep/staying asleep to be most sensitive to pain. A study examining the relationship between sleep problems and disease activity in RA reported that insomnia and poor sleep quality were positively correlated with RA disease activity, whereas daytime sleepiness was inversely correlated with RA disease activity, possibly due to pain-related alertness (26). Other cross-sectional studies have identified sleep problems as a potential cause for enhanced pain sensitivity and fatigue (27–29), and a laboratory-based study demonstrated that partial night sleep deprivation resulted in worse outcomes of fatigue, depression, and pain in RA (30). These observations highlight the reciprocal relationship between pain and sleep problems, as well as the impact of sleep problems on other important outcomes, such as fatigue and depression. Future studies are needed to determine whether treatment strategies, such as cognitive behavioral therapy, may be useful in breaking the cycle of pain begetting sleep problems, which, in turn, may lead to more pain, fatigue, and depression.

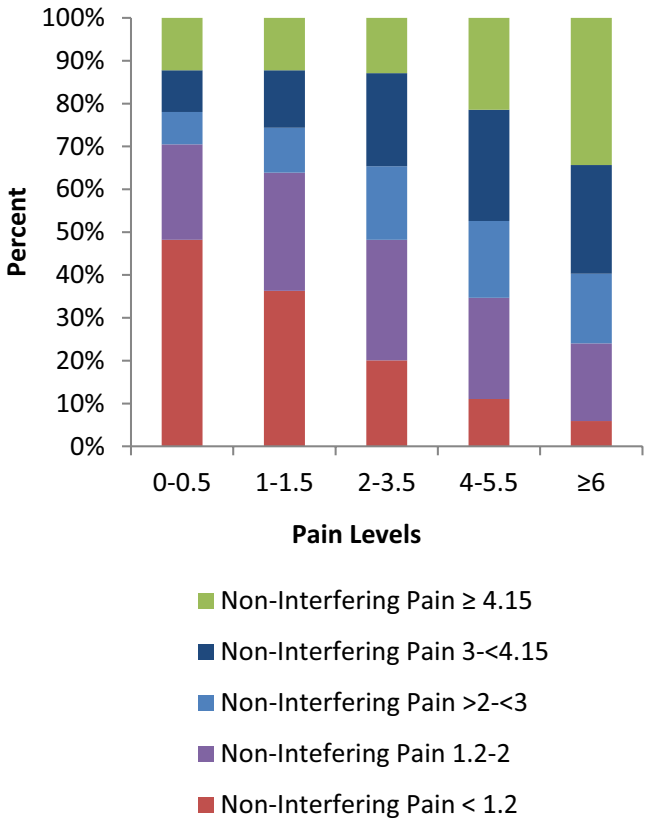


Figure 3. Percentage of patients with rheumatoid arthritis in each quintile of noninterfering pain, stratified by self-reported pain severity. Higher noninterfering pain categories indicate greater pain tolerance.

Although falling asleep/staying asleep was the activity associated with the lowest noninterfering pain level, the range between the lowest and highest noninterfering pain levels was small (mean 2.7–3.1), suggesting that, overall, the noninterfering pain level did not vary depending on activity type. Since this is the first time that noninterfering pain levels have been reported in RA, we do not have data from anchor-based methods or global ratings of change to inform what constitutes a minimum important difference (MID) in noninterfering pain. In the absence of these data, a general guideline is that the MID is approximately one-half of the SD (31). Based on the SDs for noninterfering pain in this study (Table 2), an estimate of the MID for noninterfering pain would be between 1.0–1.2. This difference is 2.5–3 times greater than the maximum difference in noninterfering pain level across different activities, supporting our clinical intuition that these differences are not clinically meaningful.

When items were combined into a single scale, the overall mean ± SD pain level that did not interfere with activities was 2.8 ± 1.9, which was slightly higher than the mean pain level of 2.3 ± 1.9 in response to the question, "At what level of pain would you be able to do everything you want to do?" This

discrepancy may be due to greater pain interference with activities that were not specifically assessed in this study. For example, we did not assess pain interference with tasks away from home (e.g., getting groceries, running errands), which was a priority item identified in a recent study of RA patients asked to select PROMIS items that addressed their treatment goals (1). Another explanation may be that RA patients have more difficulty responding to a general question about "everything you want to do," compared to questions about specific activities and functions, leading to a bias in response. Because the question is quite broad, subjects may also differ in their response depending on comorbidities, living situations, and social context.

The mean ± SD overall noninterfering pain level (2.8 ± 1.9) in this study was only slightly lower than the mean ± SD pain intensity level (3.4 ± 2.6). The magnitude of difference in these values may differ across study populations. Our study population consisted of established RA patients (mean ± SD disease duration 21.7 ± 12.6 years) with well-managed disease. It is likely that mean pain intensity would be higher in individuals with more active inflammatory disease, such as early RA patients whose treatment regimens are still in flux. Individuals may also progress through stages of reactions to chronic illness, initially experiencing shock, anxiety, and depression, before reaching a state of acknowledgement and adjustment (32,33). Based on this theory, individuals with recent-onset disease may report lower noninterfering pain than those with long disease durations because individuals with established disease have learned to adjust to their disease. Each individual's course may differ, however, because coping strategies vary and are not always adaptive (34). Further studies are needed to study the relationship between pain intensity and noninterfering pain in different study populations (e.g., early RA, active RA, socioeconomically disadvantaged patients), as well as to examine the impact of specific coping strategies on noninterfering pain in RA.

In multivariable analyses, higher pain intensity was the only characteristic independently associated with higher noninterfering pain levels. This finding may indicate that patients who routinely experience severe pain acclimate to it and are therefore able to function at higher levels of pain than those who experience less intense pain. Alternatively, the results could be the outcome of a response bias, resulting from subjects' tendencies to rate items similarly. For example, subjects who rate pain highly may also rate noninterfering pain levels highly, simply because they are high raters. The reasons underlying the association between high pain intensity and high noninterfering pain need to be further investigated prior to designing interventions to decrease the gap between pain intensity and noninterfering pain.

Limitations of this study include the absence of data on tender and swollen joint counts and acute-phase reactants, which limit our ability to evaluate the impact of inflammatory disease activity on pain interference and noninterfering pain levels. In addition, the questions regarding pain interference were prefaced by the stem, "Compared to 6 months ago." Thus, participants who had high levels of pain interference 6 months ago may have responded with lower ratings of interference than they would have without this comparison; whereas those who had no/minimal pain interference 6 months ago may have responded with higher ratings of interference than they would have without this comparison. This would have affected our results by decreasing the spread of the data. It may have also decreased the proportion of patients with pain interference if participants interpreted the "not at all" answer option to mean no change from 6 months ago, rather than no pain interference at all. Thus, our estimates of the proportion of individuals who reported pain interference are likely to be conservative estimates. Last, participants in FORWARD represent a relatively well-managed and well-educated population, resulting in a participation bias that may limit the generalizability of our results.

This study also has a number of strengths. Data were obtained from a large observational cohort of RA patients primarily recruited from routine clinical practice, enhancing the generalizability of results. The incorporation of the PROMIS pain interference items enabled a comprehensive assessment of the consequences of pain on relevant aspects of patients' lives. The addition of questions asking patients about the level of pain that would no longer interfere with specific activities, linked to the PROMIS pain interference items, enabled a novel evaluation of how patients differ in their ability to work through pain.

In conclusion, more than one-half of RA patients reported that pain interfered with function, most commonly affecting daily activities. On average, pain levels had to be <3 of 10 in order to not interfere with function. These data highlight the importance of effective pain management, with a goal of reducing pain levels to <3 of 10. Future studies are also needed to identify interventions such as cognitive behavioral therapy or mindfulness-based strategies that can increase this threshold, enabling RA patients to augment quality of life and function, notwithstanding pain.

ACKNOWLEDGMENTS

We acknowledge the participants and staff of FORWARD, The National Databank for Rheumatic Diseases.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Lee, Katz, Quebe, Sun, Patel, Gaich, Michaud.

Acquisition of data. Katz, Michaud.

Analysis and interpretation of data. Lee, Katz, Quebe, Sun, Patel, Gaich, Michaud.

ROLE OF THE STUDY SPONSOR

Employees of Eli Lilly and Company were involved in the interpretation of data, drafting the article, and revising it critically for intellectual content. They also approved the final version to be published.

REFERENCES

1. Bacalao EJ, Greene GJ, Beaumont JL, Eisenstein A, Muftic A, Mandelin AM, et al. Standardizing and personalizing the treat to target (T2T) approach for rheumatoid arthritis using the Patient-Reported Outcomes Measurement Information System (PROMIS): baseline findings on patient-centered treatment priorities. Clin Rheumatol 2017;36:1729–36.
2. Heiberg T, Finset A, Uhlig T, Kvien TK. Seven year changes in health status and priorities for improvement of health in patients with rheumatoid arthritis. Ann Rheum Dis 2005;64:191–5.
3. Ten Klooster PM, Veehof MM, Taal E, van Riel PL, van de Laar MA. Changes in priorities for improvement in patients with rheumatoid arthritis during 1 year of anti-tumour necrosis factor treatment. Ann Rheum Dis 2007;66:1485–90.
4. Lee YC, Cui J, Lu B, Frits ML, Iannaccone CK, Shadick NA, et al. Pain persists in DAS28 rheumatoid arthritis remission but not in ACR/EULAR remission: a longitudinal observational study. Arthritis Res Ther 2011;13:R83.
5. Taylor PC, Moore A, Vasilescu R, Alvir J, Tarallo M. A structured literature review of the burden of illness and unmet needs in patients with rheumatoid arthritis: a current perspective. Rheumatol Int 2016;36:685–95.
6. Altawil R, Saevarsdottir S, Wedrén S, Alfredsson L, Klareskog L, Lampa J. Remaining pain in early rheumatoid arthritis patients treated with methotrexate. Arthritis Care Res (Hoboken) 2016;68:1061–8.
7. Anderson DL. Development of an instrument to measure pain in rheumatoid arthritis: Rheumatoid Arthritis Pain Scale (RAPS). Arthritis Rheum 2001;45:317–23.
8. De Luca K, Parkinson L, Downie A, Blyth F, Byles J. Three subgroups of pain profiles identified in 227 women with arthritis: a latent class analysis. Clin Rheumatol 2017;36:625–34.
9. Riley WT, Rothrock N, Bruce B, Christodolou C, Cook K, Hahn EA, et al. Patient-reported outcomes measurement information system (PROMIS) domain names and definitions revisions: further evaluation of content validity in IRT-derived item banks. Qual Life Res 2010;19:1311–21.
10. Fayers PM, Hjermstad MJ, Klepstad P, Loge JH, Caraceni A, Hanks GW, et al. The dimensionality of pain: palliative care and chronic pain patients differ in their reports of pain intensity and pain interference. Pain 2011;152:1608–20.
11. Thomas JG, Pavlovic J, Lipton RB, Roth J, Rathier L, O'Leary KC, et al. Ecological momentary assessment of the relationship between headache pain intensity and pain interference in women with migraine and obesity. Cephalalgia 2016;36:1228–37.
12. Covic T, Adamson B, Hough M. The impact of passive coping on rheumatoid arthritis pain. Rheumatology (Oxford) 2000; 39:1027–30.
13. Thastum M, Herlin T, Zachariae R. Relationship of pain-coping strategies and pain-specific beliefs to pain experience in children with juvenile idiopathic arthritis. Arthritis Rheum 2005;53:178–84.
14. Turk DC, Swanson KS, Tunks ER. Psychological approaches in the treatment of chronic pain patients: when pills, scalpels, and needles are not enough. Can J Psychiatry 2008;53:213–23.
15. Michaud K. The National Data Bank for Rheumatic Diseases (NDB). Clin Exp Rheumatol 2016;34 Suppl 101:S100–1.
16. Wolfe F, Michaud K. The National Data Bank for rheumatic diseases: a multi-registry rheumatic disease data bank. Rheumatology (Oxford) 2011;50:16–24.
17. Amtmann D, Cook KF, Jensen MP, Chen WH, Choi S, Revicki D, et al. Development of a PROMIS item bank to measure pain interference. Pain 2010;150:173–82.
18. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137–45.
19. England BR, Sayles H, Mikuls TR, Johnson DS, Michaud K. Validation of the Rheumatic Disease Comorbidity Index. Arthritis Care Res (Hoboken) 2015;67:865–72.
20. 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.
21. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RS, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR Preliminary Diagnostic Criteria for Fibromyalgia. J Rheumatol 2011;38: 1113–22.
22. Wolfe F, Brähler E, Hinz A, Häuser W. Fibromyalgia prevalence, somatic symptom reporting, and the dimensionality of polysymptomatic distress: results from a survey of the general population. Arthritis Care Res (Hoboken) 2013;65:777–85.
23. Heiberg T, Kvien TK. Preferences for improved health examined in 1,024 patients with rheumatoid arthritis: pain has highest priority. Arthritis Rheum 2002;47:391–7.
24. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S. Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. Arthritis Care Res (Hoboken) 2010;62:647–56.
25. Van Tuyl LH, Sadlonova M, Hewlett S, Davis B, Flurey C, Goel N, et al. The patient perspective on absence of disease activity in rheumatoid arthritis: a survey to identify key domains of patient-perceived remission. Ann Rheum Dis 2017;76:855–61.
26. Westhovens R, Van der Elst K, Matthys A, Tran M, Gilloteau I. Sleep problems in patients with rheumatoid arthritis. J Rheumatol 2014;41:31–40.
27. Lee YC, Lu B, Edwards RR, Wasan AD, Nassikas NJ, Clauw DJ, et al. The role of sleep problems in central pain processing in rheumatoid arthritis. Arthritis Rheum 2013;65:59–68.
28. Katz P, Margaretten M, Trupin L, Schmajuk G, Yazdany J, Yelin E. Role of sleep disturbance, depression, obesity, and physical inactivity in fatigue in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:81–90.
29. Ulus Y, Akyol Y, Tander B, Durmus D, Bilgici A, Kuru O. Sleep quality in fibromyalgia and rheumatoid arthritis: associations with pain, fatigue, depression, and disease activity. Clin Exp Rheumatol 2011;29 Suppl 69:S92–6.
30. Irwin MR, Olmstead R, Carrillo C, Sadeghi N, Fitzgerald JD, Ranganath VK, et al. Sleep loss exacerbates fatigue, depression, and pain in rheumatoid arthritis. Sleep 2012;35:537–43.
31. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41:582–92.
32. Treharne GJ, Lyons AC, Booth DA, Mason SR, Kitas GD. Reactions to disability in patients with early versus established rheumatoid arthritis. Scand J Rheumatol 2004;33:30–8.
33. Antonak RF, Livneh H. A hierarchy of reactions to disability. Int J Rehabil Res 1991;14:13–24.
34. Hamilton NA, Zautra AJ, Reich J. Individual differences in emotional processing and reactivity to pain among older women with rheumatoid arthritis. Clin J Pain 2007;23:165–72.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 633–639
DOI 10.1002/acr.24154

American College of Rheumatology
Empowering Rheumatology Professionals

Predictors of Smoking Cessation in Patients With Rheumatoid Arthritis in Two Cohorts: Most Predictive Health Care Factors

Maria Schletzbaum,[1] Xing Wang,[1] Robert Greenlee,[2] Megan E. Piper,[1] and Christie M. Bartels[1]

Objective. Recognizing smoking as a risk factor for rheumatoid arthritis (RA) severity, the present study was undertaken to evaluate patient- and health care–level predictors of smoking cessation in patients with RA to guide implementation of smoking cessation interventions.

Methods. Electronic health record data from 2 health systems were abstracted for patients with at least 2 International Classification of Disease diagnosis codes for RA between 2005 and 2016. Patients missing smoking statuses or with <6 months of follow-up were excluded. Multivariable logistic regression was used to determine predictors of smoking cessation.

Results. Among 3,577 patients with RA, 507 smoked at baseline, and 29% quit over a median of 4.75 years. Black male patients, ages 40–59 years and enrolled in Medicaid, were significantly more likely to be baseline smokers; however, none of these factors predicted cessation. Instead, patients new to rheumatology care were 60% more likely to quit (adjusted odds ratio [OR_{adj}] 1.60 [95% confidence interval (95% CI) 1.02–2.50]), and patients in the rural community health system were 66% more likely to quit (OR_{adj} 1.66 [95% CI 1.03–2.69]). Seropositive patients were 43% less likely to quit smoking (OR_{adj} 0.57 [95% CI 0.35–0.91]).

Conclusion. Health care factors, including health system and being new to rheumatology care, were more predictive of smoking cessation in patients with RA than patient sociodemographic factors, suggesting an important role for health system cessation efforts for patients with RA. Seropositive patients were less likely to quit and may particularly benefit from cessation support. Emphasizing smoking cessation with new or seropositive RA patients and leveraging health system interventions could improve smoking cessation and outcomes in RA.

INTRODUCTION

Smoking doubles the risk of developing rheumatoid arthritis (RA) (1), particularly the risk for seropositive RA (2) (rheumatoid factor or anti–cyclic citrullinated peptide autoantibodies) and for poor prognosis. Smoking additionally contributes to cardiovascular, pulmonary, and oncologic diseases, which are the main causes of death in patients with RA (3). Continuing to smoke after being diagnosed with RA contributes to RA severity (4), treatment failure (5,6), and higher medication dose requirements (7). Patients with RA who continue smoking for at least 5 years after RA diagnosis are at even greater risk of death than smokers without RA and patients with RA who quit smoking (8). While patients are more likely to quit smoking after diagnosis of a smoking-related chronic disease, research shows that most will continue to smoke (8–10). Further, most patients with RA are not aware of the associations between smoking and RA development and complications (11), although such knowledge could influence cessation attempts.

In the US, some populations are more likely to be current smokers: men, individuals ages 25–64 years, American Indian/Alaska Natives, those of low socioeconomic status (SES), and those living in rural areas or in the Midwestern and Southern states (12). Cessation rates in the US are associated with sex, age, race, and SES (10,12–14). While smoking cessation is universally recommended, knowing which patient populations are less likely to quit is helpful for targeting interventions. Furthermore, modifiable factors that increase cessation can help guide intervention implementation. Health system–level interventions such as smoking

Supported by the Rheumatology Research Foundation (R Bridge award to Dr. Bartels), the University of Wisconsin Medical Scientist Training Program (grant T32GM008692), and the NIH (Institute for Clinical and Translational Research, University of Wisconsin Clinical and Translational Science award 1UL1-TR-002373).

[1]Maria Schletzbaum, BA, Xing Wang, PhD, Megan E. Piper, PhD, Christie M. Bartels, MD, MS: University of Wisconsin School of Medicine and Public Health, Madison; [2]Robert Greenlee, PhD, MPH: Marshfield Clinic Research Institute, Marshfield, Wisconsin.

Dr. Bartels has received research support from Pfizer. No other disclosures relevant to this article were reported.

Address correspondence to Christie M. Bartels, MD, MS, 1485 Highland Avenue, #4132, Madison, WI 53705. Email: cb4@medicine.wisc.edu.

Submitted for publication August 20, 2019; accepted in revised form January 21, 2020.

SIGNIFICANCE & INNOVATIONS

- Despite recognition of smoking as a risk factor for poor rheumatoid arthritis (RA) outcomes, little is known about what predicts smoking cessation among patients with RA.
- We found that health system factors were more predictive of smoking cessation than patient sociodemographic factors, highlighting the value of implementing health system interventions for smoking cessation.
- Patients new to rheumatology care were 60% more likely to quit smoking, while seropositive patients were 43% less likely to quit smoking. Thus, new patients and seropositive patients present opportunities to emphasize the relationship between smoking and RA and support smoking cessation.
- Overall, our findings support implementing system interventions to promote cessation for all RA patients.

cessation intervention training for providers and staff, identification of smokers and documentation reminders in health records, and designated clinic staff responsible for cessation support are recognized as key for improving cessation rates by national tobacco cessation guidelines (15,16). However, it is unknown which patient or health system factors predict smoking cessation in patients with RA.

Given the adverse outcomes in patients with RA who continue smoking, the objective of our study was to identify predictors of smoking cessation in patients with RA in 2 health systems to guide future intervention implementation efforts. Based on known disparities in smoking initiation and cessation in the general and chronic disease population, we hypothesized that patient sociodemographic factors (sex, race, SES) would predict smoking cessation in patients with RA.

PATIENTS AND METHODS

Inclusion and exclusion. Electronic health records (EHRs) at 2 Midwestern health systems, 1 suburban academic system and 1 rural community system, were screened for RA cohort eligibility. Inclusion criteria were ages ≥18 years and at least 2 RA diagnosis codes (International Classification of Diseases, Ninth Revision [ICD-9] codes 714.0–714.33, 714.4, 714.8, 714.81, or 714.89) on 2 distinct dates, at least 2 months apart, within a 24-month period using a previously validated search algorithm (17,18). ICD-10 codes were back-converted to ICD-9 codes. For the suburban academic system, the study eligibility period was from January 1, 2003 through December 31, 2016 based on when EHRs were implemented. For the rural community system, the study eligibility period was from January 1, 2005 through December 31, 2016 based on when reliable smoking status data were available in the EHR. The index date was defined as the date of the first RA diagnosis code. The 12 months preceding the first RA diagnosis code (index date) were designated as the baseline period to evaluate baseline smoking status, comorbidities, and health care utilization. In order to assess the outcome of smoking cessation, at least 6 months of observation time following the index date were required for cohort inclusion. Study end date was the date of patient death, date of last health care encounter before the first occurrence of a 24-month or longer period without any encounters, or December 31, 2016. Patients also had to have at least 1 primary care visit (family medicine, internal medicine, geriatrics, obstetrics/gynecology, or pediatrics) and at least 1 rheumatology visit in the system during the baseline or study period to help ensure equal likelihood of capturing cessation. This medical records study was approved by the Minimal Risk Institutional Review Board at the University of Wisconsin School of Medicine and Public Health with a waiver of informed consent.

Smoking and cessation outcome classification. Baseline smoking status was determined by the patient's smoking status (never, current, or former) at the baseline health care visit immediately preceding the index date. Final smoking status was determined by the patient's EHR-documented smoking status at the most recent health care visit on or before the patient's study end date. Only patients with a baseline smoking status of current smoker were included in the cohort being evaluated for smoking cessation. As the primary outcome, patients who were current smokers at baseline and former smokers at the end of their study participation were classified as cessations. Patients who reported that they were current smokers at baseline but were listed as never smokers at the end of the study were imputed as quit at the end of the study (n = 18). Patients who remained current smokers were classified as continued smokers.

Covariates. Information on patient sociodemographics and comorbidities was obtained from EHR data. Sociodemographic variables included patient age, sex, race, ethnicity, and having Medicaid coverage. Race was self-reported as Caucasian/White, African American/Black, Asian, American Indian/Alaska Native, and other, including Native Hawaiian and Pacific Islander. At 1 health system, patients could select multiple races. If multiple races were selected, patients were categorized in the indicated non-White race category. Information on Hispanic ethnicity was collected separately from race. As an indicator of SES, patients were classified as ever having Medicaid coverage if they had any EHR claims paid by Medicaid during the baseline or study period.

Comorbidity data were collected from the 12-month baseline period before the first RA coded visit. Comorbidities included prior diabetes mellitus, hypertension, hyperlipidemia, chronic obstructive pulmonary disease, asthma, and cardiovascular disease (prior myocardial infarction, revascularization procedure, ischemic heart disease, heart failure, transient ischemic attack/stroke, and peripheral vascular disease) defined using previously validated

ICD-9 or ICD-10 code definitions (19–24). Autoantibody laboratory results for rheumatoid factor and anti–cyclic citrullinated peptide were reviewed for ever being positive in the baseline or study period based on the performing laboratory's criteria at the time the test was performed. Patients with primary care visits in the baseline period were classified as having baseline primary care. Similarly, patients with rheumatology visits in the baseline period (i.e., before RA diagnosis code) were classified as having baseline rheumatology care. Patients without any baseline rheumatology visits were classified as being new to rheumatology care.

Statistical analysis. Descriptive statistics of the baseline RA cohort characteristics were reported and compared among never smokers, current smokers, and former smokers using analysis of variance for continuous variables and chi-square tests for categorical variables. Univariate and multivariate logistic regression analyses were conducted to determine predictors of smoking cessation among patients with RA who were current smokers at baseline. Length of follow-up time was controlled for in multivariable models. Additional sensitivity analysis to account for clustering by health system was conducted using generalized linear mixed models for multilevel logistic regression. Analysis was conducted using SAS software, version 9.4.

RESULTS

Description of the RA cohort. The RA cohort included 3,577 patients from 2 health systems. Overall, 26% of patients were former smokers (n = 915) (Figure 1), and 14% were current smokers (n = 507) at baseline and thus eligible for the primary outcome of cessation. As shown in Table 1, males and patients between ages 40 and 59 years were more frequently represented among current smokers than never smokers ($P < 0.001$). The proportion of current smokers who were Black was twice that of never smokers (4.1% current versus 1.5% never; $P < 0.001$), and patients who had Medicaid made up 42.8% of current smokers but only 15.2% of never smokers ($P < 0.001$). Hispanic ethnicity was not associated with baseline smoking status ($P = 0.49$). Seropositivity was most common in current smokers, followed by former smokers, and lowest in never smokers (70.7%, 64.0%, and 58.7%, respectively; $P < 0.001$). Cardiac comorbidities at baseline were equally common among current or never smokers, but twice as common in former smokers (12.4% and 12.0% versus 24.5%; $P < 0.001$). Pulmonary comorbidities were more common in current and former smokers than in never smokers (25.4% and 28.5% versus 15.7%; $P < 0.001$). Diabetes mellitus was most common in former smokers, followed by current smokers, and

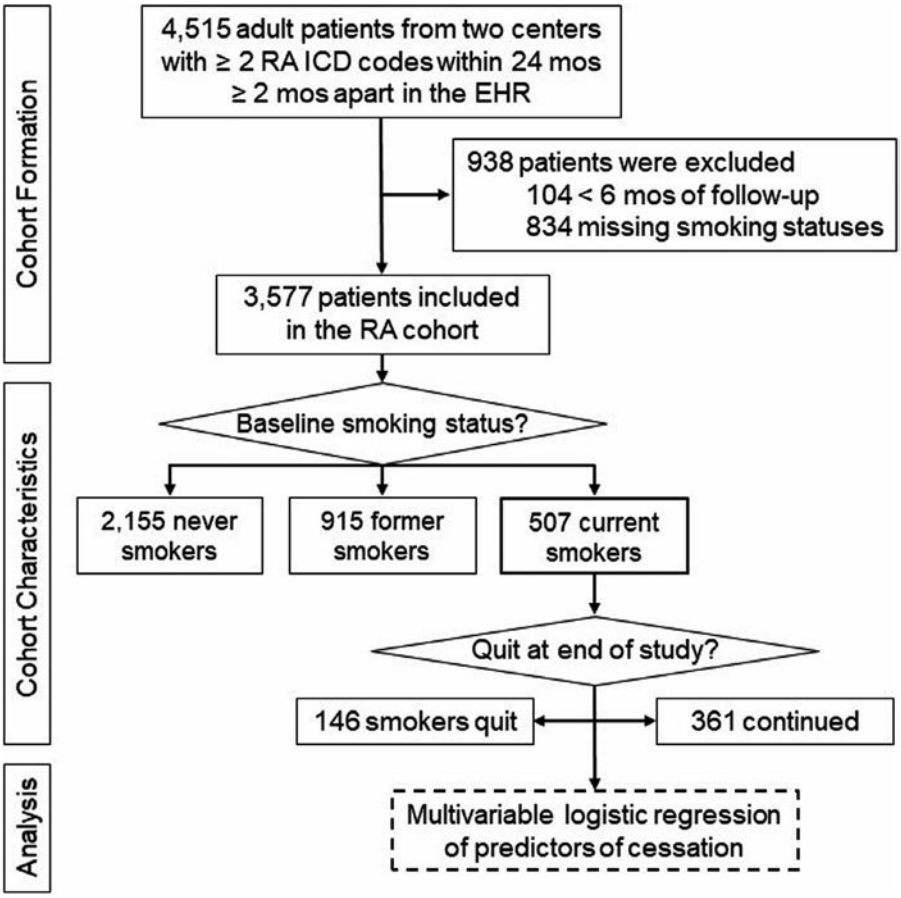


Figure 1. Rheumatoid arthritis (RA) current smoker cohort description. EHR = electronic health record; ICD = International Classification of Diseases.

Table 1. Rheumatoid arthritis (RA) patient characteristics by baseline smoking status*

Characteristic	Total cohort (n = 3,577)	Never smokers (n = 2,155)	Former smokers (n = 915)	Current smokers (n = 507)	*P*
Age at RA diagnosis, mean ± SD years†	56.4 ± 15.3	55.2 ± 16.2	61.5 ± 13.3	52.1 ± 12.1	<0.001‡
Age at index date, years					<0.001‡
18–39	535 (15.0)	390 (18.1)	62 (6.8)	83 (16.4)	
40–59	1,511 (42.2)	888 (41.2)	330 (36.1)	293 (57.8)	
60–79	1,304 (36.5)	735 (34.1)	443 (48.4)	126 (24.9)	
80+	227 (6.4)	142 (6.6)	80 (8.7)	5 (1.0)	
Male sex	941 (26.3)	419 (19.4)	350 (38.3)	172 (33.9)	<0.001‡
Race					<0.001‡
White	3,361 (94.0)	2,020 (93.7)	876 (95.7)	465 (91.7)	
Black	67 (1.9)	33 (1.5)	13 (1.4)	21 (4.1)	
Asian	52 (1.5)	44 (2.0)	6 (0.7)	2 (0.4)	
Native American/Alaska Native	37 (1.0)	19 (0.9)	7 (0.8)	11 (2.2)	
Other	12 (0.3)	7 (0.3)	2 (0.2)	3 (0.6)	
Unknown	48 (1.3)	32 (1.5)	11 (1.2)	5 (1.0)	
Ethnicity, Hispanic§	49 (1.4)	33 (1.5)	9 (1.0)	7 (1.4)	0.494
Medicaid, ever	698 (19.5)	328 (15.2)	153 (16.7)	217 (42.8)	<0.001‡
RF antibodies¶	1,398 (64.0)	752 (61.5)	378 (65.0)	268 (70.5)	0.005‡
Anti-CCP antibodies¶	1,317 (50.9)	695 (47.6)	371 (52.7)	251 (59.6)	<0.001‡
Seropositive, RF or CCP¶	1,766 (62.0)	958 (58.7)	489 (64.0)	319 (70.7)	<0.001‡
Cardiac comorbidity	546 (15.3)	259 (12.0)	224 (24.5)	63 (12.4)	<0.001‡
Pulmonary comorbidity	729 (20.4)	339 (15.7)	261 (28.5)	129 (25.4)	<0.001‡
Diabetes mellitus	406 (11.4)	198 (9.2)	149 (16.3)	59 (11.6)	<0.001‡
Baseline primary care	2,665 (74.5)	1,478 (68.6)	763 (83.4)	424 (83.6)	<0.001‡
Baseline rheumatology care	1,826 (51.1)	1,086 (50.4)	483 (52.8)	257 (50.7)	0.472
Suburban academic system	1,891 (52.9)	1,225 (64.8)	468 (24.8)	198 (10.5)	<0.001‡
Rural community system	1,686 (47.1)	930 (55.2)	447(26.5)	309 (18.3)	
Follow-up, median (IQR) years	5.67 (3.08–8.75)	6.67 (3.42–9.58)	4.83 (2.50–6.92)	4.75 (2.67–7.25)	<0.001‡
Death before end of study	356 (10.0)	215 (10.0)	98 (10.7)	43 (8.5)	0.404

* Values are the number (%) unless indicated otherwise. Anti-CCP = anti–cyclic citrullinated peptide; IQR = interquartile range; RF = rheumatoid factor.
† Due to missing values, n = 3,230 for age at RA diagnosis.
‡ Significant.
§ Due to missing values, n = 3,549 for ethnicity.
¶ Due to missing values, n = 2,191 for RF antibodies, n = 2,594 for anti-CCP antibodies, n = 2,856 for either RF or anti-CCP seropositivity.

lowest in never smokers (16.3%, 11.6%, and 9.2%; $P < 0.001$). Current and former smokers were more likely than never smokers to have had a primary care visit in the baseline year (83% and 83% versus 69%; $P < 0.001$), while baseline smoking was not associated with having baseline rheumatology visits ($P = 0.47$). The suburban academic system had fewer current smokers at baseline ($P < 0.001$). Death during the study did not differ significantly by baseline smoking status ($P = 0.40$). Over a median observation period of 4.75 years (interquartile range 2.67–7.25), 28.8% of patients (n = 146) with RA who smoked at baseline had quit.

Predictors of smoking cessation. Adjusted for observation time, being new to rheumatology care increased the odds of smoking cessation by 60% (adjusted odds ratio [OR_{adj}] 1.60 [95% confidence interval (95% CI) 1.02–2.50], $P = 0.041$) (Table 2) among patients with RA who were baseline smokers. Patients in the rural community system were 1.66 times more likely to quit smoking (OR_{adj} 1.66 [95% CI 1.03–2.69], $P = 0.039$). Conversely, seropositivity decreased the odds of quitting by ~43% (OR_{adj} 0.57 [95% CI 0.35–0.91], $P = 0.018$). Age, sex, race, ethnicity, and ever having Medicaid coverage were not predictive of smoking cessation. In generalized linear mixed-model sensitivity analysis to account for clustering by health system, seropositivity remained significant ($P = 0.018$), while being new to rheumatology care did not reach statistical significance ($P = 0.053$). Sex, age, race, and ethnicity remained nonsignificant ($P > 0.05$).

DISCUSSION

Patient sociodemographic factors are associated with smoking and cessation in some chronic diseases (10,12–14); thus, we hypothesized that patient factors would predict smoking cessation in patients with RA. Instead, we found that health care factors, such as being new to rheumatology care and the health system, were more predictive of smoking cessation. The observed increased likelihood of quitting in patients new to rheumatology care may partially be due to cessation following a new RA diagnosis, a phenomenon previously reported in RA and other chronic diseases (8–10,14,25). Despite their worse RA prognosis, seropositive patients were 43% less likely to quit smoking (26,27).

Table 2. Odds of smoking cessation at last follow-up in patients with rheumatoid arthritis*

Explanatory variable	Unadjusted OR (95% CI)	Adjusted OR (95% CI) (n = 442)	*P*
Age at index date, years			
18–39	Ref.	Ref.	–
40–59	0.68 (0.40–1.14)	0.75 (0.42–1.35)	0.336
60–79	0.80 (0.45–1.45)	0.88 (0.45–1.72)	0.700
80+	0.47 (0.05–4.66)	0.44 (0.04–4.44)	0.488
Male sex	1.02 (0.68–1.54)	0.92 (0.58–1.47)	0.735
Race			
White	Ref.	Ref.	–
Black	1.20 (0.47–3.03)	1.50 (0.54–4.20)	0.440
Other	0.25 (0.06–1.10)	0.17 (0.02–1.36)	0.095
Ethnicity, Hispanic	0.98 (0.19–5.12)	2.47 (0.38–16.07)	0.344
Medicaid, ever	0.98 (0.67–1.45)		
Seropositive, RF or anti-CCP	0.58 (0.37–0.89)	0.57 (0.35–0.91)	0.018†
Baseline primary care	0.86 (0.52–1.44)		
New to rheumatology, no baseline visits	1.17 (0.79–1.71)	1.60 (1.02–2.50)	0.041†
Cardiopulmonary comorbidity	0.84 (0.55–1.26)		
Rural community health system	1.33 (0.89–2.00)	1.66 (1.03–2.69)	0.039†
Length of study follow-up, years	1.13 (1.05–1.21)	1.15 (1.06–1.25)	<0.001†

* Cessation analysis included 507 baseline smokers. Of these baseline smokers, 146 quit by the end of study follow-up. 95% CI = 95% confidence interval; anti-CCP = anti–cyclic citrullinated peptide; OR = odds ratio; Ref. = reference; RF = rheumatoid factor.
† Significant.

Health systems can play a significant role in helping patients with RA to quit smoking. In our study, patients at the rural community system were proportionally more likely to quit smoking, potentially due to smoking cessation interventions that were implemented in that health system. For example, in the rural community health system, roughly one-half of the patients are covered by the system's health insurance plan, which offers robust smoking cessation interventions for members, including free telephone-based cessation coaching and coverage of cessation medications and nicotine replacements.

Health system policies, procedures, and programs represent modifiable factors that can be leveraged to increase cessation rates in patients with RA. Given that physicians may only counsel a minority of smokers about cessation (28–30), health system–level interventions can lead to more universal cessation counseling and improved cessation rates (15,16). In a sensitivity analysis controlling for health care system, patients new to rheumatology care no longer had a significantly higher likelihood of cessation, potentially suggesting greater importance of system-level factors versus individual patient factors (e.g., new patient) in promoting cessation. Within rheumatology, clinic-level interventions have been shown to support cessation (31) and increase referrals to cessation resources, such as state smoking quit lines (32). At the 2016 study end point, such interventions were implemented in the academic health system rheumatology clinics included in this study (32).

Patients new to rheumatology care with a diagnosis of RA were more likely to quit smoking. New patient visits are often longer in duration, potentially providing more opportunity to discuss the etiology of RA and to provide cessation counseling. Patients new to rheumatology care may also be newly diagnosed with RA and more motivated to quit, as seen in previous literature on smokers with new chronic disease diagnoses (14,25,33), and as suggested by the prevalence of cardiovascular disease being 2 times higher in former smokers than current smokers in this study. Thus, the new patient period represents a window of opportunity, and cessation efforts should be emphasized at this time. Additionally, both new and established patients with RA often are not aware of the relationship between smoking and RA (11,34). This lack of awareness has been cited as one of the primary barriers to smoking cessation (11,34) and should be stressed with all patients who are current smokers or former smokers at risk of relapse.

We observed that seropositive patients were less likely to quit smoking despite increased risk of RA disease progression (26,27), cardiovascular disease (35,36), and mortality (35,36). The reduced likelihood of cessation in seropositive patients could reflect greater smoking intensity, as smoking and greater intensity of smoking are strongly correlated with seropositivity (2). While all RA patients should receive smoking cessation support, seropositive patients may need targeted smoking cessation efforts and may require greater support or more intensive interventions.

Despite the strengths of this study, including analysis of patients receiving care in 2 health systems with rural and suburban populations, we acknowledge some limitations. First, while we adjusted for some potential differences in the patient populations between the 2 health systems, additional differences in the patient populations such as educational level, income, and location (rural, suburban, urban) were not measured. These unmeasured patient factors could contribute to the differences in smoking cessation

seen in the 2 health systems, as could the initial proportion of current smokers in each system. However, results were similar in sensitivity analysis accounting for clustering by health system. Among baseline current smokers, 92% of patients were White and non-Hispanic. Thus, we were underpowered to detect some differences by race and ethnicity in current smokers. This study could be replicated in other health systems with more diverse patient populations. Additionally, given the retrospective nature of our study and limitations of EHR data, we lacked complete information on cumulative smoking exposure (e.g., pack-years), which limits our ability to compare smoking intensity among groups. As inclusion in the study depended on reporting and documenting smoking status in the electronic medical record (EMR), there is the potential for misclassification bias for both the exposure and outcome. However, in a similar cohort of patients with systemic lupus erythematosus at 1 of the health systems in this study, manually comparing provider notes to EMR-listed smoking status resulted in the reclassification of <1% of patients, suggesting validity of our approach (37). As smoking status was evaluated only at 2 points, interim quit attempts and recidivism may not have been captured. However, the cessations that were captured were more likely to be sustained cessations, which are most clinically meaningful.

Despite these limitations, the results of our study showed an association of smoking with health care factors consistent with national guidelines endorsing health system interventions, and they suggest that such interventions may be most helpful in promoting cessation in patients with RA. As smoking cessation is a crucial secondary prevention measure for patients with RA, additional studies are warranted on what motivates cessation, including the influence of extraarticular manifestations, such as interstitial lung disease and other comorbidities, and what cessation methods are utilized by patients with RA. Additional evaluation of the impact of system-level interventions, as subsequently performed at the academic health system in this study, will also be important.

In conclusion, while patient sociodemographic factors may predict smoking initiation, in our study of patients with RA, cessation was better predicted by health care factors. Patients with RA who were new to rheumatology care were 60% more likely to quit smoking. In our study, patients receiving care at the rural community health system were also significantly more likely to quit smoking, possibly reflecting system interventions. Conversely, seropositive patients were 43% less likely to quit. Our findings identify further opportunities to target cessation efforts to patients who are new to rheumatology care or seropositive. The results of this study highlight the role of the health care system in smoking cessation, pointing toward health system–level cessation efforts implemented in rheumatology clinics as a potential path for greater smoking cessation in patients with RA.

ACKNOWLEDGMENTS

The authors thank Amanda Perez and Monica Messina for manuscript formatting.

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. Bartels 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. Schletzbaum, Greenlee, Bartels.

Acquisition of data. Greenlee, Bartels.

Analysis and interpretation of data. Schletzbaum, Wang, Greenlee, Piper, Bartels.

REFERENCES

1. Sugiyama D, Nishimura K, Tamaki K, Tsuji G, Nakazawa T, Morinobu A, et al. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. Ann Rheum Dis 2010;69:70–81.
2. Liu X, Tedeschi SK, Barbhaiya M, Leatherwood CL, Speyer CB, Lu B, et al. Impact and timing of smoking cessation on reducing risk of rheumatoid arthritis among women in the nurses' health studies. Arthritis Care Res (Hoboken) 2019;71:914–24.
3. England BR, Sayles H, Michaud K, Caplan L, Davis LA, Cannon GW, et al. Cause-specific mortality in male US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:36–45.
4. Gianfrancesco MA, Trupin L, Shiboski S, van der Laan M, Graf J, Imboden J, et al. Smoking is associated with higher disease activity in rheumatoid arthritis: a longitudinal study controlling for time-varying covariates. J Rheumatol 2019;46:370–5.
5. Saevarsdottir S, Rezaei H, Geborek P, Petersson I, Ernestam S, Albertsson K, et al. Current smoking status is a strong predictor of radiographic progression in early rheumatoid arthritis: results from the SWEFOT trial. Ann Rheum Dis 2015;74:1509–14.
6. Van den Broek M, Klarenbeek NB, Dirven L, van Schaardenburg D, Hulsmans HM, Kerstens PJ, et al. Discontinuation of infliximab and potential predictors of persistent low disease activity in patients with early rheumatoid arthritis and disease activity score-steered therapy: subanalysis of the BeSt study. Ann Rheum Dis 2011;70:1389–94.
7. Westhoff G, Rau R, Zink A. Rheumatoid arthritis patients who smoke have a higher need for DMARDs and feel worse, but they do not have more joint damage than non-smokers of the same serological group. Rheumatology (Oxford) 2008;47:849–54.
8. Sparks JA, Chang SC, Nguyen US, Barbhaiya M, Tedeschi SK, Lu B, et al. Smoking behavior changes in the early rheumatoid arthritis period and risk of mortality during thirty-six years of prospective followup. Arthritis Care Res (Hoboken) 2018;70:19–29.
9. Quinones AR, Nagel CL, Newsom JT, Huguet N, Sheridan P, Thielke SM. Racial and ethnic differences in smoking changes after chronic disease diagnosis among middle-aged and older adults in the United States. BMC Geriatr 2017;17:48.
10. Holm M, Schioler L, Andersson E, Forsberg B, Gislason T, Janson C, et al. Predictors of smoking cessation: a longitudinal study in a large cohort of smokers. Respir Med 2017;132:164–9.
11. Wattiaux A, Bettendorf B, Block L, Gilmore-Bykovskyi A, Ramly E, Piper ME, et al. Patient perspectives on smoking cessation and interventions in rheumatology clinics. Arthritis Care Res (Hoboken) 2020;72:369–77.
12. Centers for Disease Control and Prevention. Best practices user guide: health equity in tobacco prevention and control. Atlanta: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2015.
13. Van Eerd EA, van Schayck OC, Wesseling G, Kotz D. Predictors of long-term smoking cessation in patients with COPD: results from a randomised controlled trial. Eur Respir J 2017;49.

14. Freund KM, D'Agostino RB, Belanger AJ, Kannel WB, Stokes J III. Predictors of smoking cessation: the Framingham Study. Am J Epidemiol 1992;135:957–64.
15. Curry SJ, Keller PA, Orleans CT, Fiore MC. The role of health care systems in increased tobacco cessation. Annu Rev Public Health 2008;29:411–28.
16. Fiore MC, Jaén CR, Baker TB, Bailey WC, Benowitz NL, Curry SJ, et al. Treating tobacco use and dependence: 2008 update. Clinical practice guideline. Rockville (MD): US Department of Health and Human Services. Public Health Service; 2008.
17. Katz JN, Barrett J, Liang MH, Bacon AM, Kaplan H, Kieval RI, et al. Sensitivity and positive predictive value of Medicare Part B physician claims for rheumatologic diagnoses and procedures. Arthritis Rheum 1997;40:1594–600.
18. MacLean CH, Louie R, Leake B, McCaffrey DF, Paulus HE, Brook RH, et al. Quality of care for patients with rheumatoid arthritis. JAMA 2000;284:984–92.
19. Blecker S, Herbert R, Brancati FL. Comorbid diabetes and end-of-life expenditures among Medicare beneficiaries with heart failure. J Card Fail 2012;18:41–6.
20. Centers for Medicare & Medicaid Services. Measure information form: admissions for chronic obstructive disease (COPD) or asthma in older adults; 2016.
21. Borzecki AM, Wong AT, Hickey EC, Ash AS, Berlowitz DR. Identifying hypertension-related comorbidities from administrative data: what's the optimal approach? Am J Med Qual 2004;19:201–6.
22. Tu K, Campbell NR, Chen ZL, Cauch-Dudek KJ, McAlister FA. Accuracy of administrative databases in identifying patients with hypertension. Open Med 2007;1:e18–26.
23. Levy L, Fautrel B, Barnetche T, Schaeverbeke T. Incidence and risk of fatal myocardial infarction and stroke events in rheumatoid arthritis patients: a systematic review of the literature. Clin Exp Rheumatol 2008;26:673–9.
24. Newton KM, Wagner EH, Ramsey SD, McCulloch D, Evans R, Sandhu N, et al. The use of automated data to identify complications and comorbidities of diabetes: a validation study. J Clin Epidemiol 1999;52:199–207.
25. Rigotti NA, Singer DE, Mulley AG Jr, Thibault GE. Smoking cessation following admission to a coronary care unit. J Gen Intern Med 1991;6:305–11.
26. Kaltenhauser S, Wagner U, Schuster E, Wassmuth R, Arnold S, Seidel W, et al. Immunogenetic markers and seropositivity predict radiological progression in early rheumatoid arthritis independent of disease activity. J Rheumatol 2001;28:735–44.
27. Scott DL, Coulton BL, Popert AJ. Long term progression of joint damage in rheumatoid arthritis. Ann Rheum Dis 1986;45:373–8.
28. Vreede AP, Johnson HM, Piper M, Panyard DJ, Wong JC, Bartels CM. Rheumatologists modestly more likely to counsel smokers in visits without rheumatoid arthritis control: an observational study. J Clin Rheumatol 2017;23:273–7.
29. Naranjo A, Khan NA, Cutolo M, Lee SS, Lazovskis J, Laas K, et al. Smoking cessation advice by rheumatologists: results of an international survey. Rheumatology (Oxford) 2014;53:1825–9.
30. Cokkinides VE, Halpern MT, Barbeau EM, Ward E, Thun MJ. Racial and ethnic disparities in smoking-cessation interventions: analysis of the 2005 National Health Interview Survey. Am J Prev Med 2008;34:404–12.
31. Naranjo A, Bilbao A, Erausquin C, Ojeda S, Francisco FM, Rua-Figueroa I, et al. Results of a specific smoking cessation program for patients with arthritis in a rheumatology clinic. Rheumatol Int 2014;34:93–9.
32. Bartels CM, Ramly E, Panyard D, Lauver D, Johnson H, Li Z, et al. Rheumatology clinic smoking cessation protocol markedly increases quit line referrals [abstract]. Arthritis Rheumatol 2017;69 Suppl 10: URL: https://acrabstracts.org/abstract/rheumatology-clinic-smoking-cessation-protocol-markedly-increases-quit-line-referrals/.
33. Danielsen SE, Lochen ML, Medbo A, Vold ML, Melbye H. A new diagnosis of asthma or COPD is linked to smoking cessation: the Tromso study. Int J Chron Obstruct Pulmon Dis 2016;11:1453–8.
34. Aimer P, Stamp L, Stebbings S, Valentino N, Cameron V, Treharne GJ. Identifying barriers to smoking cessation in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2015;67:607–15.
35. Gabriel SE. Cardiovascular morbidity and mortality in rheumatoid arthritis. Am J Med 2008;121 Suppl S:9S–14S.
36. Goodson NJ, Wiles NJ, Lunt M, Barrett EM, Silman AJ, Symmons DP. Mortality in early inflammatory polyarthritis: cardiovascular mortality is increased in seropositive patients. Arthritis Rheum 2002;46:2010–9.
37. McKown T, Schletzbaum M, Unnithan R, Wang X, Ezeh N, Bartels CM. The effect of smoking on cumulative damage in systemic lupus erythematosus: an incident cohort study. Lupus 2021;30:620–9.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 640–648
DOI 10.1002/acr.24178

A Population-Based Approach to Reporting System–Level Performance Measures for Rheumatoid Arthritis Care

Claire E. H. Barber,[1] Deborah A. Marshall,[1] Elena Szefer,[2] Cheryl Barnabe,[1] Natalie J. Shiff,[3] Vivian Bykerk,[4] Joanne Homik,[5] J. Carter Thorne,[6] Vandana Ahluwalia,[7] Susanne Benseler,[8] Dianne Mosher,[9] Marinka Twilt,[8] and Diane Lacaille[10]

Objective. To operationalize and report on nationally endorsed rheumatoid arthritis (RA) performance measures (PMs) using health administrative data for British Columbia (BC), Canada.

Methods. All patients with RA in BC ages ≥18 years were identified between January 1, 1997 and December 31, 2009 using health administrative data and followed until December 2014. PMs tested include: the percentage of incident patients with ≥1 rheumatologist visit within 365 days; the percentage of prevalent patients with ≥1 rheumatologist visit per year; the percentage of prevalent patients dispensed disease-modifying antirheumatic drug (DMARD) therapy; and time from RA diagnosis to DMARD therapy. Measures were reported on patients seen by rheumatologists, and in the total population.

Results. The cohort included 38,673 incident and 57,922 prevalent RA cases. The percentage of patients seen by a rheumatologist within 365 days increased over time (35% in 2000 to 65% in 2009), while the percentage of RA patients under the care of a rheumatologist seen yearly declined (79% in 2001 to 39% in 2014). The decline was due to decreasing visit rates with increasing follow-up time rather than calendar effect. The percentage of RA patients dispensed a DMARD was suboptimal over follow-up (37% in 2014) in the total population but higher (87%) in those under current rheumatologist care. The median time to DMARD in those seen by a rheumatologist improved from 49 days in 2000 to 23 days in 2009, with 34% receiving treatment within the 14-day benchmark.

Conclusion. This study describes the operationalization and reporting of national PMs using administrative data and identifies gaps in care to further examine and address.

INTRODUCTION

Rheumatoid arthritis (RA) is a potentially debilitating disease that affects ~1% of the population (1,2). Early and targeted treatment is the paradigm of RA care endorsed by guidelines (3–5). In 2016, the Arthritis Alliance of Canada (AAC) (6) developed a set of system-level performance measures (PMs) that were created to reflect system adherence to these treatment principles (7). The 6 measures capture timely access to rheumatologist care, ongoing follow-up, and timely and ongoing treatment for patients with inflammatory arthritis. To date, the measures have been tested in a longitudinal, early RA cohort (8) as well as using clinical data from 5 different models of care across Canada (9).

The objective of the current study was to operationalize and report on 4 of the 6 AAC PMs using population-level data from administrative health databases, and to explore the impact of different denominators on measure-reporting to better understand the impact this has on measure attribution at the level of rheumatologist

All inferences, opinions, and conclusions drawn herein are those of the authors and do not reflect the opinions or policies of the data stewards.

Supported by the Canadian Initiative for Outcomes in Rheumatology Care. Dr. Marshall's work was supported by a Canada Research Chair in Health Systems and Services Research and the Arthur J. E. Child Chair in Rheumatology Outcomes Research. Dr. Lacaille's work was supported by the Arthritis Society of Canada and the University of British Columbia (Mary Pack Chair in Rheumatology Research).

[1]Claire E. H. Barber, MD, PhD, FRCPC, Deborah A. Marshall, PhD, Cheryl Barnabe, MD, MSc, FRCPC: University of Calgary, Calgary, Alberta, Canada, and Arthritis Research Canada, Richmond, British Columbia, Canada; [2]Elena Szefer, MSc: Emmes Canada, Burnaby, British Columbia, Canada; [3]Natalie J. Shiff, MD, MHSc: University of Saskatchewan, Saskatoon, Saskatchewan, Canada; [4]Vivian Bykerk, MD, FRCPC: Hospital for Special Surgery and Cornell University, New York, New York; [5]Joanne Homik, MD, MSc, FRCPC: University of Alberta, Edmonton, Alberta, Canada; [6]J. Carter Thorne, MD, FRCPC: Southlake Regional Health Care, Newmarket, Ontario, Canada; [7]Vandana Ahluwalia, MD, FRCPC: William Osler Health System, Brampton, Ontario, Canada; [8]Susanne Benseler, MD, PhD, Marinka Twilt, MD, MSCE, PhD: Alberta Children's Hospital, University of Calgary, Calgary, Alberta, Canada; [9]Dianne Mosher, MD, FRCPC: University of Calgary, Calgary, Alberta, Canada; [10]Diane Lacaille, MD, MHSc, FRCPC: Arthritis Research Canada, Richmond, and University of British Columbia, Vancouver, British Columbia, Canada.

Dr. Shiff owns stock in AbbVie, Gilead, Iovance, Novartis, and Mylan. No other disclosures relevant to this article were reported.

Address correspondence to Claire E. H. Barber, MD, PhD, FRCPC, Division of Rheumatology, University of Calgary, HRIC 3AA20-3280, Hospital Drive NW, Calgary, Alberta T2N 4Z6, Canada. Email: cehbarbe@ucalgary.ca.

Submitted for publication October 4, 2019; accepted in revised form February 25, 2020.

SIGNIFICANCE & INNOVATIONS

- This is the first time that the nationally endorsed performance measures have been operationalized using administrative data, allowing us to report on quality of care in rheumatoid arthritis at a population level.
- At the population level, improvements over time were seen in access to rheumatologist care and timeliness of early treatment, while suboptimal rates of rheumatology follow-up persisted.
- Higher performance rates on treatment measures were observed in patients accessing rheumatologist care over the measurement period.

and the provincial health care system. Populations of interest for applying the PMs include the total RA population in BC and those who received rheumatologist care using different definitions of rheumatologist care. The measures evaluated included: 1) the percentage of newly diagnosed RA patients seen by a rheumatologist within 1 year of diagnosis; 2) the percentage of prevalent RA patients previously seen by a rheumatologist seen in yearly rheumatology follow-up; 3) the percentage of prevalent RA patients treated with a disease-modifying antirheumatic drug (DMARD); and 4) time from RA diagnosis to DMARD therapy for newly diagnosed RA patients.

MATERIALS AND METHODS

Study design and cohort definition. The PMs were evaluated in a longitudinal, population-based RA cohort using administrative health data from the province of British Columbia (BC), Canada. All prevalent patients ages ≥18 years who received care for RA in BC between January 1, 1997 and December 31, 2009 were identified using administrative health data and were followed until December 2014. Patients with RA were identified using an established case definition (10), with inclusion criteria of at least 2 physician visits ≥8 weeks apart within 5 years, and with an International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) code for RA (714.X). The following exclusion criteria were applied over a 5-year period after the index date (i.e., the second RA visit when patients meet the inclusion criteria): 1) ≥2 subsequent physician visits for the same type of other inflammatory arthritis, including systemic lupus erythematosus, connective tissue diseases, psoriatic arthritis, ankylosing spondylitis, or other spondyloarthropathy (ICD-9-CM 710.x, 696.x; 720.x; 713.1, 555.x, 556.x); and 2) if patients saw a rheumatologist and RA was never confirmed. Individuals were determined to be incident patients if they first met RA criteria between January 1997 and December 2009 based on data from January 1990 onwards. Additionally, individuals were required to have ≥7 years of registration data prior to their index date to be considered an incident patient.

As the purpose of this study was to evaluate the operationalization of PMs, different denominator definitions were used. Where relevant, measures were evaluated in the following RA samples: 1) all RA patients in BC (i.e., the total RA population regardless of the specialty of the physician providing RA care) and 3 variations in the denominator definition to identify RA patients having received rheumatologist care; 2) those ever seeing a rheumatologist during follow-up (defined as having ≥1 rheumatologist visit at any time during follow-up); and 3) those under rheumatologist care (defined as having previously had ≥2 rheumatologist visits before the measurement year). The requirement of 2 visits was to ensure that the rheumatologist had assumed care of the patient with RA, i.e., to exclude patients referred for evaluation of possible RA but never seen again; and 4) those under active rheumatologist care (i.e., with a rheumatologist visit during the measurement year).

Data sources. For this study, data were obtained from administrative databases of the Ministry of Health of BC, through Population Data BC, including: all physician visits since January 1990, with 1 diagnostic code per visit representing the reason for the visit and a specialty code, introduced in 1997, allowing identification of rheumatologists from the Medical Service Plan database (11), demographic information from the health registration database (12), and PharmaNet data (13), which includes information on all medications dispensed from pharmacies for all individuals, regardless of funding source, since January 1996. Data were available for all individuals receiving universal health care coverage in the province of BC (~4.6 million people) until December 2014.

Ethics. Ethics approval for the project was obtained from the University of British Columbia (REB No. H00-80305). No personal identifying information was provided, and all procedures were compliant with BC's Freedom of Information and Protection of Privacy Act.

Calculation of PMs. Four PMs from the AAC (7) were operationalized in this study. Measures and samples are listed in Table 1. PMs calculated using the incident cohort are reported for calendar years 2000–2009, 2009 being the latest year of incidence available because 5 years are required after the index date to apply the case definition exclusion criteria.

For the first PM, the incident RA cohort was used to ascertain the percentage of patients with newly diagnosed RA with at least 1 visit to a rheumatologist within 365 days of their first RA diagnostic code. If the first RA code was from a rheumatologist, then the PM was met. If the patient was first coded as having RA by another type of practitioner and was then seen within 365 days by a rheumatologist, but the 2 events fell within different measurement years, then the measure was reported as met in the year of diagnosis. Patients who died or relocated within 1 year post-RA

Table 1. List of performance measures and the samples of rheumatoid arthritis (RA) patients to whom they were applied*

Performance measures	RA sample (denominator indicated by X)†				
	Incident vs. prevalent	Total RA population‡	Ever seeing a rheumatologist§	Under care of a rheumatologist (ever in past)¶	Under active care of a rheumatologist#
Percentage of RA patients ages ≥18 years seen by a rheumatologist within 1 year of diagnosis	Incident	X	X		
Percentage of RA patients ages ≥18 years previously seen by a rheumatologist seen in yearly rheumatology follow-up	Prevalent	NA		X	
Percentage of RA patients ages ≥18 years treated with a DMARD	Prevalent	X	X	X	X
Time from RA diagnosis to DMARD therapy for newly diagnosed patients ages ≥18 years	Incident	X	X		

* DMARD = disease-modifying antirheumatic drug; NA = not applicable.
† RA cases ages ≥18 years: ≥2 physician visits ≥8 weeks apart within 5 years, with an International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) code for RA (714.X). Exclusion criteria (applied over a 5-year period after the index date): 1) ≥2 subsequent physician visits for the same type of other inflammatory arthritis, including systemic lupus erythematosus, connective tissue diseases, psoriatic arthritis, ankylosing spondylitis, or other spondyloarthropathy (ICD-9-CM 710.x, 696.x; 720.x; 713.1, 555.x, 556.x); 2) if patient saw a rheumatologist and RA was never confirmed.
‡ All RA patients in the province of British Columbia regardless of the specialty of the physician providing RA care.
§ Defined as ≥1 rheumatologist visit any time during follow-up (i.e., on or after initial RA diagnosis, but not necessarily before the measurement year).
¶ Defined as ≥2 rheumatologist visits on or after RA diagnosis and before the measurement year.
# Defined as ≥1 rheumatologist visit during the measurement year.

diagnosis were excluded from the denominator. The measure was reported in the entire RA population and in patients seen by a rheumatologist (those ever seen in follow-up and those seen within the first 5 years of follow-up). Equalizing follow-up time prevents differences across calendar years due to the greater likelihood of seeing a rheumatologist with increasing follow-up time. The median and 90th percentile number of days between the first RA visit and the first rheumatologist visit was also reported in the subset of patients seeing a rheumatologist within 5 years.

The second PM was estimated using the proportion of patients seen in follow-up by a rheumatologist at least once during each measurement year among prevalent RA patients who first saw a rheumatologist between 2000 and 2014. Once patients were seen by a rheumatologist at least twice, they were considered to be under the care of a rheumatologist until the end of follow-up and remained in the denominator for all subsequent measurement years. The requirement for 2 rheumatologist visits was to avoid including patients referred to a rheumatologist for consideration of RA in whom the diagnosis was not confirmed. Patients were excluded from the denominator if they died or relocated to another province within the measurement year, or if they were hospitalized for the entire measurement year.

The second PM was also operationalized to better capture gaps in follow-up care (8,14). A gap in follow-up care was defined as a period of at least 14 months between 2 consecutive rheumatologist visits. A 14-month rather than 12-month window was chosen because yearly follow-ups may be scheduled shortly after the 12-month mark due to scheduling availability and/or billing incentives, and thus visits slightly over the 12-month mark would not represent true deficits in care. The PM was considered to be met for all calendar years during time periods in which consecutive rheumatologist visits were ≤14 months apart. The PM was considered not met in the calendar year of the patient's first missed visit, and the measure continued to be considered not met until 2 consecutive visits ≤14 months apart occurred.

The third PM captures the percentage of prevalent RA patients who were dispensed a DMARD during the measurement year. It is reported in 1) the total RA population, 2) patients ever seeing a rheumatologist, 3) patients under rheumatology care (i.e., ever in the past), and 4) patients under active rheumatologist care (i.e., with a visit during the measurement year). DMARD therapy in this measure includes conventional DMARD medications (e.g., methotrexate, leflunomide, sulfasalazine, hydroxychloroquine, and gold), biologic agents, and small-molecule inhibitors (list shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract) (7). Patients with at least 1 DMARD prescription during the measurement year met the PM. Patients with a cancer diagnosis, HIV, or pregnancy during the measurement year were excluded from the denominator, as these diagnoses may preclude DMARD treatment. Exclusions from the denominator due to malignancy or HIV were applied from the date of first diagnostic code onwards, including all subsequent measurement years. Malignancy was defined as at least 1 hospitalization or physician visit with codes ICD-9-CM 140-208, or ICD-10-CA C00-26,30-41, 43-58, 60-69, 7A, 7B. HIV was defined as 3 physician visits or hospital admissions within 3 years ICD-9-CM codes 042, 043, 044 or ICD-10-CM B20-24 (15). Patients with pregnancy were excluded for 365 days before and after the delivery date to account for potential DMARD discontinuation due to

Table 2. Cohort demographics*

	Total RA population	Ever seeing a rheumatologist	
		Yes	No
Incident RA cohort demographics			
No. of RA cases	38,673	19,460 (50)	19,213 (50)
Age at index date, mean ± SD years	58.4 ± 16.6	56.6 ± 16	60.1 ± 17
Length of follow-up, mean ± SD years	10.7 ± 4.3	10.5 ± 4.1	10.9 ± 4.5
Female sex	25,629	13,727 (54)	11,902 (46)
Prevalent RA cohort demographics			
No. of RA cases	57,922	29,638 (51)	28,284 (49)
Age at index date, mean ± SD years	58 ± 16.8	56 ± 16	60.2 ± 17.4
Length of follow-up, mean ± SD years	11.4 ± 5	11.7 ± 4.8	11.1 ± 5.1
Female sex	39,031	21,416 (55)	17,615 (45)

* Values are the number (%) unless indicated otherwise. RA = rheumatoid arthritis.

pregnancy planning or breastfeeding. Pregnancy was defined as at least 1 physician visit or hospital admission with a pregnancy or delivery code (ICD-9-CM 630-639, 640-648, 670-679, V22, V23, V24.0-V24.2, V27, V30-V39 or ICD-10-CA O00-O99, Z37). If there were multiple visits or hospitalizations with delivery codes within 45 days, the last date was used.

The fourth PM calculated the time from first RA diagnosis to the first DMARD prescription for incident RA patients who received a DMARD (median and 90th percentile times to DMARD start). The measure was reported in the total RA population and in patients ever seeing a rheumatologist. If the start of treatment with a DMARD preceded the first RA diagnosis, a time of 0 was assigned. If the 2 events were not in the same measurement year, the measure was reported in the year of RA diagnosis. Additionally, the percentage of patients who were dispensed a DMARD within the 14-day benchmark from RA diagnosis was calculated.

RESULTS

The cohort included 38,673 incident and 57,922 prevalent RA patients. Cohort demographic characteristics are shown in Table 2. In the total RA population, the percentage of patients seeing a rheumatologist within the first year of diagnosis was suboptimal but improved over time from 35% in 2000 to 65% in 2009 (Table 3; see run chart in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract). In the sample of patients ever seeing a rheumatologist during follow-up, improved performance was also seen (74% in 2000 and 96% in 2009). However, lower performance in earlier years could be due to more patients seeing a rheumatologist in later years due to longer follow-up, thus increasing the denominator. Therefore, the measure was also reported in patients seeing a rheumatologist within the first 5-years of follow-up, which improved measure performance in the earlier years, with no impact in later years (88% in 2000 to 97% in 2009). The median number of days between the first RA diagnosis and the first rheumatologist visit was 0 for all years of measurement, indicating that the diagnosis of RA was made by the rheumatologist in >50% of cases. An improvement in the 90th percentile number of days between these 2 visit dates was seen over time but remained suboptimal (from 451 days in 2000 to 154 days in 2009). This was also calculated for RA patients ever seeing a rheumatologist (i.e., at any time during follow-up); but without ensuring equal follow-up for all measurement years, a length time bias was observed, with longer wait times seen in earlier cohort years (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract).

The percentage of patients entering the care of a rheumatologist in 2000 or later and seen in yearly follow-up by a rheumatologist declined in each subsequent calendar year, whether it was measured using the fixed-interval method (i.e., at least 1 rheumatologist visit in the measurement year) (Table 4) or using gaps of ≥14 months between consecutive rheumatologist visits (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract). By definition, the measure is met 100% of

Table 3. Percentage of rheumatoid arthritis (RA) patients seen by a rheumatologist within 1 year of diagnosis*

Calendar year of RA onset	Total RA population	RA patients ever seeing a rheumatologist	RA patients seeing a rheumatologist in first 5 years of follow-up
2000	2,944 (35)	1,411 (74)	1,182 (88)
2001	2,945 (37)	1,424 (77)	1,257 (88)
2002	2,889 (39)	1,377 (83)	1,258 (90)
2003	2,945 (38)	1,357 (82)	1,231 (91)
2004	2,962 (45)	1,525 (87)	1,430 (93)
2005	2,922 (45)	1,496 (89)	1,424 (93)
2006	2,891 (50)	1,587 (92)	1,535 (95)
2007	2,491 (47)	1,309 (90)	1,270 (93)
2008	2,169 (51)	1,185 (94)	1,169 (96)
2009	1,299 (65)	877 (96)	875 (97)

* The values in each column represent the denominator for each RA population, and the percentages (in parentheses) represent the proportion of this denominator meeting the performance measure. The denominator excludes patients who died or left the province within 365 days of RA onset.

Table 4. Percentage of prevalent patients with rheumatoid arthritis (RA) under the care of a rheumatologist seen in yearly follow-up, by calendar year, and fixed-interval method*

	Measurement year														
Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
2000	2,111 (100)	2,089 (58)	2,055 (44)	2,012 (42)	1,971 (40)	1,923 (40)	1,868 (38)	1,824 (37)	1,759 (36)	1,702 (37)	1,657 (36)	1,596 (36)	1,544 (34)	1,483 (35)	1,430 (35)
2001		2,085 (100)	2,043 (58)	2,002 (46)	1,961 (43)	1,920 (42)	1,884 (39)	1,850 (39)	1,804 (39)	1,756 (37)	1,688 (36)	1,639 (35)	1,59 (35)	1,561 (35)	1,512 (36)
2002			1,910 (100)	1,873 (60)	1,834 (49)	1,792 (46)	1,753 (42)	1,700 (42)	1,656 (40)	1,611 (40)	1,564 (39)	1,512 (38)	1,47 (38)	1,424 (37)	1,367 (37)
2003				2,008 (100)	1,978 (63)	1,943 (51)	1,907 (47)	1,853 (44)	1,814 (41)	1,769 (41)	1,723 (39)	1,680 (37)	1,625 (38)	1,569 (38)	1,502 (38)
2004					2,044 (100)	2,008 (65)	1,970 (55)	1,932 (48)	1,895 (44)	1,851 (41)	1,791 (40)	1,749 (38)	1,677 (38)	1,632 (36)	1,586 (36)
2005						2,051 (100)	2,020 (66)	1,978 (54)	1,942 (45)	1,904 (43)	1,853 (42)	1,812 (41)	1,758 (38)	1,709 (38)	1,644 (39)
2006							1,962 (100)	1,923 (68)	1,879 (54)	1,843 (49)	1,801 (44)	1,754 (42)	1,715 (41)	1,665 (38)	1,620 (38)
2007								1,701 (100)	1,658 (66)	1,628 (53)	1,591 (45)	1,551 (42)	1,506 (40)	1,462 (39)	1,413 (39)
2008									1,529 (100)	1,501 (69)	1,477 (52)	1,447 (46)	1,409 (42)	1,375 (41)	1,336 (39)
2009										1,279 (100)	1,260 (68)	1,235 (54)	1,206 (49)	1,175 (47)	1,137 (45)
2010											395 (100)	386 (49)	373 (40)	363 (36)	358 (36)
2011												337 (100)	325 (45)	316 (35)	306 (32)
2012													288 (100)	282 (44)	268 (37)
2013														279 (100)	270 (38)
2014															245 (100)
All	2,111 (100)	4,174 (79)	6,008 (66)	7,895 (62)	9,788 (59)	11,637 (58)	13,364 (56)	14,761 (54)	15,936 (51)	16,844 (49)	16,800 (45)	16,698 (42)	16,493 (40)	16,295 (39)	15,994 (39)

* Values are the number (%). Number represents the measure denominator; percentage represents the proportion meeting the measure shown. Year defined as the year entering rheumatologist care. Fixed-interval method: patients met the performance measure if they had at least 1 visit during the measurement year. Patients were excluded from denominator if they died or left the province in the current year or in a preceding year, or if they were hospitalized for the entire measurement year.

the time in the first measurement year (i.e., 2000). When looking at all RA patients under rheumatologist care regardless of when they entered care, the percentage declined from 79% in 2001 to 39% in 2014 (fixed-interval method) (Table 4) or from 82% in 2001 to 42% in 2014 (gaps method) (see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract). However, the decline in yearly rheumatologist visits over time was a function of loss of follow-up from rheumatologist care with increasing follow-up time rather than a calendar-year effect, as shown when results are stratified by the year that patients entered rheumatologist care.

The percentage of prevalent RA patients who were dispensed a DMARD during the measurement year is shown in Table 5 (see run chart in Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract). Overall, DMARD use was suboptimal in patients not under active rheumatologist care, with little improvement observed over time. Among all RA patients, regardless of the specialty of the physician providing care, only 37% received a DMARD in 2014. In patients ever seeing a rheumatologist (at any time during follow-up) and in patients under rheumatologist care (at least 2 visits prior to the measurement year), the percentage receiving a DMARD increased to 57% in 2014. The highest rates of DMARD use were seen in patients under active rheumatologist care (i.e., with a rheumatologist visit during the measurement year), at 87% in 2014.

In contrast, among those who received a DMARD, the median time from RA diagnosis to starting a DMARD among all incident RA patients who eventually saw a rheumatologist over follow-up was 49 days in 2000 and improved over time to 23 days in 2009 (Table 6; see run chart in Supplementary Figure 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract). A decline in the 90th percentile time was also seen from 1,644 days in 2000 to 118 days in 2009. The percentage of patients who received a DMARD within the benchmark of 14 days from RA diagnosis increased from 21% in 2000 to 34% in 2009. In a sensitivity analysis, using a 30-day benchmark, this increased from 28% in 2000 to 43% in 2009, respectively. Because lower performance in earlier calendar years could be due to a length time bias (i.e., patients in earlier years have a longer follow-up time available to start a DMARD), the measure was also calculated for patients who received a DMARD within 1 year of diagnosis. The percentage treated within 14 days increased from 39% in 2000 to 45% in 2009 (Table 6). When the PM was reported for the entire RA population (regardless of the specialty of the physician providing care), performance was lower, with only 25% of incident RA patients meeting the 14-day benchmark in 2009 (see Supplementary Table 4, available on

Table 5. Percentage of prevalent patients with rheumatoid arthritis (RA) who were dispensed a disease-modifying drug*

Measurement year	Total RA population†	Ever seeing a rheumatologist‡	Under care of a rheumatologist§	Under active care of a rheumatologist¶
2000	23,344 (36)	12,262 (60)	9,678 (66)	7,006 (75)
2001	25,461 (36)	13,581 (59)	11,294 (64)	7,506 (76)
2002	27,310 (35)	14,738 (59)	12,618 (63)	7,868 (79)
2003	29,190 (35)	15,828 (59)	13,962 (62)	8,569 (79)
2004	30,786 (36)	16,984 (59)	15,314 (62)	9,285 (81)
2005	32,151 (37)	18,045 (60)	16,569 (62)	9,905 (82)
2006	33,595 (38)	19,190 (60)	17,760 (63)	10,378 (84)
2007	34,619 (38)	20,007 (60)	18,643 (62)	10,685 (84)
2008	35,274 (38)	20,637 (60)	19,270 (62)	10,558 (85)
2009	35,104 (39)	20,862 (61)	19,682 (63)	10,783 (87)
2010	33,607 (39)	20,035 (60)	19,183 (62)	9,908 (89)
2011	31,771 (36)	18,950 (56)	18,365 (57)	9,139 (85)
2012	30,365 (36)	18,185 (56)	17,816 (57)	8,670 (86)
2013	28,845 (36)	17,379 (57)	17,204 (57)	8,335 (87)
2014	27,537 (37)	16,695 (57)	16,692 (57)	8,087 (87)

* The values in each column represent the denominator for each RA population, and the percentages (in parentheses) represent the proportion of this denominator meeting the performance measure. The denominators exclude patients who died, left the province, met the exclusion criteria for the performance measure (malignancy or HIV during the current year or in a preceding year; pregnancy during the measurement year). Disease-modifying antirheumatic drugs (azathioprine, chloroquine, cyclophosphamide, cyclosporin, gold, hydroxychloroquine, leflunomide, methotrexate, minocycline, mycophenolate mofetil, and sulfasalazine), biologics, and oral small-molecule inhibitors were dispensed (a complete list is shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract).
† All RA patients in British Columbia regardless of the specialty of the physician providing RA care.
‡ Defined as ≥1 rheumatologist visit at any time during follow-up (i.e., on or after initial RA diagnosis, but not necessarily before the measurement year).
§ Defined as ≥2 rheumatologist visits on or after RA diagnosis and before the measurement year.
¶ Defined as ≥1 rheumatologist visit during the measurement year.

Table 6. Time from diagnosis of rheumatoid arthritis (RA) to the start of disease-modifying antirheumatic drug (DMARD) therapy for incident RA patients who had a rheumatologist visit at some point during follow-up*

	RA patients treated with a DMARD at any time over follow-up				RA patients treated with a DMARD during the first year			
Year of RA onset	No. of incident RA cases	Median (90th percentile) time to DMARD start, days	% treated within 14 days†	% treated within 30 days	No. of incident RA cases	Median (90th percentile) time to DMARD start, days	% treated within 14 days†	% treated within 30 days
2000	1,158	49 (1,644)	21	28	612	27 (185)	39	53
2001	1,196	53.5 (1,521)	22	30	644	26 (178)	41	56
2002	1,092	41 (1,058)	26	33	638	22.5 (180)	44	57
2003	1,134	43 (1,115)	23	32	678	27 (181)	38	54
2004	1,237	31 (825)	27	37	787	21 (145)	43	59
2005	1,181	26 (579)	30	41	771	18 (161)	46	62
2006	1,278	23 (411)	31	42	855	17 (139)	47	62
2007	1,029	29 (399)	28	37	668	21 (149)	43	57
2008	946	26 (339)	29	40	644	21 (153)	42	58
2009	657	23 (188)	34	43	498	21 (139)	45	57

* Excludes patients who died, left the province, or met the performance measure exclusion criteria (malignancy or HIV exclusion criteria in the incident year or in a preceding year, or pregnancy in the incident year). DMARDs and other immunosuppressive agents were as follows: azathioprine, chloroquine, cyclophosphamide, cyclosporin, gold, hydroxychloroquine, leflunomide, methotrexate, minocycline, mycophenolate mofetil, sulfasalazine, biologics, and oral small-molecule inhibitors (a complete list is shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract).
† Benchmark is 14 days from diagnosis to the start of DMARD therapy. A 30-day window was also evaluated in a sensitivity analysis.

the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24178/abstract).

DISCUSSION

This study represents the first time that the AAC system-level PMs for RA have been operationalized using administrative data. The study has revealed a number of gaps in RA care, including loss of follow-up from rheumatology care and suboptimal DMARD use. While some delays in DMARD initiation in new-onset RA were observed, timeliness of DMARD use improved over the study period (although was still suboptimal), as did the percentage of RA patients seen by a rheumatologist within the first year of diagnosis. Importantly, this study reveals high rates of DMARD dispensation in patients under active rheumatology care, with much lower rates in the other RA samples evaluated. Given current rheumatologist shortages in many regions, this work highlights ongoing need for monitoring these measures if RA patients are returned to primary care.

This study also represents the first time that the percentage of patients seen within a year of diagnosis by a rheumatologist was tested. Population-level data, such as health administrative data, are required to evaluate this measure. This PM is designed to understand to what extent RA patients are not seen by rheumatologists and is not intended as a wait-time measure. When looking at the entire RA population, a dramatic improvement was observed over time (35% in 2000 to 65% in 2009), as well as in patients ever seeing a rheumatologist during follow-up (74% to 96%); however, a length time bias was observed, causing poorer performance in earlier calendar years, as discussed below. This improvement coincides with the paradigm shift in RA treatment emphasizing early diagnosis and DMARD treatment for RA (16–19) and the Ministry of Health's Chronic Disease Management Strategy for Arthritis in BC, an initiative to improve the care of arthritis (20). The availability of rheumatologists in BC only increased from 38 to 41 over this period (21). It is still possible, however, that the measure underestimates the problem, as primary care physicians may use different billing codes when they suspect a new diagnosis of RA prior to rheumatologist assessment.

This is not the first population-based study to demonstrate gaps in referrals to rheumatologists for patients with RA. This gap was initially described in RA patients receiving care from 1996–2000 in BC using administrative health data (14). Similarly, in a population-based study in Quebec, 27% of incident RA patients in 2000 saw a rheumatologist within the next 2.5–3.5 years (22). A population-based study in Ontario between 2000–2009 using the Ontario Rheumatoid Arthritis Administrative Database (23), where RA patients had at least 1 claim by a musculoskeletal specialist, found similarly high rates of RA patients seeing rheumatologists within 1 year of diagnosis, with improvement over time (81% in 2000; 89% in 2009) (23). The wait time from symptom onset or from primary care referral to a rheumatologist could not be examined in our study, as these data are not captured in administrative databases. Other sources, such as primary care databases and triage databases (9), have been used to capture wait times for rheumatologist care, including from symptom onset (24).

Low rates of ongoing rheumatologist care were observed. This is consistent with our previous work evaluating the consistency of rheumatologist care for RA in BC, which showed that patients under rheumatologist care (defined as in the current study) had low rates of yearly follow-up (only 34% had yearly follow-up over 5 years, and only 30% over 9 years) (14). Furthermore, low consistency in

rheumatologist care was associated with lower DMARD use (22% in patients without rheumatologist care in the preceding 5 years compared to 92% in those under continuous care) (14). Declining follow-up over time was also seen in a longitudinal study of an early RA cohort in which follow-up rates declined from 100% in 2008 to 85% in 2015 (8). The higher follow-up rates may be due to follow-up protocol (8) and participation bias. In both studies, rates of yearly visits declined with increasing length of follow-up since the first rheumatology visit. The measure highlights a gap in rheumatology care that can be tracked over time and monitored following interventions to improve follow-up.

Rates of DMARD use varied based on whether patients saw rheumatologists and visit timing in relation to measurement year. Rates were low in the entire RA population (35–39% across measurement years). Rates improved but remained suboptimal for patients seeing rheumatologists ever in follow-up or in the past (56–63%) and were highest (87% in 2014) for patients under active rheumatologist care (i.e., ≥1 visit in the measurement year). Similar findings were previously reported in BC over 1996–2000, with only 43% of the entire RA population using a DMARD over 5 years, and 31% over 1 year, versus 76% for patients under continuous rheumatologist care (10). In Ontario, 67% of seniors with RA (age ≥65 years) under care of a rheumatologist (≥1 visit) received treatment with DMARDs in 2006 compared to 21% with no rheumatology care (25). Similarly, high DMARD rates (87–95%) were previously reported in the early RA rheumatology cohort (8) and in the Rheumatology Informatics System for Effectiveness (RISE) registry (91%) (26), indicating that the measure may be less useful for quality improvement among rheumatologists and more useful for health system monitoring. We did not examine prednisone use in patients who were not treated with DMARDs; however, this could serve as a marker of unmet needs and poor quality of care.

In contrast to the rate of DMARD use, timeliness for those who received DMARDs was better and improved over the study period. Nonetheless, only 34% of incident RA patients seeing rheumatologists during follow-up met the 14-day benchmark in 2009, and this only improved to 43% when a 30-day benchmark was used. Time from RA diagnosis to starting DMARDs improved over calendar years, with the median time in 2009 (23 days) approaching the 14-day benchmark. However, caution is needed in interpreting results as indicating worse performance in earlier calendar years given the length time bias, as described below.

Lessons were learned from operationalizing the PMs using administrative data. Choosing the appropriate denominator is challenging. Trying to avoid selection bias has to be balanced with identifying the sample influenced by the process of care being evaluated or under the control of the specialist whose performance is being assessed. Because both perspectives are relevant, depending on evaluation purpose (e.g., public health policy versus rheumatologist performance), we reported the PMs on the entire RA population (the least selection bias) and in samples seen by rheumatologists. In the latter, the timing of rheumatologist visits in relation to measurement year was relevant and led to operationalization with different denominators. Length time bias is another important issue when comparing performance across calendar years. Longer follow-up times available in earlier years for a process of care to occur can increase the likelihood of entering the numerator (improving performance) or the denominator (worsening performance) or increase the median time to an event. Hence, there is a need to choose denominator definitions equalizing follow-up time, especially when evaluating timeliness of care (e.g., PMs 1 and 4).

While this study represents the first comprehensive evaluation of these AAC system-level PMs using population-based data and highlights a number of important findings, a number of limitations need to be outlined, namely those inherent to administrative data. Identification of rheumatologists using administrative data may have been incomplete due to some rheumatologists using internist fee codes, leading to underreporting of the first 2 measures. However, the estimate is consistent with our knowledge of rheumatologists in BC. Additionally, it is also possible that incomplete or inaccurate data on patient registration could have impacted measure results. While we used a valid, previously published RA case definition, it is possible that some of the identified patients did not have RA. Additionally, the use of exclusion criteria applied over 5 years reduced the ability to report on measures in more recent years, as did the lag in the availability of administrative data. This also makes it challenging to use these results for real-time quality improvement. However, this work is still valuable for monitoring practice trends over time.

In conclusion, the results of this study will inform further reporting on PMs nationally and help serve in benchmarking when planning quality improvement and advocacy work. Additionally, future work will examine the predictors of measure performance, including geographic variation in performance and also patient outcomes associated with high versus low levels of measure performance, to better understand the implications of measuring performance and to facilitate the introduction of the measures into public reporting and health policy decision-making. Timely communication of performance at the practice level could be used to influence clinical care, and at the provincial level, it could inform health policy.

ACKNOWLEDGMENTS

We thank the Ministry of Health of British Columbia and Population Data BC for providing access to the administrative data.

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. Barber 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. Barber, Marshall, Szefer, Barnabe, Shiff, Bykerk, Homik, Thorne, Ahluwalia, Benseler, Mosher, Twilt, Lacaille.

Acquisition of data. Lacaille.
Analysis and interpretation of data. Barber, Szefer, Lacaille.

REFERENCES

1. O'Donnell S, McRae L, Toews J, Pelletier L. CDDSS Arthritis Working Group. National Surveillance of Arthritis in Canada, results from the Canadian Chronic Disease Surveillance System (CCDSS) [abstract]. J Rheumatol 2019;46:794–95.
2. Public Health Agency of Canada. Canadian Chronic Disease Surveillance System (CCDSS). 2018. URL: https://infobase.phac-aspc.gc.ca/CCDSS-SCSMC/data-tool/.
3. Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheumatol 2016;68:1–26.
4. Bykerk VP, Akhavan P, Hazlewood GS, Schieir O, Dooley A, Haraoui B, et al. Canadian Rheumatology Association recommendations for pharmacological management of rheumatoid arthritis with traditional and biologic disease-modifying antirheumatic drugs. J Rheumatol 2012;39:1559–82.
5. Smolen JS, Landewe 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.
6. Arthritis Alliance of Canada. URL: http://arthritisalliance.ca/en/.
7. Barber CE, Marshall DA, Mosher DP, Akhavan P, Tucker L, Houghton K, et al. Development of system-level performance measures for evaluation of models of care for inflammatory arthritis in Canada. J Rheumatol 2016;43:530–40.
8. Barber CE, Schieir O, Lacaille D, Marshall DA, Barnabe C, Hazlewood G, et al. High adherence to system-level performance measures for rheumatoid arthritis in a national early arthritis cohort over eight years. Arthritis Care Res (Hoboken) 2018;70:842–50.
9. Barber CE, Thorne JC, Ahluwalia V, Burt J, Lacaille D, Marshall DA, et al. Feasibility of measurement and adherence to system performance measures for rheumatoid arthritis in 5 models of care. J Rheumatol 2018;45:1501–8.
10. Lacaille D, Anis AH, Guh DP, Esdaile JM. Gaps in care for rheumatoid arthritis: a population study. Arthritis Rheum 2005;53:241–8.
11. Population Data BC, for the British Columbia Ministry of Health. Medical Services Plan payment information. 2016. URL: https://www.popdata.bc.ca/data/health/msp.
12. Population Data BC, for the British Columbia Ministry of Health. Consolidation file (MSP registration and premium billing) data set. 2016. URL: https://www.popdata.bc.ca/data/population/consolidationfile.
13. Population Data BC, for the British Columbia Ministry of Health. PharmaNet data set. 2016. URL: https://www.popdata.bc.ca/data/health/pharmanet.
14. Lacaille D, Lipson R, Thompson D. Rheumatologist care in rheumatoid arthritis: are we dropping the ball? J Rheumatol 2015;42: 1264–65.
15. Antoniou T, Zagorski B, Loutfy MR, Strike C, Glazier RH. Validation of case-finding algorithms derived from administrative data for identifying adults living with human immunodeficiency virus infection. PLoS One 2011;6:e21748.
16. Lard LR, Visser H, Speyer I, vander Horst-Bruinsma IE, Zwinderman AH, Breedveld FC, et al. Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. Am J Med 2001;111: 446–51.
17. Möttönen T, Hannonen P, Korpela M, Nissilä M, Kautiainen H, Ilonen J, et al. Delay to institution of therapy and induction of remission using single-drug or combination–disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. Arthritis Rheum 2002;46: 894–8.
18. Nell VP, Machold KP, Eberl G, Stamm TA, Uffmann M, Smolen JS. Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis. Rheumatology (Oxford) 2004;43:906–14.
19. Finckh A, Liang MH, van Herckenrode CM, de Pablo P. Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: a meta-analysis. Arthritis Rheum 2006;55:864–72.
20. British Columbia Ministries of Health. Building the vision in British Columbia: an update from the Ministries of Health. 2002. URL: https://www.health.gov.bc.ca/library/publications/year/2002/cdm/updatedec02.pdf.
21. Canadian Institute for Health Information. Scott's medical database metadata. 2017. URL: https://www.cihi.ca/en/scotts-medical-database-metadata.
22. Feldman DE, Bernatsky S, Haggerty J, Leffondré K, Tousignant P, Roy Y, et al. Delay in consultation with specialists for persons with suspected new-onset rheumatoid arthritis: a population-based study. Arthritis Rheum 2007;57:1419–25.
23. Widdifield J, Paterson JM, Bernatsky S, Tu K, Thorne JC, Ivers N, et al. Access to rheumatologists among patients with newly diagnosed rheumatoid arthritis in a Canadian universal public healthcare system. BMJ Open 2014;4:e003888.
24. Widdifield J, Tu K, Thorne JC, Bombardier C, Paterson MJ, Jaakkimainen RL, et al. Patterns of care among patients referred to rheumatologists in Ontario, Canada. Arthritis Care Res (Hoboken) 2017; 69:104–14.
25. Widdifield J, Bernatsky S, Paterson JM, Thorne JC, Cividino A, Pope J, et al. Quality care in seniors with new-onset rheumatoid arthritis: a Canadian perspective. Arthritis Care Res (Hoboken) 2011;63:53–7.
26. Yazdany J, Bansback N, Clowse M, Collier D, Law K, Liao KP, et al. Rheumatology Informatics System for Effectiveness: a national informatics-enabled registry for quality improvement. Arthritis Care Res (Hoboken) 2016;68:1866–73.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 649–657
DOI 10.1002/acr.24163

American College of Rheumatology
Empowering Rheumatology Professionals

Outcomes Reported in Prospective Long-Term Observational Studies and Registries of Patients With Rheumatoid Arthritis Worldwide: An Outcome Measures in Rheumatology Systematic Review

Maria A. Lopez-Olivo,[1] Richard J. Zogala,[1] Jude des Bordes,[1] Natalia V. Zamora,[1] Robin Christensen,[2] Davesh Rai,[1] Niti Goel,[3] Loreto Carmona,[4] Gregory Pratt,[1] Vibeke Strand,[5] and Maria E. Suarez-Almazor[1]

Objective. Prospective long-term observational studies (LOS) in rheumatoid arthritis (RA) lack a core set of universally collected outcome measures, particularly patient-centered outcomes, precluding accurate comparisons across studies. Our aim was to identify long-term outcome measures collected and reported in these studies.

Methods. We conducted a systematic review of registries and LOS of patients with RA, searching in ClinicalTrials.gov, the Agency for Healthcare Research and Quality Registry of Patient Registries, and Google Scholar. The names and acronyms of registries and LOS were further searched in the Medline and Embase databases to retrieve published articles. Two independent reviewers undertook data collection, quality appraisal, and data extraction.

Results. We identified 88 registries/LOS that met our eligibility criteria. These were divided into 2 groups: disease-based (52 [59%]) and therapy-based (36 [41%]). Methodologic and reporting standards varied across the eligible studies. For clinical outcomes, disease activity was recorded in 88 (100%) of all LOS/registries. The most commonly reported measure (86 [98%]) was the composite outcome Disease Activity Score using 28 joints. Of the patient-centered outcomes collected, physical functioning was most frequently reported (75 [85%]) with the Health Assessment Questionnaire (75 [85%]) as the most commonly used instrument within this domain. Other domains of patient-centered outcomes were comparatively infrequently recorded: mental (29 [33%]), social (20 [23%]), and health-related quality of life (37 [42%]).

Conclusion. Most registries/LOS collect measures of disease activity and physical function. However, there is substantial heterogeneity in the collection of relevant patient-centered outcomes that measure symptom burden and mental and social ramifications of RA.

INTRODUCTION

Over the past few decades, there has been growing interest and need for prospective long-term observational studies (LOS) and registries pertaining to rheumatoid arthritis (RA). Randomized controlled trials (RCTs) are generally held in the highest esteem because they are likely to provide the best evidence for causality. However, they most often focus on addressing one specific question, and unless they are designed as community-based pragmatic trials, they may not provide real-world data. The strict inclusion and exclusion criteria of RCTs ensure internal validity but can lead to uncertainty about generalizability. In addition, RCTs provide information on the efficacy of therapies for RA in the shorter term but may not be ideal to address longer-term effectiveness. Prospective LOS and patient registries can address questions about long-term effectiveness and collect multiple outcomes as well as rare adverse events associated with therapy, which is typically not feasible in RCTs. Numerous RA cohorts and registries around the world are collecting longitudinal data to complement evidence obtained from RCTs. A few studies examining the features of

The Parker Institute, Bispebjerg and Frederiksberg Hospital, was supported by the Oak Foundation (core grant OCAY-13-309).

[1]Maria A. Lopez-Olivo, MD, PhD, Richard J. Zogala, MD, Jude des Bordes, MBChB, DrPH, Natalia V. Zamora, MD, Davesh Rai, MD, Gregory Pratt, DDS, MLS, Maria E. Suarez-Almazor, MD, PhD: University of Texas MD Anderson Cancer Center, Houston; [2]Robin Christensen, MSc, PhD: Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark, and University of Southern Denmark and Odense University Hospital, Odense, Denmark; [3]Niti Goel, MD, PhD: Duke University School of Medicine, Durham, North Carolina; [4]Loreto Carmona, MD, PhD: Instituto de Salud Musculoesquéletica, Madrid, Spain; [5]Vibeke Strand, MD, MACR, FACP: Stanford University School of Medicine, Palo Alto, California.

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

Address correspondence to Maria E. Suarez-Almazor, MD, PhD, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 1467, Houston, TX 77030. Email: MSalmazor@mdanderson.org.

Submitted for publication July 1, 2019; accepted in revised form January 28, 2020.

SIGNIFICANCE & INNOVATIONS

- We identified 88 prospective long-term observational studies and registries across the world reporting outcomes in patients with rheumatoid arthritis.
- Globally, there is significant heterogeneity of collected and reported outcomes across observational studies and registries, varying according to the type of registry (i.e., disease-based versus therapy-based).
- Patient-centered outcomes measuring symptom burden and mental and social aspects of disease are not consistently collected and/or reported.

selected registries/LOS in RA have found significant heterogeneity in the outcomes collected, creating challenges in the comparability of findings across studies (1–3). Although there have been efforts to reduce the variability in data collection and analysis, a well-defined and universally accepted core set of outcomes to be measured in LOS that includes important patient-centered domains with specific relevance to long-term outcomes has yet to be agreed upon (4–6). The European Alliance of Associations for Rheumatology proposed a core set that primarily included pathophysiologic measures. Although they recommended measuring quality of life and function, specific subdomains were not proposed (5). Barber et al in Canada also proposed a core set of measures to be collected in clinical practice to improve quality of care, rather than for longitudinal outcome studies (6).

Using the Strengthening the Reporting of Observational Studies in Epidemiology statement, a previous study reviewed registries and cohorts of RA patients receiving biologic therapy in the US and Europe to compare differences in study design and methods that may explain heterogeneous results (1). The review compared methodologic domains such as recruitment methods and inclusion data among selected therapy-based registries. However, only selected clinical outcomes could be evaluated due to the heterogeneity of outcomes collected, and no data on outcomes potentially important to patients (e.g., fatigue, sleep, productivity) were assessed except for physical function and health-related quality of life (HRQoL).

In 2017, a European Alliance of Associations for Rheumatology task force agreed upon a set of 21 core set domains and instruments for observational studies in RA (5). Many domains important to patients, such as productivity, social engagement, and survival, were not included as core measures, but merely as desirable or complementary. Furthermore, how RA patients view the importance and relevance of outcomes reported in studies is not clear (7,8). The Outcome Measures in Rheumatology (OMERACT) initiative has designed and implemented strategies to develop and validate outcomes to be reported in rheumatic diseases such as RA, and continues to be a significant driving force behind this effort (9). OMERACT relies on the inclusion of the patient's voice in discussions regarding the relevance and appropriateness of outcome measures, recognizing and including the patient perspective (10). Although designed for use in RCTs and LOS, most of the RA outcome measures agreed upon in OMERACT have been adopted in the setting of an RCT or short-term studies. A wide consensus on what outcomes, especially patient-centered outcomes, should be collected in RA registries has yet to be reached.

To build upon the interests of research groups, a first step is to identify outcome domains and measures, including patient-centered outcomes, currently collected in long-term studies of RA patients. We therefore conducted a systematic review of registries/LOS of patients with RA, primarily evaluating data collection and reporting patient-centered outcomes.

MATERIALS AND METHODS

Eligibility criteria. We included both registries and prospective LOS. Although the distinction is not always clear, registries are generally considered to be databases with ongoing longitudinal data collection of individual patients, with data not necessarily collected to answer specific research questions; they are often population-based (11). In contrast, LOS usually include patients in specific settings and often aim to answer defined research questions.

To be included in our review, registries/LOS had to include patients with RA, assess outcomes or prognosis, include clinical outcomes or patient-centered outcomes in their data collection (we used the definitions and concepts provided by the Patient-Centered Outcomes Research Institute) (12), and have appeared in at least 1 publication written in English since 2013. Registries/LOS were excluded if they were an open label extension of a clinical trial, the purpose of the registry was to answer a particular question unrelated to clinical, patient-centered, or safety outcomes (e.g., biomarkers, lifestyle habits), or entry into the registry was limited to those with a specific articular or extraarticular manifestation of RA (e.g., anemia) or a certain study subpopulation (e.g., those with interstitial lung disease).

Registry identification and selection. Our search strategy started with a Google Scholar search using the names of the 193 United Nations member states as listed on the organization's website (13). Then, for each member state, the name was combined with the keywords "rheumatoid" and "registry," and the first 15 results were selected for review. To help reduce false hits in the search results, we searched multiple-word names as a phrase (e.g., "Marshall Islands"), and a few single-word names were searched within quotation marks (e.g., "Niger" to avoid retrieving "Nigeria"). Some member states were searched using both the formal and common names (e.g., "Côte D'Ivoire" and "Ivory Coast"), and others were searched using only a simplified name (e.g., "Bolivia" rather than "The Plurinational State of Bolivia"). We also searched

2 databases of registries: the Agency for Healthcare Research and Quality (AHRQ) Registry of Patient Registries (RoPR) and ClinicalTrials.gov, using terms related to RA and registries. Additional handsearching was performed for identified registries/LOS when URLs were not readily available through the previous searches. This search strategy resulted in 2,996 URLs, of which 2,766 were excluded as per Figure 1, leaving 230 URLs corresponding to 88 unique registries/LOS. The decision to include an LOS/registry in the review was made by 2 independent pairs of reviewers (either RJZ and JdB or NVZ and DR). Consensus was reached by discussion or third-party adjudication (MES-A).

Data collection process. Data sources to retrieve information from the selected registries/LOS included websites and publications in the medical literature. For each registry included in the review, we identified public websites and the corresponding URLs. Initial sources of data, when available, included information from the websites or in the databases of registries (RoPR and ClinicalTrials.gov). For the next step, an expert health sciences librarian (GP) conducted searches in the Medline and Embase databases (via the Ovid platform) using the names and acronyms of the identified registries/LOS for all publications until August 2018. When retrieval was sparse, proximity operators were used in the search strings for names of registries. When available, ClinicalTrials.gov identifiers were also included. For example, literature involving the Consortium of Early Arthritis Cohorts USA was searched in all fields of database records using (Consortium adj3 "Early Arthritis" adj3 Cohort* or "CATCH US" or CATCH-US or NCT02386527.af.) To further identify relevant citations, those retrieved were cross-referenced with subject heading (National Library of Medicine [MeSH] or Embase [Emtree]) terms related to RA and registry or cohort keywords. Preference was given to literature describing the registries/cohorts themselves and to citations published from January 2013 to August 2018.

Publications related to each registry were compiled in EndNote (Clarivate Analytics), and all citations were grouped by registry. Only English language publications were reviewed. To extract variables of interest, 2 reviewers (RJZ and JdB) independently examined the websites, databases of registries/LOS, and full-text publications related to each registry. For scientific publications, we

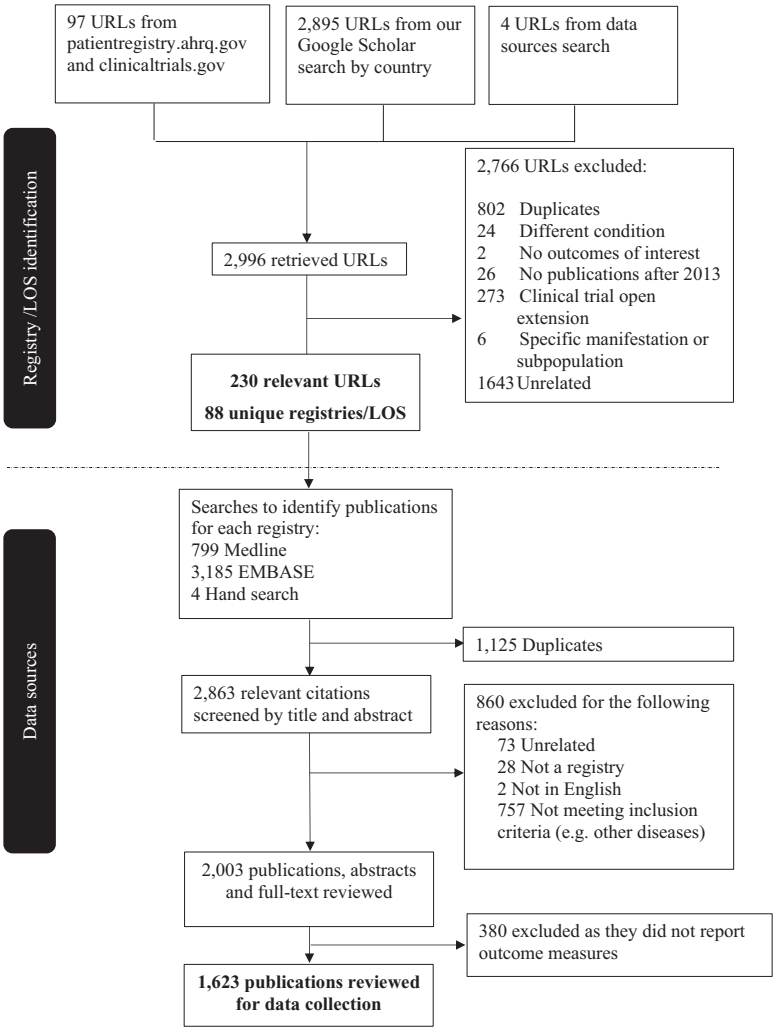


Figure 1. Flow diagram of the identification and selection process. LOS = long-term observational studies.

extracted variables of interest from the Methods and Results sections. Disagreements regarding the data collected were resolved by consensus or third-party adjudication (MES-A).

We collected general registry information, including country, types of patients included, and the purpose of the registry/LOS when specified from all sources available, including websites and publications. Because we were primarily interested in patient-centered outcomes, we evaluated documentation of specific outcome domains and sociodemographic data broadly, covering commonly identified risk factors for RA outcomes, clinical outcomes, and patient-centered outcomes, primarily patient-reported outcomes. These included: socioeconomic status (e.g., education and income); comorbidities, including smoking; rheumatoid factor and/or anti–citrullinated protein antibody levels; clinical outcomes (e.g., radiographic evaluation and clinician-based disease activity indices); safety outcomes (e.g., serious adverse events [SAEs] and death); and patient-centered outcomes (e.g., measures of physical function or HRQoL, as well as assessments of symptom burden, such as pain, fatigue, stiffness, sleep, mental anguish [e.g., depression, anxiety], and social participation [e.g., working status]).

Quality appraisal. Two pairs of reviewers independently appraised the registries/LOS (either RJZ and JdB or NVZ and DR). Disagreements were resolved by consensus or third-party adjudication (MES-A). To appraise the quality of each registry, we used a guide developed by AHRQ (14) that includes the following items: 1) planning (written registry protocol with goals, a defined target population, specific methods for collecting information, and appropriate personnel and storage of data); 2) design (appropriate review of the literature, description of the target population, defined inclusion and exclusion criteria of patients, and estimated follow-up time); 3) data elements and resources (including appropriate and validated scales for assessing outcomes); and 4) ethics (including protection of human subjects such as privacy and informed consent, and review and approval by oversight committees).

Synthesis of results. Characteristics and reported variables were summarized overall and by type of registry/LOS (disease-based or therapy-based). Descriptive statistics were used to synthesize the data collected; unweighted frequencies and percentages were used for categorical variables.

RESULTS

We identified 97 URLs for registries/LOS from RoPR and ClinicalTrials.gov. The Google Scholar search identified 2,895 URLs, and an additional 4 were identified through handsearching. After cross-referencing and selection, we included 230 relevant URLs corresponding to 88 registries/LOS (Figure 1 and Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24163/abstract). One of these was a collaboration among registries (not a registry on its own), which we decided to include because it provided data from registries with no individual data from different countries (15).

Table 1. Characteristics of the registries and prospective long-term observational studies included in our analysis*

Characteristic	Disease-based (n = 52)	Therapy-based (n = 36)	Total (n = 88)
Location			
Asia	9 (17)	5 (14)	14 (16)
Europe	29 (56)	24 (67)	53 (60)
North America	9 (17)	3 (8)	12 (14)
South America	3 (6)	3 (8)	6 (7)
Oceania	1 (2)	–	1 (1)
International	1 (2)	1 (3)	2 (2)
Data source			
Publications	52 (100)	36 (100)	88 (100)
Website	22 (42)	9 (25)	31 (35)
ClinicalTrials.gov	9 (17)	7 (19)	16 (18)
Patients with only RA	39 (75)	14 (39)	53 (60)

* Values are the number (%). RA = rheumatoid arthritis.

To ascertain outcome measures, in addition to reviewing websites, we conducted a publication search that yielded 2,863 publications after deduplication (Figure 1). Titles and abstracts of 2,863 publications were reviewed; 860 of these articles were excluded for reasons detailed in Figure 1. The full text of the remaining 2,003 publications was reviewed, and an additional 380 were excluded because they did not report outcome measures, leaving 1,623 publications eligible for review.

Characteristics of registries/LOS. Table 1 shows the characteristics of the included registries/LOS. Of the 88, 52 (59%) were disease-based, and 36 (41%) were therapy-based. The origin of the registries/LOS included 36 different countries across South America, North America, Asia, Oceania, and Europe, with most originating from the US. Registries/LOS primarily included patients with RA; however, 34 (13 disease-based and 21 therapy-based) also included patients with diseases other than RA, including myositis, systemic lupus erythematosus, ankylosing spondylitis, psoriatic arthritis, osteoporosis, osteoarthritis, gout, fibromyalgia, juvenile idiopathic arthritis, Crohn's disease, and ulcerative colitis.

Quality assessment. Data elements and resources was the most frequent quality domain in compliance with guidelines in both types of registries/LOS: 39 (75%) for disease-based, and 29 (81%) for therapy-based (Table 2). Planning and design were the

Table 2. Compliance with AHRQ quality domains (ref. 14)*

Registry type	Planning	Design	Data elements and resources	Ethics
Disease-based (n = 52)	7 (13)	10 (19)	39 (75)	32 (62)
Therapy-based (n = 36)	8 (22)	8 (22)	29 (81)	14 (39)
Total (n = 88)	15 (17)	18 (20)	68 (77)	46 (52)

* Values are the number (%). AHRQ = Agency for Healthcare Research and Quality; ref. = reference.

least frequently reported domains: 7 (13%) in disease-based and 8 (22%) in therapy-based for planning, and 10 (19%) in disease-based and 8 (22%) in therapy-based for design.

Data collected. A summary of the data collected in the registries/LOS is shown in Tables 3 and 4. We evaluated a total of 43 outcomes, including clinical outcomes (disease activity, imaging, safety), and patient-centered outcomes comprising 3 domains: physical, mental, and social wellbeing, as well as HRQoL. The mean ± SD of outcomes collected by disease-based registries/LOS was 12.1 ± 5.0 and by therapy-based registries was 10.1 ± 4.0. The mean ± SD of patient-centered outcomes collected by disease-based registries was 4.8 ± 3.0 and by therapy-based registries was 2.4 ± 2.2.

Sociodemographic data and risk factors. All registries/LOS reported data collection of sociodemographic data and risk factors. The collection of serologic markers, including rheumatoid factor and anti–cyclic citrullinated peptide (73 [83%]), was the most frequently reported variable, with comorbidities (72 [82%]) being the second most frequently reported variable. Alcohol consumption was reported in 10 (11%) of registries/LOS, whereas smoking was reported in 63 (73%).

Clinical outcomes. All registries/LOS reported the collection of at least 1 disease activity measure or composite index. The most commonly reported disease activity outcome was the Disease Activity Score using 28 joints (DAS28) or 1 of its versions (86 [98%]) (16). Other indices reported included the American College of Rheumatology 20% improvement criteria (10 [11%]) (17,18), the Simplified Disease Activity Index (35 [40%]) (19), the Clinical Disease Activity Index (38 [43%]) (20), and the physician global assessment (37 [42%]). Other markers of disease activity included erythrocyte sedimentation rate, C-reactive protein level, and tender and/or swollen joint counts. Patient global assessments of disease activity scores were reported in 40 (77%) and 26 (2%) of condition- and therapy-based registries/LOS, respectively. More disease-based than therapy-based registries reported imaging data: 37 (71%) versus 18 (50%). The most common instrument score reported was the Sharp/van der Heijde score (21,22).

Table 3. Variables and outcomes reported in the registries and prospective long-term observational studies in our analysis*

	Disease-based (n = 52)	Therapy-based (n = 36)	Total (n = 88)
Sociodemographic data			
Education status	21 (40)	7 (19)	28 (32)
Income	16 (31)	–	16 (18)
Financial measures (nonspecified)	17 (33)	6 (17)	23 (26)
Lifestyle factors			
Alcohol	9 (17)	1 (3)	10 (11)
Smoking	40 (77)	24 (67)	64 (73)
Clinical characteristics			
RF/ACPA	46 (88)	27 (75)	73 (83)
Comorbidities	41 (79)	31 (86)	72 (82)
Disease activity	52 (100)	36 (100)	88 (100)
DAS28/DAS28-CRP/ESR	50 (96)	36 (100)	86 (98)
DAS28-CRP	17 (33)	10 (28)	27 (31)
DAS28-ESR	11 (21)	11 (31)	22 (25)
ACR20	6 (12)	4 (11)	10 (11)
SDAI	23 (44)	12 (33)	35 (40)
CDAI	21 (40)	17 (47)	38 (43)
Patient global assessment	40 (77)	26 (72)	66 (75)
Physician global assessment	26 (50)	11 (31)	37 (42)
RADAI	9 (17)	1 (3)	10 (11)
RAID	1 (2)	4 (11)	5 (6)
RAPID	12 (23)	4 (11)	16 (18)
Drug safety			
Serious adverse events	17 (33)	36 (100)	53 (60)
Deaths	16 (31)	17 (47)	33 (38)
Imaging	37 (71)	18 (50)	55 (63)

* Values are the number (%). ACPA = anti–citrullinated protein antibody; ACR20 = American College of Rheumatology 20% improvement criteria (refs. 17,18); CDAI = Clinical Disease Activity Index (ref. 20); DAS28 = Disease Activity Score using 28 joints (ref. 16); DAS28-CRP = DAS28 using the C-reactive protein level (ref. 31); DAS28-ESR = DAS28 using the erythrocyte sedimentation rate (ref. 32); RADAI = Rheumatoid Arthritis Disease Activity Index; RAID = Rheumatoid Arthritis Impact of Disease (ref. 33); RAPID = Routine Assessment of Patient Index Data (ref. 34); RF = rheumatoid factor; SDAI = Simplified Disease Activity Index (ref. 19).

Table 4. The frequencies of patient-centered outcomes*

	Disease-based (n = 52)	Therapy-based (n = 36)	Total (n = 88)
Health-related quality of life	27 (52)	10 (28)	37 (42)
EQ-5D	19 (37)	7 (19)	26 (30)
SF-6D	3 (6)	2 (6)	5 (6)
SF-36	16 (31)	8 (22)	24 (27)
SF-12	–	2 (4)	2 (2)
AIMS2	–	1 (2)	1 (1)
RAQoL	2 (4)	–	2 (2)
EQ-VAS	3 (6)	–	3 (3)
PROMIS-29	1 (2)	2 (6)	3 (3)
Physical domain	48 (92)	28 (78)	76 (86)
Function	48 (92)	27 (75)	75 (85)
HAQ†	48 (92)	27 (75)	75 (85)
HAQ DI	45 (87)	25 (69)	70 (80)
MDHAQ	9 (17)	4 (11)	13 (15)
HAQ-II	5 (10)	1 (3)	6 (7)
PAS-II	1 (2)	–	1 (1)
FFbH	–	1 (3)	1 (1)
VAS function	8 (15)	–	8 (9)
Symptom burden	46 (88)	19 (53)	65 (74)
Pain	43 (83)	18 (50)	61 (69)
Sleep	12 (23)	–	12 (14)
Fatigue	22 (42)	9 (25)	31 (35)
Stiffness	23 (44)	6 (17)	29 (33)
Mental domain	23 (44)	6 (17)	29 (33)
Depression	19 (37)	6 (17)	25 (28)
Anxiety	8 (15)	1 (3)	9 (10)
Fear	2 (4)	–	2 (2)
Coping	2 (4)	1 (3)	3 (3)
Helplessness	1 (2)	–	1 (1)
Social domain	12 (23)	8 (22)	20 (23)
Working status	9 (17)	6 (17)	15 (17)
WPAI	3 (6)	3 (8)	6 (7)

* Values are the number (%). AIMS2 = Arthritis Impact Measurement Scale 2 (ref. 40); EQ-5D = EuroQol 5-domain questionnaire (ref. 24); EQ-VAS = EuroQol visual analog scale (ref. 42); FFbH = Funktionsfragebogen Hannover; HAQ = Health Assessment Questionnaire; HAQ DI = HAQ disability index (ref. 23); MDHAQ = multidimensional HAQ (ref. 35); PAS-II = Patient Activity Scale II (ref. 37); PROMIS-29 = Patient-Reported Outcomes Measurement Information System 29; RAQoL = Rheumatoid Arthritis Quality of Life questionnaire (ref. 41); SF-6D = Short Form 6 dimensions (ref. 38); SF-12 = SF 12-item questionnaire (ref. 39); SF-36 = SF 36-item questionnaire (ref. 25); VAS = visual analog scale; WPAI = Work Productivity and Activity Impairment questionnaire (ref. 43).
† Three forms of HAQ collected included HAQ DI, MDHAQ, and HAQ II (ref. 36).

Safety. We assessed the reporting of SAEs and deaths. Only one-third of disease-based registries reported the collection of SAEs, compared with all therapy-based registries. Death was also less frequently reported as being recorded in the disease-based than in the therapy-based registries: 16 (31%) compared with 17 (47%).

Patient-centered outcomes. We examined patient-centered outcomes because they pertain to HRQoL and its 3 major domains: physical (comprising function and symptom burden), mental, and social. Among different Health Assessment Questionnaire (HAQ) scales used, the HAQ disability index (HAQ DI) (23) was reported in 45 (87%) of the disease-based and 25 (69%) of the therapy-based registries. Among the 8 different scales used to assess HRQoL, the EuroQol 5-domain questionnaire (24) was most commonly reported in disease-based registries (19 [37%]). In therapy-based registries, the Medical Outcomes Study Short Form 36-item questionnaire (25) was reported most frequently, used by 8 (22%) of these registries.

Registries/LOS reported collection of patient-centered outcomes pertaining to the physical domain (76 [86%]) more frequently than either the mental (29 [33%]) or social (20 [23%]) domains. Within the physical domain, outcomes related to function (disease-based 48 [92%], therapy-based 27 [75%]) were collected more frequently than symptom burden (disease-based 46 [88%], therapy-based 19 [53%]). The mental domain was reported least frequently in the therapy-based registries/LOS (6 [17%]) when compared to all other patient-centered outcomes. The reporting of outcomes in the social domain was similar for disease-based and therapy-based registries/LOS, 12 (23%) versus 8 (22%), respectively (Table 4). The most frequently reported outcome within the subgroup of function was HAQ DI, where 45 (87%) of disease-based and 25 (69%) of therapy-based registries/LOS reported collection. The most frequently reported outcome within the subgroup of symptom burden was pain, which was collected in 43 (83%) of disease-based and 18 (50%) of therapy-based registries/LOS. The most frequently reported outcome within the mental domain was depression, with 19 (37%) of disease-based and 6 (17%) of therapy-based registries/LOS reporting collection. Patient-centered outcomes identified in our search relating to the social domain were working status (disease-based 9 [17%], therapy-based 6 [17%]) and the Work Productivity and Activity Impairment questionnaire (disease-based 3 [6%], therapy-based 3 [8%]) (Table 4). Considerable heterogeneity across registries within close geographic proximity was observed with respect to the types of instruments used to measure disease activity, HRQoL, and patient-centered outcomes (data not shown).

DISCUSSION

This study was the first step from an OMERACT initiative to identify outcome domains and measures, including patient-centered outcomes, currently collected in long-term studies of RA patients. We therefore conducted one of the largest systematic reviews scrutinizing the data collection of RA registries/LOS worldwide. We found substantial heterogeneity in the collected outcome measures and variability in the instruments used to define these outcomes. We inferred the perceived importance of these variables by how frequently they were reported. For the purpose of our study, we divided registries into 2 groups: disease-based and therapy-based.

We observed differences in the quality domains between the types of registries. Planning and design were more commonly reported in the therapy-based than in the disease-based

registries. This finding is likely because data collection surrounding drug administration is more stringently regulated. Regarding the data collected, disease-based registries collected more variables than therapy-based registries. This practice may be because therapy-based registries often focus on pharmacovigilance and adverse events related to therapy, limiting their scope of collection. A therapy-based registry focus on adverse-events related to therapy explains why safety data annotating serious adverse reactions were recorded in 17 (33%) of the disease-based registries versus 36 (100%) of the therapy-based registries. Imaging data were recorded in 55 (63%) of all registries, with a higher proportion of disease-based than therapy-based registries collecting this information. These data may have been omitted from the therapy-based registries because the efficacy of the drug had already been proven in an RCT. Given the additional expense of imaging, the administrators of the therapy-based registries may have considered it unnecessary. Conceivably, given the fact that imaging is a surrogate marker for long-term outcomes, if a registry was collecting alternative long-term outcomes, the collection of imaging data might also have been considered redundant.

More than half of the registries were from European countries. This fact was not surprising because many European countries are under a national health system, allowing for data collection across their populations, and some registries were introduced as a requirement for the pharmacovigilance of biologic agents. However, the type of disease activity, HRQoL, and patient-centered outcomes instruments varied within close geographic regions. Variation in the outcomes instruments used within close geographic regions could be due to an individual registry's conceptualization and provenance occurring independently in a similar time frame, without collaboration at early design and implementation stages.

One or more patient-centered outcomes were collected by the majority of registries. However, the degree of heterogeneity for patient-centered outcomes was considerably greater than that of outcomes related to disease activity. When patient-centered outcomes were divided into domains, we noted significant differences, with those relating to physical function and symptom burden being most frequently reported compared to the mental and social domains. Interestingly, within the social domain we only identified 2 specific reported outcomes outside generic instruments, and both focused on productivity. All other aspects of social participation that are not already captured by HRQoL outcomes were not reported. Other aspects of social participation were not reported, which illuminates the limited emphasis placed on collecting outcomes focused on assessing both the mental and social domains, despite the fact that, over the past decade, there has been increasing pressure from government agencies and the research community to provide a more holistic view of disease (26). There has also been a change in the doctor-patient relationship, moving away from medical paternalism towards shared decision-making, autonomy, and inclusion. This shift does not appear to be well reflected yet in registry design and data collection. Evidence suggests that the use of outcomes that are relevant to patients increases patient satisfaction and improves patient–provider communication, as well as overall patient HRQoL (27). In the clinical setting, patient-centered outcomes may help patients in making informed decisions about their care (by providing results across time and assessing their perspective to these results and treatment) and aid clinicians in monitoring the progress of care.

Our findings are consistent with previous analyses of RA registries. Radner et al (3) evaluated the variables contained in 27 registries across 16 European countries and determined that the most frequently recorded variables were those that described disease activity; DAS28 was reported as being recorded in 100% of the registries. The researchers' work also demonstrated a substantial amount of heterogeneity in collected outcome measures across the European continent (3). Curtis et al (1) performed an analysis of European and American registries, comparing patient characteristics, drug therapy, and adverse events recorded in the various registries. This analysis also reported heterogeneity in the outcomes collected. The authors of that article used a similar strategy to ours of identifying which outcomes were collected by the registries of interest (1).

Registries are created to collect specific, pertinent data. An important purpose of the data collected is to improve patient experience, outcome, and quality of life. However, if data are recorded but never scrutinized, then the purpose of these data may be questioned. Thus, even if data were not captured by our search, the absence of these data from publication lends credence to our salient point that homogeneity in data collection and research is required. In the current review, we did not identify how often data were recorded by each registry. Although the first step in ensuring homogeneity of data collection is to define a core set of outcome measures, ideally the frequency of recording such data will also be standardized.

As the importance of registry data continues to increase, an effort has been made to help define what should be universally recorded by registries (28,29). In 2017, the European Alliance of Associations for Rheumatology outlined 21 variables as the minimum number of data points that should be collected by any registry. This list provides an excellent framework for researchers to determine which outcome measures are important not only to clinicians but also to their patients. However, within this recommendation, there is a relative paucity of guidance on patient-centered outcomes. Of the 21 recommended variables, only 3 are included: HAQ, EuroQol 5-domain questionnaire, and pain (5). We found that 2 of these measures were the most frequently reported (HAQ: 75 [85%], and pain: 61 [64%]). The HAQ might be most frequently reported because functional status is an outcome of relevance for both patients and providers and has been available for decades. Evidence suggests that the HAQ is also a useful monitoring tool that is easily completed by patients in the clinical setting (30).

Some limitations to our study merit further discussion. Unfortunately, there was no single data collection form or method used by all of the registries to assemble the collected variables. For this reason, we relied on information gleaned from published articles, Google Scholar searches, registry websites, ClinicalTrials.gov, and RoPR to determine which variables were recorded by the individual registries. Another limitation of the study is also true for our quality assessment, which was done only using the data reported in articles meeting our eligibility criteria. Studies published before the year 2013 could have reported registry data that were not captured in subsequent publications. Also, given the scope of our study, we were unable to directly contact the administrators of individual registries. Therefore, we cannot ensure that our search was able to fully capture all variables recorded by the individual registries. However, the frequencies we obtained are similar to those noted in previous studies in which registry administrators were contacted (3). In addition, our search strategy included only English publications, and this limitation may have resulted in relevant publications being omitted.

Registries provide pertinent information about the long-term trajectory of disease and disease burden. The current study demonstrates the heterogeneity of collected variables among international registries and indicates that a strategy is needed to reduce the variability of data collection. Further, the study highlights the need for greater emphasis to be placed on the collection of patient-centered outcomes other than physical function and symptom burden. Although we found that patient-centered outcomes that assess physical function (76 [86%]) and symptom burden (65 [74%]) are collected with some regularity, those outcomes that assess other aspects of disease burden were collected inconsistently. Patient-centered outcomes measures provide vital information regarding the patient experience; hopefully, their collection can be routinely incorporated into registry design in the future. Additional studies will be needed to further assess the acceptability and comfort that patients may have answering questions related to psychosocial domains, and whether social desirability factors may impact the completeness and validity of data collection and analysis.

Due to the long-term outlook of registries, outcome measures that are essential in RCTs may not be as important to registries and may not provide the information most relevant to patients. A core set of clearly defined outcomes that are relevant to patients would allow for collaborative research and comparisons across registries and would facilitate analyses that could potentially identify geographic, racial, and cultural differences in disease outcomes.

ACKNOWLEDGMENTS

The authors thank Dr. Karla Criner for her assistance in data collection. They also thank Erica Goodoff, Scientific Editor in the Department of Scientific Publications at The University of Texas MD Anderson Cancer Center, for editing assistance.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Suarez-Almazor 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. Lopez-Olivo, Zogala, Christensen, Goel, Carmona, Strand, Suarez-Almazor.

Acquisition of data. Lopez-Olivo, Zogala, des Bordes, Zamora, Rai, Pratt, Suarez-Almazor.

Analysis and interpretation of data. Lopez-Olivo, Zogala, Christensen, Goel, Carmona, Strand, Suarez-Almazor.

REFERENCES

1. Curtis JR, Jain A, Askling J, Bridges SL Jr, Carmona L, Dixon W, et al. A comparison of patient characteristics and outcomes in selected European and U.S. rheumatoid arthritis registries. Semin Arthritis Rheum 2010;40:2–14.
2. Kearsley-Fleet L, Zavada J, Hetland ML, Nordstrom DC, Aaltonen KJ, Listing J, et al. The EULAR Study Group for Registers and Observational Drug Studies: comparability of the patient case mix in the European biologic disease modifying anti-rheumatic drug registers. Rheumatology (Oxford) 2015;54:1074–9.
3. Radner H, Dixon W, Hyrich K, Askling J. Consistency and utility of data items across European rheumatoid arthritis clinical cohorts and registers. Arthritis Care Res (Hoboken) 2015;67:1219–29.
4. Baillet A, Gossec L, Carmona L, Wit M, van Eijk-Hustings Y, Bertheussen H, et al. Points to consider for reporting, screening for and preventing selected comorbidities in chronic inflammatory rheumatic diseases in daily practice: a EULAR initiative. Ann Rheum Dis 2016;75:965–73.
5. Radner H, Chatzidionysiou K, Nikiphorou E, Gossec L, Hyrich KL, Zabalan C, et al. 2017 EULAR recommendations for a core data set to support observational research and clinical care in rheumatoid arthritis. Ann Rheum Dis 2018;77:476–9.
6. Barber CE, Mosher DP, Ahluwalia V, Zummer M, Marshall DA, Choquette D, et al. Development of a Canadian core clinical dataset to support high-quality care for Canadian patients with rheumatoid arthritis. J Rheumatol 2017;44:1813–22.
7. Dougados M, Nataf H, Steinberg G, Rouanet S, Falissard B. Relative importance of doctor-reported outcomes vs patient-reported outcomes in DMARD intensification for rheumatoid arthritis: the DUO study. Rheumatology (Oxford) 2013;52:391–9.
8. Primdahl J, Jensen DV, Meincke RH, Jensen KV, Ziegler C, Nielsen SW, et al. Patients' views on routine collection of patient-reported outcomes in rheumatology outpatient care: a multicenter focus group study. Arthritis Care Res (Hoboken) 2020;72:1331–8.
9. Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. J Rheumatol Suppl 1994;41:86–9.
10. Hsiao B, Fraenkel L. Incorporating the patient's perspective in outcomes research. Curr Opin Rheumatol 2017;29:144–9.
11. Curbelo RJ, Loza E, de Yebenes MJ, Carmona L. Databases and registers: useful tools for research, no studies. Rheumatol Int 2014;34:447–52.
12. Patient-Centered Outcomes Research Institute. Establishing the definition of patient-centered outcomes research. 2012. URL: https://www.pcori.org/research-results/about-our-research/patient-centered-outcomes-research/establishing-definition.

13. The United Nations. List of member states. URL: http://www.un.org/en/member-states/.
14. Gliklich RE, Dreyer NA, eds. Registries for evaluating patient outcomes: a user's guide. Rockville (MD): Agency for Healthcare Research and Quality; 2014.
15. Chatzidionysiou K, Lie E, Lukina G, Hetland ML, Hauge EM, Pavelka K, et al. Rituximab retreatment in rheumatoid arthritis in a real-life cohort: data from the CERERRA collaboration. J Rheumatol 2017; 44:162–9.
16. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight–joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44–8.
17. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum 1993;36:729–40.
18. American College of Rheumatology Committee to Reevaluate Improvement Criteria. A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. Arthritis Rheum 2007;57:193–202.
19. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology (Oxford) 2003;42:244–57.
20. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther 2005;7:R796–806.
21. Van der Heijde DM, van Riel PL, Nuver-Zwart IH, Gribnau FW, van de Putte LB. Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis. Lancet 1989;1:1036–8.
22. Van der Heijde DM, van Leeuwen MA, van Riel PL, Koster AM, van 't Hof MA, van Rijswijk MH, et al. Biannual radiographic assessments of hands and feet in a three-year prospective followup of patients with early rheumatoid arthritis. Arthritis Rheum 1992;35:26–34.
23. Cole JC, Khanna D, Clements PJ, Seibold JR, Tashkin DP, Paulus HE, et al. Single-factor scoring validation for the Health Assessment Questionnaire-Disability Index (HAQ-DI) in patients with systemic sclerosis and comparison with early rheumatoid arthritis patients. Qual Life Res 2006;15:1383–94.
24. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol 1997;36:551–9.
25. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I: conceptual framework and item selection. Med Care 1992;30:473–83.
26. Weldring T, Smith SM. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). Health Serv Insights 2013;6:61–8.
27. Locklear T, DeBar LL, Willig J, Rundell S, Blackhall L, Zatzick D, et al. Case studies from the clinic: initiating and implementing patient-reported outcome measures. EGEMS (Wash DC) 2017; 5:7.
28. Zavada J, Dixon WG, Askling J. Launch of a checklist for reporting longitudinal observational drug studies in rheumatology: a EULAR extension of STROBE guidelines based on experience from biologics registries. Ann Rheum Dis 2014;73:628.
29. Dixon WG, Carmona L, Finckh A, Hetland ML, Kvien TK, Landewe R, et al. EULAR points to consider when establishing, analysing and reporting safety data of biologics registers in rheumatology. Ann Rheum Dis 2010;69:1596–602.
30. Greenwood MC, Doyle DV, Ensor M. Does the Stanford Health Assessment Questionnaire have potential as a monitoring tool for subjects with rheumatoid arthritis? Ann Rheum Dis 2001;60:344–8.
31. Wells G, Becker JC, Teng J, Dougados M, Schiff M, Smolen J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009;68:954–60.
32. Van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum 1998;41:1845–50.
33. Gossec L, Paternotte S, Aanerud GJ, Balanescu A, Boumpas DT, Carmona L, et al. Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient-derived composite measure of impact of rheumatoid arthritis: a EULAR initiative. Ann Rheum Dis 2011;70:935–42.
34. Pincus T, Swearingen CJ, Bergman M, Yazici Y. RAPID3 (Routine Assessment of Patient Index Data 3), a rheumatoid arthritis index without formal joint counts for routine care: proposed severity categories compared to disease activity score and clinical disease activity index categories. J Rheumatol 2008;35:2136–47.
35. Pincus T, Swearingen C, Wolfe F. Toward a multidimensional Health Assessment Questionnaire (MDHAQ): assessment of advanced activities of daily living and psychological status in the patient-friendly health assessment questionnaire format. Arthritis Rheum 1999;42:2220–30.
36. Wolfe F, Michaud K, Pincus T. Development and validation of the Health Assessment Questionnaire II: a revised version of the Health Assessment Questionnaire. Arthritis Rheum 2004;50:3296–305.
37. Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies, and clinical trials: the patient activity scale (PAS/PAS-II). J Rheumatol 2005;32:2410–5.
38. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002; 21:271–92.
39. Ware J Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220–33.
40. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. AIMS2: the content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. Arthritis Rheum 1992;35:1–10.
41. Tijhuis GJ, de Jong Z, Zwinderman AH, Zuijderduin WM, Jansen LM, Hazes JM, et al. The validity of the Rheumatoid Arthritis Quality of Life (RAQoL) questionnaire. Rheumatology (Oxford) 2001;40:1112–9.
42. Brooks R. EuroQol: the current state of play. Health Policy 1996; 37:53–72.
43. Zhang W, Bansback N, Boonen A, Young A, Singh A, Anis AH. Validity of the work productivity and activity impairment questionnaire: general health version in patients with rheumatoid arthritis. Arthritis Res Ther 2010;12:R177.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 658–665
DOI 10.1002/acr.24214

American College of Rheumatology
Empowering Rheumatology Professionals

Findings on Coronary Angiographies in Patients With Rheumatoid Arthritis and Ischemic Heart Disease: Are They Different From Patients Without Rheumatoid Arthritis?

Marie Holmqvist,[1] Ängla Mantel,[1] Solveig Wållberg-Jonsson,[†] Stefan James,[2] Tomas Jernberg,[3] and Johan Askling[1]

Objective. Patients with rheumatoid arthritis (RA) are at increased risk of coronary artery disease (CAD) and seem to develop more severe acute coronary syndromes (ACS) than the general population. Because few studies have investigated the CAD distribution in the context of acute or stable CAD in RA, the objective was to investigate whether this risk is due to a different distribution and severity of coronary stenoses (versus non-RA), resulting in clinical manifestation of CAD.

Methods. We performed a population-based study using linkages of nationwide clinical, health, and demographics registers. We compared 1 cohort of patients with RA, and 1 matched cohort of patients without RA, undergoing a first coronary angiography from 2006 through 2015. Cardiovascular (CV) characteristics and the presence and distribution of clinically significant stenoses were compared (through odds ratios [ORs]), stratified by indication (stable CAD, ST-elevation myocardial infarction [STEMI], and non–ST-elevation ACS [NSTACS]), using logistic regression.

Results. We identified 2,985 patients with RA and 10,290 patients without RA who underwent a first coronary angiography. A higher proportion of patients with RA (75% versus 69%) had STEMI and NSTACS as indication for angiography. We found no difference in the presence and distribution of clinically significant coronary stenoses in RA compared with the patients without RA, regardless of the CAD type (for having any significant stenosis in stable CAD OR 0.9, STEMI OR 0.8, and NSTACS OR 1.1), stratification by RA duration, sex, or burden of concomitant CV risk factors.

Conclusion. Although RA may accelerate the development of clinical CAD events, the underlying angiographic characteristics are similar to those in patients without RA.

INTRODUCTION

Patients with rheumatoid arthritis (RA) are at increased risk of acute coronary syndrome (ACS). This excess risk cannot be readily explained by traditional cardiovascular (CV) risk factors (1) but points to a role for factors associated with the RA disease (2,3). Yet despite improved RA disease control during recent years, and despite declining rates for ACS in the general population as well as in patients with RA, the excess risk of ACS in RA remains (4).

Besides the elevated incidence, the clinical/phenotypic characteristics of ACS in RA are somewhat different from those of ACS in the general population (5). Studies further suggest inferior clinical outcomes in RA after ACS, including reinfarction and survival (5,6). Whereas the increased ACS incidence in RA might be explained by an increased presence of CV risk factors, and the increased long-term mortality may be explained by the excess mortality related to the RA disease itself, the somewhat different ACS phenotype and the inferior short-term outcomes of ACS in RA raise the hypothesis that ACS in RA is, at least in part, the result of other pathogenic processes than in the general population.

With regard to CV characteristics in RA, studies indicate that patients with RA have an increased carotid intima-media thickness (IMT) (7–11), and that the prevalence and perhaps severity of coronary artery calcification is increased (12–17). Notably, these results pertain to patients with RA in the absence of acute ACS. With respect to coronary angiography, only 1 study, of 203

Supported by the Swedish Research Council, the Heart Lung Foundation, and Region Stockholm/Karolinska Institutet.

[1]Marie Holmqvist, MD, PhD, Ängla Mantel, MD, PhD, Johan Askling, MD, PhD: Solna, Karolinska Institutet, Stockholm, Sweden; [2]Stefan James, MD, PhD: Uppsala University, Uppsala, Sweden; [3]Tomas Jernberg, MD, PhD: Danderyd Hospital, Karolinska Institutet, Stockholm, Sweden.

[†]Dr. Wållberg-Jonsson is deceased.

Drs. Holmqvist and Mantel contributed equally to this work.

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

Address correspondence to Marie Holmqvist, MD, PhD, Karolinska Institutet, Department of Medicine, Solna, Clinical Epidemiology Division, T2, SE-171 76 Stockholm, Sweden. Email: marie.holmqvist@ki.se.

Submitted for publication September 27, 2019; accepted in revised form April 7, 2020.

SIGNIFICANCE & INNOVATIONS

- Our study suggests that overall, among individuals with acute or stable coronary artery disease, the pattern of cardiovascular risk factors and clinically significant coronary stenoses is largely the same in rheumatoid arthritis (RA) as in the non-RA population.
- Despite these similarities, we also noted a higher prevalence of ST-elevation myocardial infarction among patients with RA, a finding that supports the existing evidence of a different acute coronary syndromes (ACS) phenotype in RA, which in turn should be considered when assessing and risk stratifying ACS patients in clinical practice.

patients with symptoms of angina pectoris during exercise, suggested more significant coronary artery involvement in RA than among non-RA controls (18). Studies of the distribution of coronary stenoses in patients with RA at the time of an ACS are lacking, as are studies comparing any characteristics of coronary angiographic patterns in RA versus the general population across ACS and stable ischemic heart disease. Such data would, however, be important for our understanding of whether ACS in RA is largely an effect of an increased force of development of usual ACS in these individuals ("more of the same"), or whether ACS in RA displays RA-specific features.

The aim of this study was therefore to assess the angiographic pattern in patients with RA and in matched subjects from the general population who underwent coronary angiography because of an acute ACS. For contextualization, we assessed the corresponding characteristics in individuals undergoing angiography due to stable coronary artery disease (CAD).

PATIENTS AND METHODS

Study design and setting. We performed a nationwide population-based cohort study of 1 cohort of patients with RA (of any duration), and 1 general-population cohort, both of which had undergone coronary angiography due to ACS or stable CAD. Health care services in Sweden are publicly funded. Patients with RA are treated at specialized internal medicine or rheumatology clinics, patients with acute coronary conditions are treated at coronary intensive care units. All Swedish residents are assigned a personal identification number at birth or immigration, which can be used to link different national health register sources together.

Data sources. The National Patient Register includes main and secondary diagnoses for specialized inpatient care (with full coverage since 1987) and also covers specialized outpatient care since 2001. Diagnoses are coded according to the Swedish version of the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision. Information on dispensed pharmacotherapies was collected from the Prescribed Drug Register (PDR), which stores information on all dispensed drugs from Swedish pharmacies since July 2005. Drugs are coded according to Anatomical Therapeutic Chemical classifications. The clinical Swedish Rheumatology Quality Register (SRQ) is run by the Swedish Rheumatology Association and holds clinical information on patients with RA. The Swedish Coronary Angiography and Angioplasty Register (SCAAR) is part of Swedeheart, the national quality-of-care register for coronary intensive care. SCAAR contains information on all angiographies performed in Sweden and includes information on baseline risk factors, indications, and angiographic findings. The Total Population Register holds demographic information for all Swedish residents.

Study population and exposure definition. Prevalent RA (exposure) was defined as individuals age ≥18 years with at least 2 visits listing RA in the National Patient Register between 2001 and 2015, of which at least 1 had to be from an internal medicine or rheumatology clinic or listed in the SRQ. The positive predictive value for this exposure definition is 91%, based on a retrospective chart-review using current classification criteria as the gold standard (19). Each patient with RA was matched, using incidence density sampling, to up to 5 general-population subjects by age, sex, and residential area, using the Total Population Register.

Identification of coronary angiographies and covariates. We followed all patients with RA and their matched subjects in SCAAR from 2006 through December 31, 2015 to identify all individuals who underwent coronary angiography. For subjects with >1 angiography, only the first was included. Based on the indication registered in SCAAR, we stratified subjects into stable CAD or ACS, which was in turn further stratified into ST-elevation myocardial infarction (STEMI) or non–ST-elevation ACS (NSTACS). To exclude patients with unstable coronary disease from the stable CAD group, we excluded all subjects who had an ACS diagnosis registered in the National Patient Register >30 days prior to the angiography date. We collected information on cardiovascular risk factors (smoking, body mass index, diabetes mellitus, and hypertension) from SCAAR, and information on pharmacotherapies (dispensed prescription >30 days prior to the angiography) serving as proxies for cardiovascular risk factors from the PDR. Figure 1 shows a flow-chart of the identification of the study population.

Outcome definition. Coronary arteries were categorized by the angiographer as significantly affected if visually >50% of the vessel lumen was obstructed or the fractional flow reserve was ≤0.8 (20). Findings were stratified into the number of affected vessels and/or main stem.

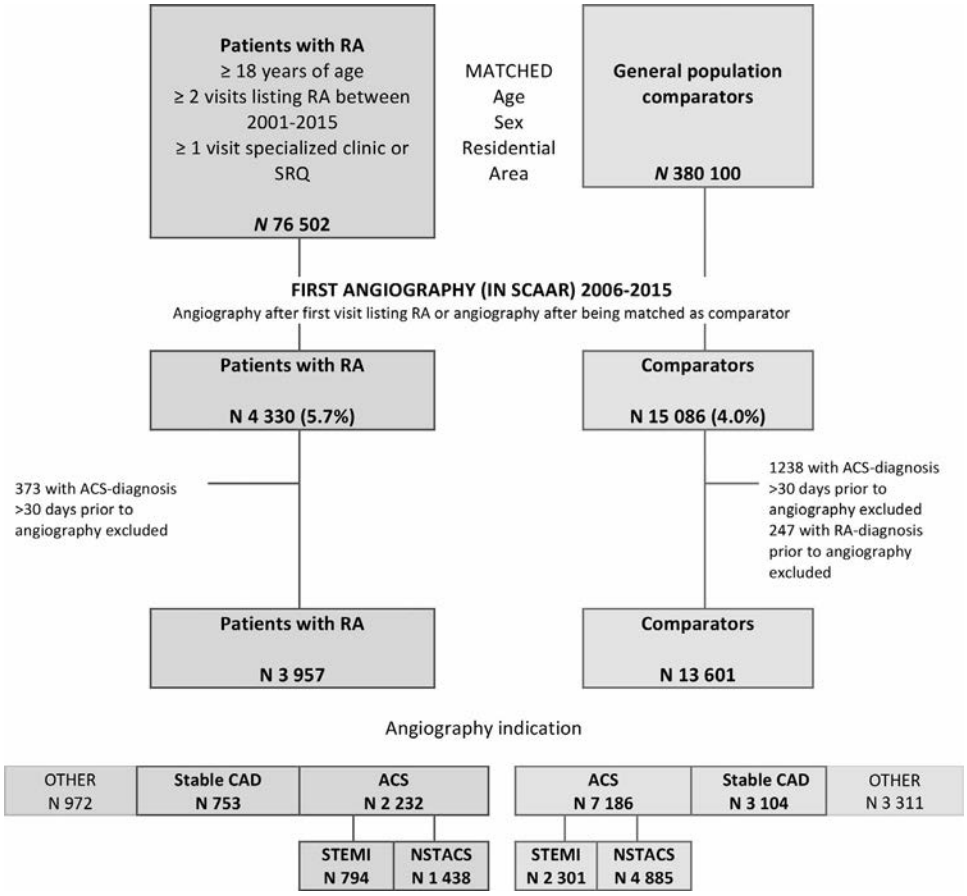


Figure 1. Identification of the study population. ACS = acute coronary syndrome; CAD = coronary artery disease; NSTACS = non–ST-elevation acute coronary syndrome; RA = rheumatoid arthritis; SCAAR = Swedish Coronary Angiography and Angioplasty Register; SRQ = Swedish Rheumatology Quality Register; STEMI = ST-elevation myocardial infarction.

Statistical analysis. We used logistic regression models to calculate age- and sex-adjusted odds ratios (ORs) as a measure of the association between RA and significant stenosis. Each individual was classified based on the number of affected coronary vessels and the location of the stenoses as 1) without any significant stenosis, 2) with one vessel 3) with two vessels, or 4) with three vessels with significant stenosis of the left coronary artery (LAD), left circumflex artery (LCX), or right coronary artery (RCA), and 5) with left main stem and 1 or 2 of LAD, LCX, or RCA. In separate models, we created a composite outcome defined as any of the above vessels affected. Normal findings without any significant stenosis were used as reference in all models. We performed analyses overall and stratified by sex, rheumatoid factor (RF) status, and RA disease duration by the time of angiography (<5 versus ≥5 years).

To investigate whether CV risk factors would potentiate coronary atherosclerosis in RA (compared with non-RA) we performed sensitivity analyses restricted to study subjects with at least 2, or at least 3, CV risk factors, as well as analyses adjusted (complete case analyses) for the same CV risk factors. All analyses were carried out with SAS software, version 9.3.

RESULTS

In total, 3,957 patients with RA (5.2%) and 13,601 patients without RA (3.6%) underwent a first angiography between 2006 and 2015. Of those, 972 patients with RA (24%) and 3,311 patients without RA (32%) did so for other reasons than ACS or CAD (for example as work-up for valvar surgery or heart failure) and were excluded from further analyses.

Distribution of ischemic indications for coronary angiography. Of those who underwent coronary angiography due to ASC or stable CAD, 2,232 patients with RA (75%) and 7,186 without RA (69%) had ACS as an indication ($P < 0.0001$, adjusted for difference). Among those with ACS, 794 patients with RA (36%) and 7,186 without RA (32%) were diagnosed with STEMI ($P = 0.0009$, adjusted for difference) (Figure 1).

Characteristics at the time of coronary angiography. Demographic data, RA disease characteristics, CV risk factors, and preexisting comorbidities for patients with RA and population comparators, stratified by indication for angiography, are shown

Table 1. Demographic information, cardiovascular risk factors, and rheumatoid arthritis (RA) treatment in subjects at the time of coronary angiography between 2006 and 2015, stratified by indication for angiography*

	ACS–STEMI		ACS–NSTACS		Stable CAD	
	RA (n = 794)	Non-RA (n = 2,301)	RA (n = 1,438)	Non-RA (n = 4,885)	RA (n = 753)	Non-RA (n = 3,104)
Year of angiography						
2006–2007	151 (19.0)	422 (18.3)	238 (16.6)	878 (18.0)	139 (18.5)	594 (19.1)
2008–2009	185 (23.3)	445 (19.3)	257 (17.9)	881 (18.0)	144 (19.1)	615 (19.8)
2010–2011	168 (21.2)	479 (20.8)	293 (20.4)	954 (19.5)	166 (22.1)	605 (19.5)
2012–2013	158 (19.9)	487 (21.2)	314 (21.8)	1,076 (22.0)	142 (18.9)	636 (20.5)
2014–2015	132 (16.6)	468 (20.3)	336 (23.4)	1,096 (22.4)	162 (21.5)	654 (21.1)
Women	451 (56.8)	1,158 (50.3)	785 (54.6)	2,577 (52.8)	446 (59.2)	1,829 (58.9)
Age overall, mean ± SD years	71.5 ± 9.9	72.3 ± 10.3	70.4 ± 9.0	71.6 ± 9.1	66.8 ± 9.3	68.2 ± 8.8
Age for women, mean ± SD years	73.1 ± 9.8	73.9 ± 10.4	70.9 ± 9.1	72.2 ± 9.2	66.4 ± 9.5	67.7 ± 8.9
Age for men, mean ± SD years	69.4 ± 9.6	70.6 ± 9.8	69.8 ± 8.8	70.9 ± 8.9	67.3 ± 8.8	69.0 ± 8.6
Disease duration <5 years	245 (30.9)	NA	427 (29.7)	NA	251 (33.3)	NA
RF positivity	701 (88.3)	NA	1,222 (85.0)	NA	601 (79.8)	NA
RA treatment†						
Glucocorticoid	377 (47.5)	NA	685 (47.6)	NA	332 (44.1)	NA
DMARD, any	476 (60.0)	NA	848 (59.0)	NA	453 (60.2)	NA
Biologic drug	96 (12.1)	NA	166 (11.5)	NA	93 (12.4)	NA
NSAID	326 (41.1)	NA	547 (38.0)	NA	278 (36.9)	NA
Cardiovascular risk factors						
Smoking status						
Smoker	105 (13.2)	332 (14.4)	183 (12.7)	502 (10.3)	40 (5.3)	162 (5.2)
Former smoker	141 (17.7)	376 (16.3)	364 (25.3)	1,118 (22.9)	212 (28.2)	784 (25.3)
Never smoker	171 (21.5)	586 (25.5)	361 (25.1)	1,411 (28.9)	209 (27.8)	934 (30.1)
Missing	377 (47.5)	1,007 (43.8)	530 (36.9)	1,854 (37.9)	292 (38.8)	1,224 (39.4)
Preexisting comorbidities‡						
Diabetes mellitus	89 (11.2)	314 (13.7)	248 (17.3)	948 (19.4)	135 (17.9)	555 (17.9)
Hypertension	372 (46.9)	1,017 (44.2)	805 (56.0)	2,801 (57.4)	469 (62.3)	2,011 (64.8)
Preexisting CV/risk factor drugs§						
Insulin + oral antidiabetics	88 (11.1)	287 (12.5)	245 (17.0)	843 (17.3)	132 (17.5)	493 (15.9)
Anticoagulants	275 (34.7)	710 (30.9)	656 (45.6)	2,167 (44.4)	555 (73.7)	2,375 (76.5)
Nitroglycerine	70 (8.8)	206 (9.0)	282 (19.6)	1,016 (20.8)	458 (60.8)	1,961 (63.2)
Beta blockers	306 (38.5)	690 (30.0)	623 (43.3)	2,047 (41.9)	514 (68.3)	2,052 (66.1)
Diuretics	304 (38.3)	666 (28.9)	582 (40.5)	1,706 (34.9)	325 (43.2)	1,115 (35.9)
ACE inhibitors	303 (38.2)	786 (34.2)	653 (45.4)	2,131 (43.6)	361 (47.9)	1,501 (48.4)
Lipid-lowering agents	169 (21.3)	596 (25.9)	473 (32.9)	1,766 (36.2)	409 (54.3)	1,903 (61.3)

* Values are the number (%) unless indicated otherwise. ACE = angiotensin-converting enzyme; ACS = acute coronary syndrome; CAD = coronary artery disease; CV = cardiovascular; DMARD = disease-modifying antirheumatic drug; NA = not applicable; NSAID = nonsteroidal antiinflammatory drug; NSTACS = non–ST-elevation acute coronary syndrome; RF = rheumatoid factor; STEMI = ST-elevation myocardial infarction.
† Filled prescription in Prescribed Drug Register in a window from 6 months to 30 days prior to angiography.
‡ Based on information in the Swedeheart registry.
§ Filled prescription in Prescribed Drug Register ≥30 days prior to angiography.

in Table 1. Among patients with RA, 88% of patients with STEMI, 85% of patients with NSTACS, and 80% of patients with stable CAD were RF positive. One-third of the patients with RA had a disease duration of <5 years, approximately 60% were treated with a disease-modifying antirheumatic drug (DMARD), and 10% received a biologic DMARD between 6 months to 30 days before the coronary angiography. A slightly higher proportion of patients with RA (versus comparators) were former smokers.

Among patients with STEMI, 57% of patients with RA were women as compared with 50% of the comparators, whereas for NSTACS and stable CAD, the sex distributions were similar. Based on registration in SCAAR, there were no major differences in the prevalence of diabetes mellitus between patients with RA and population comparators. Similarly, the use of insulin and oral antidiabetics, based on fulfilled prescriptions in the PDR, were similar. Among patients with STEMI, 47% of patients with RA versus 44% without RA were diagnosed with hypertension, and 38% of patients with RA versus 34% without RA filled a prescription of angiotensin-converting enzyme inhibitors prior to the angiography. The use of diuretics prior to angiography was also somewhat higher among patients with RA (Table 1).

Presence and distribution of coronary stenoses on angiography. Overall, we did not note any difference in the occurrence and distribution of significant stenoses between patients with and without RA, neither for patients with STEMI nor for patients with NSTACS or CAD. All ORs for any and for

Table 2. Odds ratios (ORs) and 95% confidence intervals (95% CIs) from 2 models comparing findings on angiography in patients with rheumatoid arthritis (RA) and patients without RA who underwent coronary angiography 2006–2015, stratified by indication for angiography; no stenosis is used as the reference group in both models*

	ACS–STEMI (RA/non-RA = 794/2,301)		ACS–NSTACS (RA/non-RA = 1,438/4,885)		Stable CAD (RA/non-RA = 753/3,104)	
	RA/non-RA	OR (95% CI)†	RA/non-RA	OR (95% CI)†	RA/non-RA	OR (95% CI)†
Model 1						
No significant stenosis	43/102	Ref.	235/799	Ref.	404/1,695	Ref.
Any vessel	726/2,132	0.8 (0.6–1.2)	1,171/3,967	1.1 (0.9–1.3)	326/1,433	0.9 (0.8–1.1)
Model 2						
No significant stenosis	43/102	Ref.	235/799	Ref.	404/1,695	Ref.
1 vessel, not left main stem	374/1,025	0.9 (0.6–1.3)	481/1,656	1.0 (0.9–1.2)	127/584	0.9 (0.7–1.1)
2 vessels, not left main stem	192/618	0.8 (0.5–1.2)	300/1,011	1.1 (0.9–1.3)	86/349	1.0 (0.8–1.3)
3 vessels, not left main stem	122/368	0.9 (0.6–1.3)	266/870	1.1 (0.9–1.4)	87/311	1.2 (0.9–1.5)
Left main stem and 1 vessel	7/16	1.2 (0.5–3.2)	11/55	0.7 (0.4–1.4)	1/22	0.2 (0.1–1.4)
Left main stem and 2 vessels	9/34	0.7 (0.3–1.6)	38/109	1.3 (0.9–2.0)	6/44	0.6 (0.3–1.0)

* ACS = acute coronary syndrome; CAD = coronary artery disease; NSTACS = non–ST-elevation acute coronary syndrome; Ref. = reference; STEMI = ST-elevation myocardial infarction.
† Adjusted for age at angiography and sex.

vessel-specific stenosis were between 0.2 and 1.2 (Table 2). Adjusting for smoking, hypertension, and diabetes mellitus did not alter these results (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24214/abstract). Table 3 describes results stratified by RF status. Among patients with ACS, RF-positive RA was associated with a 20% higher odds of at least 1 significant stenosis (OR 1.2 [95% confidence interval (95% CI) 1.0–1.5]), which remained significantly increased after adjusting for the type of ACS (STEMI versus NSATCS; OR 1.1 [95% CI 1.0–1.4]). In patients with RF-negative RA, however, we noted a lower odds of stenosis in any vessel compared to the comparators (OR 0.7 [95% CI 0.5–1.0]), a finding that was not formally statistically significant yet was mirrored by ORs <1 in each category of vessel involvement. Among subjects with stable CAD, we noted no association with RA and no difference between RF-positive and RF-negative subjects. Sensitivity analyses by RA duration, and analyses restricted to individuals with ≥2 or ≥3 concomitant CV risk factors, revealed little effect modification by these factors (Tables 4 and 5). Stratifications by sex did not reveal any heterogeneities (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24214/abstract).

Table 3. Odds ratios (ORs) and 95% confidence intervals (95% CIs) comparing findings on angiography in patients with rheumatoid arthritis (RA) and patients without RA who underwent coronary angiography 2006–2015 with indication of acute coronary syndrome or stable CAD, stratified by RF status; no stenosis is used as the reference group in all models*

	ACS				Stable CAD		
	RA	Non-RA	OR (95% CI)†	OR (95% CI)‡	RA	Non-RA	OR (95% CI)†
RF positive							
No significant stenosis	227 (12.1)	901 (12.8)	Ref.	Ref.	322 (55.3)	1,595 (53.0)	Ref.
Any vessel	1,642 (87.7)	6,099 (86.9)	1.2 (1.0–1.4)	1.1 (0.9–1.3)	257 (44.2)	1,433 (47.3)	1.0 (0.8–1.2)
1 vessel, not main stem	743 (39.7)	2,681 (38.2)	1.2 (0.9–1.4)	1.1 (0.9–1.3)	96 (16.5)	584 (19.3)	0.9 (0.7–1.1)
2 vessels, not main stem	431 (23.0)	1,629 (23.2)	1.2 (0.9–1.4)	1.1 (0.9–1.3)	71 (12.2)	349 (11.5)	1.1 (0.8–1.5)
3 vessels, not main stem	333 (17.8)	1,238 (17.7)	1.2 (1.0–1.5)	1.1 (1.0–1.4)	68 (11.7)	311 (10.3)	1.2 (0.9–1.6)
Main stem and 1 vessel	16 (0.9)	71 (1.0)	1.0 (0.6–1.8)	1.0 (0.6–1.7)	1 (0.2)	22 (0.7)	0.3 (0.1–1.9)
Main stem and 2 vessels	40 (2.1)	143 (2.0)	1.3 (0.9–1.9)	1.2 (0.9–1.8)	5 (0.9)	44 (1.5)	0.6 (0.3–1.6)
RF negative							
No significant stenosis	51 (16.7)	901 (12.8)	Ref.	Ref.	82 (54.3)	1,595 (53.0)	Ref.
Any vessel	255 (83.3)	6,099 (86.9)	0.7 (0.5–1.0)	0.7 (0.5–1.0)	69 (45.7)	1,433 (47.3)	1.1 (0.7–1.5)
1 vessel, not main stem	112 (36.6)	2,681 (38.2)	0.7 (0.5–1.0)	0.7 (0.5–1.0)	31 (20.5)	584 (19.3)	1.0 (0.7–1.6)
2 vessels, not main stem	61 (19.9)	1,629 (23.2)	0.6 (0.4–0.9)	0.6 (0.4–1.0)	15 (9.9)	349 (11.5)	0.8 (0.5–1.5)
3 vessels, not main stem	55 (18.0)	1,238 (17.7)	0.8 (0.5–1.2)	0.8 (0.5–1.2)	19 (12.6)	311 (10.3)	1.1 (0.6–2.0)
Main stem and 1 vessel	2 (0.7)	71 (1.0)	–	–	0	22 (0.7)	–
Main stem and 2 vessels	7 (2.3)	143 (2.0)	0.9 (0.4–1.9)	0.9 (0.4–1.9)	1 (0.7)	44 (1.5)	0.4 (0.1–3.3)

* Values are the number (%) unless indicated otherwise. ACS = acute coronary syndrome; CAD = coronary artery disease; Ref. = reference; RF = rheumatoid factor.
† Adjusted for age by time of angiography and sex.
‡ Adjusted for age by time of angiography, sex, and type of acute coronary events (ST-elevation myocardial infarction versus non–ST-elevation acute coronary syndrome).

Table 4. Odds ratios (ORs) and 95% confidence intervals (95% CIs) comparing findings on angiography in patients with rheumatoid arthritis (RA) and patients without RA who underwent coronary angiography 2006–2015 with indication of acute coronary syndrome or stable CAD, stratified by RA disease duration at the time of angiography (<5 vs. ≥5 years); no stenosis is used as the reference group in all models*

	ACS			Stable CAD		
	RA	Non-RA	OR (95% CI)†	RA	Non-RA	OR (95% CI)
<5 years						
No significant stenosis	81 (12.6)	901 (12.8)	Ref.	133 (57.1)	1,595 (53.0)	Ref.
Any vessel	562 (87.3)	6,099 (86.9)	1.0 (0.8–1.3)	100 (42.9)	1,433 (47.3)	0.8 (0.6–1.1)
1 vessel, not main stem	251 (39.0)	2,681 (38.2)	1.0 (0.8–1.4)	38 (16.3)	584 (19.3)	0.8 (0.6–1.2)
2 vessels, not main stem	146 (22.7)	1,629 (23.2)	1.0 (0.8–1.4)	22 (9.4)	349 (11.5)	0.8 (0.5–1.3)
3 vessels, not main stem	123 (19.1)	1,238 (17.7)	1.2 (0.9–1.6)	34 (14.6)	311 (10.3)	1.4 (1.0–2.1)
Main stem and 1 vessel	5 (0.8)	71 (1.0)	0.8 (0.3–2.1)	0	22 (0.7)	–
Main stem and 2 vessels	14 (2.2)	143 (2.0)	1.2 (0.7–2.2)	0	44 (1.5)	–
>5 years						
No significant stenosis	197 (12.8)	901 (12.8)	Ref.	271 (54.2)	1,595 (53.0)	Ref.
Any vessel	1,335 (87.0)	6,099 (86.9)	1.1 (0.9–1.3)	226 (45.2)	1,433 (47.3)	1.0 (0.8–1.3)
1 vessel, not main stem	604 (39.4)	2,681 (38.2)	1.0 (0.9–1.2)	89 (17.8)	584 (19.3)	0.9 (0.7–1.2)
2 vessels, not main stem	346 (22.5)	1,629 (23.2)	1.0 (0.8–1.2)	64 (12.8)	349 (11.5)	1.2 (0.9–1.6)
3 vessels, not main stem	265 (17.3)	1,238 (17.7)	1.0 (0.8–1.2)	53 (10.6)	311 (10.3)	1.1 (0.8–1.5)
Main stem and 1 vessel	13 (0.9)	71 (1.0)	0.9 (0.5–1.6)	1 (0.2)	22 (0.7)	0.3 (0.1–2.2)
Main stem and 2 vessels	33 (2.2)	143 (2.0)	1.1 (0.8–1.5)	6 (1.2)	44 (1.5)	0.9 (0.4–2.1)

* Values are the number (%) unless indicated otherwise. ACS = acute coronary syndrome; CAD = coronary artery disease; Ref. = reference.
† Adjusted for age by time of angiography and sex.

DISCUSSION

In this nation-wide and population-based study of 2,985 patients with RA undergoing coronary angiography due to acute (ACS) or stable (CAD) coronary heart disease, we made 4 important observations: 1) Among patients with RA, the proportion with ACS was higher than in the patients without RA. 2) Among patients with ACS, the proportion with STEMI was higher than in the patients without RA. These findings add to the notion (5) that the ACS phenotype in RA is different from that of the patients without RA. 3) As expected, whereas the distribution of cardiovascular risk factors and smoking differs between the different indications for angiography, within each indication there were relatively modest differences between patients with

Table 5. Odds ratios (ORs) and 95% confidence intervals (95% CIs) comparing findings on angiography in patients with rheumatoid arthritis (RA) and patients without RA who underwent coronary angiography 2006–2015 with indication of acute coronary syndrome or stable CAD, stratified by the number of preexisting cardiovascular risk factors; no stenosis is used as the reference group in all models*

	ACS			Stable CAD		
	RA	Non-RA	OR (95% CI)†	RA	Non-RA	OR (95% CI)
≥2 CV risk factors, no.†	874	2,955	–	409	1,761	–
No significant stenosis	97 (11.1)	334 (11.3)	Ref.	192 (46.9)	808 (45.9)	Ref.
Any vessel	775 (88.7)	2,616 (88.5)	1.0 (0.8–1.3)	215 (52.6)	951 (54.0)	1.0 (0.8–1.2)
1 vessel, not main stem	310 (35.5)	1,016 (34.4)	1.0 (0.8–1.4)	75 (18.3)	342 (19.4)	0.9 (0.7–1.3)
2 vessels, not main stem	194 (22.2)	714 (24.2)	1.0 (0.7–1.3)	57 (13.9)	242 (13.7)	1.0 (0.7–1.4)
3 vessels, not main stem	190 (21.7)	610 (20.6)	1.1 (0.8–1.5)	64 (15.7)	236 (13.4)	1.2 (0.8–1.6)
Main stem and 1 vessel	11 (1.3)	34 (1.2)	1.1 (0.6–2.4)	1 (0.2)	14 (0.8)	0.3 (0.1–2.4)
Main stem and 2 vessels	27 (3.1)	72 (2.5)	1.4 (0.8–2.3)	5 (1.2)	31 (1.7)	0.7 (0.3–1.9)
≥3 CV risk factors, no.†	305	1,155	–	163	737	–
No significant stenosis	36 (11.8)	112 (9.8)	Ref.	71 (43.6)	302 (41.0)	Ref.
Any vessel	268 (87.9)	1,042 (90.2)	0.8 (0.6–1.2)	91 (55.8)	435 (59.0)	0.9 (0.6–1.2)
1 vessel, not main stem	101 (33.1)	368 (31.9)	0.9 (0.6–1.4)	28 (17.2)	150 (20.4)	0.8 (0.5–1.3)
2 vessels, not main stem	76 (24.9)	278 (24.1)	0.8 (0.5–1.2)	23 (14.1)	98 (13.3)	1.0 (0.6–1.7)
3 vessels, not main stem	61 (20.0)	270 (23.4)	0.9 (0.2–3.5)	35 (21.5)	123 (16.7)	1.2 (0.7–1.9)
Main stem and 1 vessel	3 (1.0)	11 (1.0)	1.1 (0.5–2.7)	1 (0.6)	4 (0.5)	1.0 (0.1–9.1)
Main stem and 2 vessels	10 (3.3)	30 (2.6)	0.8 (0.6–1.2)	2 (1.2)	17 (2.3)	0.5 (0.1–2.2)

* Values are the number (%) unless indicated otherwise. ACS = acute coronary syndrome; CAD = coronary artery disease; CV = cardiovascular; Ref. = reference.
† Combination of at least 2 or 3 cardiovascular risk factors (hypertension, smoking, diabetes mellitus, and hyperlipidemia). Analyses were adjusted for age at angiography and sex.

and without RA. This finding suggests that much of the differences observed between RA populations at large versus the patients without RA (e.g., a higher prevalence of smoking) are evened out in populations with manifest CV events. 4) Within each indication for angiography, the distribution of significant stenosis was largely similar for individuals with versus without RA, with 1 exception: patients with seronegative RA and ACS (versus non-RA) had a lower prevalence of significant stenosis. Within each indication for angiography, we found that RA duration, sex, or the load of CV risk factors did not markedly modify the occurrence and distribution of stenoses in RA versus those in population comparators.

Whereas most previous studies of CV characteristics in RA have been performed in patients without ongoing acute events or stable ischemic heart disease (21), our study extends these findings to patients with acute ACS, or stable CAD, for which very little is known. Our results may, however, be compared to those of Warrington et al, in which the angiographic pattern among 75 patients with RA complicated by angina pectoris between 1985 and 1998 were compared to 128 controls with angina pectoris (18). The results indicated that patients with RA had more extensive coronary disease than controls. In an autopsy study by Aubry et al (22), individuals who had a history of cardiovascular disease were compared (25 patients with RA, 51 controls). In that study, patients with RA were less likely to have multiple vessel disease and had less severe coronary atherosclerosis, both in terms of extent (the number of vessels with stenosis) and degree (obstruction of cross-sectional area by percentage). As suggested by the authors, possibly patients with fatal cardiovascular disease have a less complicated coronary disease profile but die from sudden deaths due to rupture of more vulnerable plaques. Interestingly, in the same study, the authors reported increased inflammation in the coronary arterial walls and an increased frequency of vulnerable plaques (22) among patients with RA compared to patients without RA, supporting the hypothesis that vulnerable plaques may be of greater importance in clinical outcomes than the actual amount of coronary atherosclerosis.

Patients with RA have previously been shown to have an increased frequency of vulnerable plaques (22). The underlying cause of plaque rupture is related to the accumulation of oxidized low-density lipoprotein–containing inflammatory cells and activation of inflammatory mechanisms, leading to thinning of the cap. Patients with RA also have increased levels of activated systemic inflammatory mediators, of which several were associated with the presence of plaques (23) or with carotid IMT (24). In our study, we noted no increased occurrence of significant coronary stenoses in seropositive RA complicated by CAD, offering some support to the hypothesis that patients with RA may be more prone to develop thrombi, plaque activation, and rupture. By contrast, we noted a tendency toward a decreased occurrence of significant stenoses in patients with seronegative RA and STEMI or NSTACS, suggesting the possible role of other factors behind some of these ACS.

Our study has a number of limitations. We were able to study the degree of clinically significant stenosis, as reported by the angiographist, but not the degree of stenoses of any magnitude, nor the full influence of potential risk factors for atherosclerosis. Several traditional ACS risk factors are correlated with an increased IMT in patients with RA (25), and we noted generally modest differences in the presence of CV risk factors between those with and without RA, but we did not fully accommodate these factors in our statistical modeling. However, adjusting for smoking, hypertension, and diabetes mellitus did not alter the interpretation of our findings. By definition, our findings are based on, and generalizable to, patients who undergo angiography, not all patients with CAD or ACS.

Our study also has a number of strengths, including its size, the nation-wide and population-based setting, and the occurrence of prospectively recorded data that were registered independently of RA status. Patients were recruited from rheumatology clinics all over the country, including private clinics and smaller hospital clinics. This method of recruitment prevented us from preferentially selecting patients with more severe RA; all investigators were unaware of the study hypothesis when the angiography was performed, which further strengthens the validity of our study. Because of the population-based sampling, our results should be generalizable to typical RA populations but might not be directly applicable to populations with different levels of RA disease control, or different use of antirheumatic therapies.

To conclude, our results extend our understanding of coronary characteristics in patients with RA, from characteristics in the absence of acute coronary events to a comparison of characteristics among individuals with acute events or stable CAD. Our results demonstrate a tendency toward relatively more STEMI and ACS in RA than in the patients without RA, but also indicate that some of the differences in CV risk factor distribution between RA and the patients without RA are attenuated in the context of manifest CV events, suggesting common pathways for coronary ischemic events in RA and in the patients without RA (or shared risk factors for different types of coronary ischemic events). Importantly, our study demonstrates little difference between RA and the general population in relation to the presence, number, or location of significant stenoses at a time point when there is a clinical indication for angiography, but our study does indicate a potential difference in the presence of significant stenoses by RF status (less often than expected in RF-negative RA), suggesting different pathways toward ACS in different types of RA inflammation.

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. Holmqvist 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. Holmqvist, Mantel, Wållberg-Jonsson, Askling.

Acquisition of data. Holmqvist, Mantel, Wållberg-Jonsson, James, Jernberg, Askling.
Analysis and interpretation of data. Holmqvist, Mantel, Wållberg-Jonsson, James, Jernberg, Askling.

REFERENCES

1. Baghdadi LR, Woodman RJ, Shanahan EM, Mangoni AA. The impact of traditional cardiovascular risk factors on cardiovascular outcomes in patients with rheumatoid arthritis: a systematic review and meta-analysis. PLoS One 2015;10:e0117952.
2. Radovits BJ, Popa-Diaconu DA, Popa C, Eijsbouts A, Laan RF, van Riel PL, et al. Disease activity as a risk factor for myocardial infarction in rheumatoid arthritis. Ann Rheum Dis 2009;68:1271–6.
3. Mantel Ä, Holmqvist M, Nyberg F, Tornling G, Frisell T, Alfredsson L, et al. Risk factors for the rapid increase in risk of acute coronary events in patients with new-onset rheumatoid arthritis: a nested case–control study. Arthritis Rheumatol 2015;67:2845–54.
4. Holmqvist M, Ljung L, Askling J. Acute coronary syndrome in new-onset rheumatoid arthritis: a population-based nationwide cohort study of time trends in risks and excess risks. Ann Rheum Dis 2017; 76:1642–7.
5. Mantel Ä, Holmqvist M, Jernberg T, Wållberg-Jonsson S, Askling J. Rheumatoid arthritis is associated with a more severe presentation of acute coronary syndrome and worse short-term outcome. Eur Heart J 2015;36:3413–22.
6. Mantel Ä, Holmqvist M, Jernberg T, Wållberg-Jonsson S, Askling J. Long-term outcomes and secondary prevention after acute coronary events in patients with rheumatoid arthritis. Ann Rheum Dis 2017;76:2017–24.
7. Alkaabi JK, Ho M, Levison R, Pullar T, Belch JJ. Rheumatoid arthritis and macrovascular disease. Rheumatology (Oxford) 2003;42:292–7.
8. Daza L, Aguirre M, Jimenez M, Herrera R, Bollain JJ. Common carotid intima-media thickness and von Willebrand factor serum levels in rheumatoid arthritis female patients without cardiovascular risk factors. Clin Rheumatol 2007;26:533–7.
9. Grover S, Sinha RP, Singh U, Tewari S, Aggarwal A, Misra R. Subclinical atherosclerosis in rheumatoid arthritis in India. J Rheumatol 2006;33:244–7.
10. Roman MJ, Moeller E, Davis A, Paget SA, Crow MK, Lockshin MD, et al. Preclinical carotid atherosclerosis in patients with rheumatoid arthritis. Ann Intern Med 2006;144:249–56.
11. Jonsson SW, Backman C, Johnson O, Karp K, Lundström E, Sundqvist KG, et al. Increased prevalence of atherosclerosis in patients with medium term rheumatoid arthritis. J Rheumatol 2001; 28:2597–602.
12. Chung CP, Oeser A, Raggi P, Gebretsadik T, Shintani AK, Sokka T, et al. Increased coronary-artery atherosclerosis in rheumatoid arthritis: relationship to disease duration and cardiovascular risk factors. Arthritis Rheum 2005;52:3045–53.
13. Kao AH, Krishnaswami S, Cunningham A, Edmundowicz D, Morel PA, Kuller LH, et al. Subclinical coronary artery calcification and relationship to disease duration in women with rheumatoid arthritis. J Rheumatol 2008;35:61–9.
14. Wang S, Yiu KH, Mok MY, Ooi GC, Khong PL, Mak KF, et al. Prevalence and extent of calcification over aorta, coronary and carotid arteries in patients with rheumatoid arthritis. J Intern Med 2009;266:445–52.
15. Yiu KH, Wang S, Mok MY, Ooi GC, Khong PL, Lau CP. Role of circulating endothelial progenitor cells in patients with rheumatoid arthritis with coronary calcification. J Rheumatol 2010;37:529–35.
16. Giles JT, Szklo M, Post W, Petri M, Blumenthal RS, Lam G, et al. Coronary arterial calcification in rheumatoid arthritis: comparison with the Multi-Ethnic Study of Atherosclerosis. Arthritis Res Ther 2009; 11:R36.
17. Hromádka M, Seidlerová J, Baxa J, Suchý D, Rajdl D, Šedivý J, et al. Relationship between hsTnI and coronary stenosis in asymptomatic women with rheumatoid arthritis. BMC Cardiovasc Disord 2016;16:184.
18. Warrington KJ, Kent PD, Frye RL, Lymp JF, Kopecky SL, Goronzy JJ, et al. Rheumatoid arthritis is an independent risk factor for multivessel coronary artery disease: a case control study. Arthritis Res Ther 2005;7:R984–91.
19. Waldenlind K, Eriksson JK, Grewin B, Askling J. Validation of the rheumatoid arthritis diagnosis in the Swedish National Patient Register: a cohort study from Stockholm County. BMC Musculoskelet Disord 2014;15:432.
20. Wong JT, Molloi S. Determination of fractional flow reserve (FFR) based on scaling laws: a simulation study. Phys Med Biol 2008; 53:3995–4011.
21. Ambrosino P, Lupoli R, Di Minno A, Tasso M, Peluso R, Di Minno MN. Subclinical atherosclerosis in patients with rheumatoid arthritis: a meta-analysis of literature studies. Thromb Haemost 2015;113:916–30.
22. Aubry MC, Maradit-Kremers H, Reinalda MS, Crowson CS, Edwards WD, Gabriel SE. Differences in atherosclerotic coronary heart disease between subjects with and without rheumatoid arthritis. J Rheumatol 2007;34:937–42.
23. Wallberg-Jonsson S, Trifunovic Cvetkovic J, Sundqvist KG, Lefvert AK, Rantapää-Dahlqvist S. Activation of the immune system and inflammatory activity in relation to markers of atherothrombotic disease and atherosclerosis in rheumatoid arthritis. J Rheumatol 2002;29:875–82.
24. Sodergren A, Karp K, Boman K, Eriksson C, Lundström E, Smedby T, et al. Atherosclerosis in early rheumatoid arthritis: very early endothelial activation and rapid progression of intima media thickness. Arthritis Res Ther 2010;12:R158.
25. Gerli R, Sherer Y, Vaudo G, Schillaci G, Gilburd B, Giordano A, et al. Early atherosclerosis in rheumatoid arthritis: effects of smoking on thickness of the carotid artery intima media. Ann N Y Acad Sci 2005;1051:281–90.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 666–670
DOI 10.1002/acr.24185

BRIEF REPORT

Association of Rural Setting With Poorer Disease Outcomes for Patients With Rheumatic Diseases: Results From a Systematic Review of the Literature

Rosemary J. Hollick and Gary J. Macfarlane

Objective. To assess whether clinical and patient-reported outcomes are poorer for individuals with inflammatory and noninflammatory rheumatic diseases living in rural locations.

Methods. We searched 6 databases for articles that reported on primary peer-reviewed research, published in English between 1990 and 2019, that focused on selected rheumatic diseases (rheumatoid arthritis [RA], psoriatic arthritis, axial spondyloarthritis, or osteoarthritis [OA]) and quantified either patient-reported or clinically measured outcomes by a measure of rurality or remoteness. Selected articles were synthesized narratively.

Results. Eight eligible publications, including 753 rural and 929 urban patients, evaluated outcomes in RA (5 studies) and OA (3 studies). Studies were small, single center, and rarely provided a definition of rurality. Aspects relating to rurality, such as access to services, were not measured. In RA, some studies suggested greater functional disability and disease activity in rural dwellers. In OA, there was some evidence to suggest that rural dwellers presented with more advanced degenerative hip changes, and that illness perceptions and coping differed between rural and urban dwellers. No studies examined work outcomes. Potentially important confounding factors such as socioeconomic status were rarely considered.

Conclusion. There remains considerable uncertainty whether outcomes differ for patients with rheumatic disease in rural settings. There is a need for larger scale studies characterizing participants in relation to place of residence in order to determine whether rurality is an independent predictor of outcome or a surrogate marker for socioeconomic factors.

INTRODUCTION

Equitable and timely access to specialist, multidisciplinary care, and support for those with rheumatic disease is essential to prevent poor outcomes such as joint deformities, functional limitations, and disability (1). Most specialist services are located in urban areas, yet the aging, multimorbid population is increasing faster in rural areas (2). Geographic location has a significant impact on health inequalities, with social exclusion and isolation, access to and awareness of health services, poor housing, low income, travel distance, and availability and accessibility of transport impacting disproportionately on rural communities (2).

Delivering and sustaining quality health care services to rural areas to manage the needs of patients with chronic, complex conditions is challenging (e.g., recruiting and retaining an appropriately skilled workforce, and difficulties realizing economies of scale while adequately serving sparsely populated areas) (2). Many studies have highlighted inadequate access to specialist health care services for those with rheumatic and musculoskeletal disease (RMD) living in rural and remote locations (3). However, whether patients with RMD living in rural settings have poorer outcomes remains unknown, and there has been, as far as we are aware, no review of the evidence. This systematic review aims to assess whether clinical and patient-reported outcomes are poorer for individuals with inflammatory and noninflammatory rheumatic diseases living in rural locations.

MATERIALS AND METHODS

Literature search. Medline, Embase, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), PsycInfo, Web of Science, and Cochrane Library were searched using Medical Subject Headings (MeSH) and keywords spanning the following

Rosemary J. Hollick, BSc (Hons), MBChB, PhD, Gary J. Macfarlane, BSc (Hons), MBChB, PhD, MD (Hons), DSc: University of Aberdeen, Aberdeen, UK.

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

Address correspondence to Rosemary J. Hollick, MBChB, PhD, Epidemiology Group, University of Aberdeen, Room 107, Health Sciences Building, Foresterhill, Aberdeen AB25 2ZN, UK. E-mail: rhollick@abdn.ac.uk.

Submitted for publication October 30, 2019; accepted in revised form March 3, 2020.

SIGNIFICANCE & INNOVATIONS

- Few studies and studies of poor quality have examined treatment outcomes in patients with rheumatic disease living in rural settings.
- Rurality (and related factors) are poorly defined in studies that have been conducted.
- The role of potentially important confounding factors, e.g., social and economic factors, has not been considered.
- Mixed-methods approaches are necessary to understand the complex interplay between rurality and health outcomes.

fields: selected rheumatic diseases (rheumatoid arthritis [RA], psoriatic arthritis [PsA], axial spondyloarthritis [SpA]/ankylosing spondylitis, and osteoarthritis [OA]), rural or urban area of residence, and disease outcomes. The latter included clinical status/disease activity, patient or physician global assessment (including quality of life), and measures of function. The search strategy was initially developed for Medline and adapted for each database.

Study eligibility. Publications were eligible for inclusion if they met the following criteria: 1) reported primary research and appeared in a peer-reviewed journal; 2) focused on adult patients with RA, PsA, axial SpA/ankylosing spondylitis, and/or OA, or these patient groups could be separately identified; 3) compared and quantified at least 1 measure of disease outcome (clinical or patient reported) in patients resident in rural and urban areas; and 4) were published in English between January 1990 and July 2019. We chose 1990 as the earliest date because studies conducted prior to this would have involved considerably different approaches to the clinical management of eligible patients.

Study selection. After duplicate removal, 1 reviewer screened all records by title, abstract, and subsequently full text to determine inclusion. Uncertainties were resolved by consensus. Bibliographies of all included publications were manually searched to obtain additional relevant publications. Relevant data were extracted by 1 reviewer and checked by a second reviewer (RJH and GJM). Due to the heterogeneity of study design, the diseases investigated, and the outcomes measured, a meta-analysis was not conducted. Data were extracted and summarized narratively.

RESULTS

A total of 8 publications were identified as including an eligible study. No additional publications were included after screening the reference lists of eligible publications (Figure 1). Eligible studies included a total of 753 and 929 patients in rural and urban locations, respectively, from 8 countries (2 from Europe, 3 Asia, 2 Africa, and 1 from the Americas). Of the eligible studies, 5 investigated RA, and 3 studies investigated OA (1 hip OA, 2 knee OA); no studies were identified on PsA or axial SpA. With the exception of 1 study, which was longitudinal in design, all others were cross-sectional. Most studies examined patients attending specialist rheumatology or orthopedic services.

In describing studies, focus has principally been on recognized measures of clinical status (e.g., disease activity or degree of degeneration), patient/clinician global assessment of disease (including quality of life), measures of function, and work productivity. Specifically, we also note the definition used (if any) of rural areas, and whether comparisons between populations are adjusted for factors that could potentially confound the relationship, namely demographic and socioeconomic factors.

RA. Puchner et al (4) enrolled rheumatologists, primarily, whose practice involved rural patients across 3 provinces of Austria. The authors provided a questionnaire that was distributed to consecutive patients with RA and then completed at home. Rurality was measured by 3 parameters: 1) size of the settlement in which patients lived; 2) time to travel to the provincial capital; and 3) time to travel to the patients' rheumatologists. Of 124 participants, 103 described that they lived in a settlement of <50,000 persons. There were no differences in patient-reported health status according to any of the above measures of rurality.

Lekpa et al (5) recruited patients with RA (according to American College of Rheumatology [ACR] 1987 criteria [6]) from the rheumatology outpatient department in Dakar, Senegal. The primary purpose of the report was to compare urban and rural patients. Urban patients were defined as those resident in the capital city or in an administrative center ("chef-lieu") of the region, while all other patients were classified as living in rural settings. The study included 180 patients, of which 143 (79%) lived in urban areas. Comparing the 2 groups, the authors reported no differences in the presence or type of hand deformities. In those from rural areas, the median level of swollen joints (6 versus 4) and the Disease Activity Score in 28 joints (7.2 versus 6.4) was higher, although neither difference was statistically significant. A greater proportion of rural patients had extraarticular manifestations (70.3% versus 49%). There was no difference in the presence of rheumatoid factor or anti–citrullinated protein antibody between the 2 groups. There was no account of confounding factors made in comparing the groups, but it was noted that there were marked differences in sex, with men more likely to be from rural areas than women (41% versus 18%).

Zhao et al (7) recruited 607 patients who met the 1987 ACR criteria for RA (6) from a rheumatology outpatient department in Chengdu, China. Residence was categorized as urban (n = 222), suburban (n = 116), or rural (n = 269), although no details were provided about how this classification was made. Clinical and self-reported information was collected and related to the Health Assessment Questionnaire disability index (HAQ DI). Functional disability significantly increased across people living in urban,

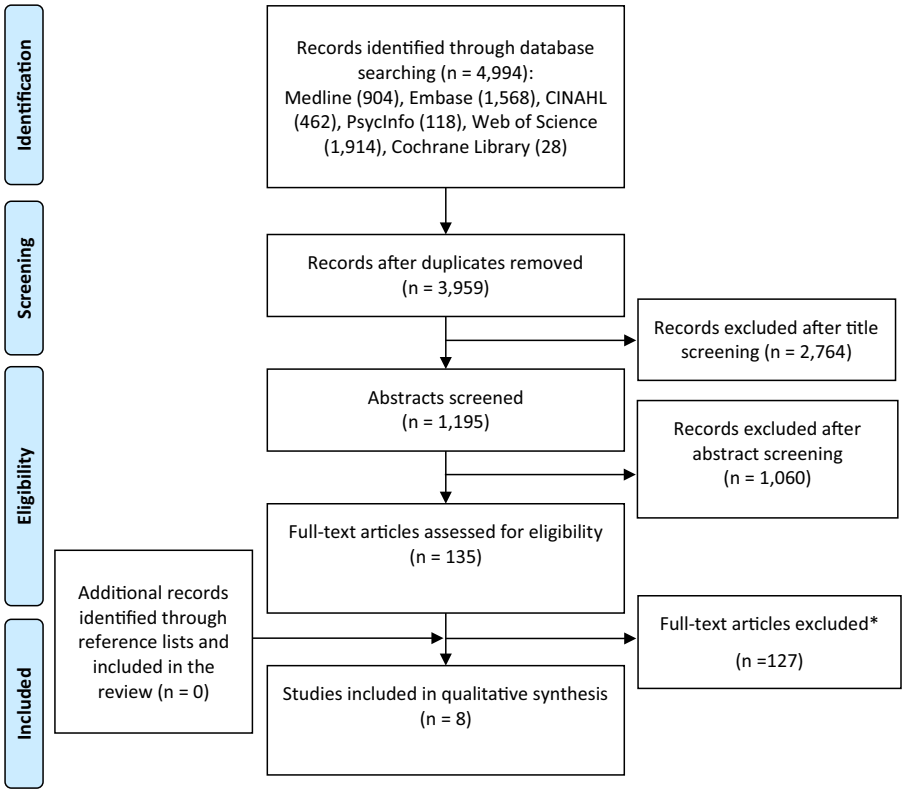


Figure 1. Flow chart identifying eligible studies. * = exclusion of full-text articles due to the following reasons: full text not English (n = 1); conference abstract only (n = 9); no primary peer-reviewed study (n = 10); not investigating disease of interest/did not differentiate between different types of arthritis (n = 29); no comparison of rural and urban patients (n = 44); not investigating disease outcomes (n = 34). CINAHL = Cumulative Index to Nursing and Allied Health Literature.

suburban, and rural settings. However, this study additionally performed a multivariable analysis in which living in a rural setting was an independent predictor of increased functional disability (equivalent to a 1.23-point increase in HAQ DI score). Other independent predictors were lack of available social support, older age, pain, number of times hospitalized, and disease duration (note that this interpretation is from the tables, as the text in this study gives contradictory interpretation of the data). Across the study population, 70% were educated below junior high school, and 40% had a household income monthly per capita <$160. However, education level and household income was not adjusted for in the analysis.

Alarcón et al (8) recruited 189 patients from the rheumatology clinic of a referral center in 1 region of Chile. This center also acted as a referral center for other regions. Rural residence was defined as "living in the scattered agricultural communities," and 61 participants were classified as such. Disability was measured using the Spanish version of the HAQ and dichotomized into moderate/severe versus slight due to sparse data. Rural residence was associated with moderate/severe disability (crude odds ratio [OR] 3.3 [95% confidence interval (95% CI) 1.2–11.6]), but with considerable uncertainty around the level of association. On multivariable analysis, the strength of association was reduced and was not significant (adjusted OR 2.1 [95% CI 0.6–7.1]) after adjusting for socioeconomic status, ethnicity, sex, and age. This study was probably too small to be able to conduct a robust multivariable analysis.

In a study that involved comparing Black Zimbabwean patients and White UK patients with RA (according to the American Rheumatism Association 1958 criteria [9]), Chikanza et al (10) presented data comparing urban (n = 41) and rural (n = 43) Zimbabwean participants attending a tertiary rheumatology clinic. No definition of how rural status was defined was included in the report. Across a variety of clinical, serologic, and radiologic measures of RA, the data show no differences of large magnitude, and the text reports no significant differences.

OA. Roopsawang and Aree-Ue (11) undertook an interview survey in 3 communities in Bangkok, Thailand and 3 communities in its vicinity. Participants were recruited through health centers and community leaders and were considered to have knee OA based on symptoms and signs using ACR criteria (12). There were 116 and 112 participants classified as rural and urban, respectively, but there were no details given on the classification. There were no confounding factors considered in comparisons, but it was noted that those in rural areas were considerably more likely to be of normal weight (35% versus 19%) and less likely to have comorbidities (35% versus 47%). Only 5.2% of rural dwellers

were educated to high school level compared to 34.8% of urban dwellers. This study only reported on illness representation and coping behavior. Rural dwellers were less likely to perceive their symptoms as curable, more likely to use spiritual coping methods, and less likely to use cognitive-focused coping behaviors, e.g., information seeking and self-care. Those in rural areas reported a lower level of emotional impact of symptoms.

Rapała et al (13) reported on 200 patients who were about to undergo total hip arthroplasty for OA. A total of 79 and 121 patients were classified as rural and urban, respectively, but there was no indication of how the classification was made. In terms of clinical status, using a scale proposed by Garlicki and Kreczka, patients from rural settings were much more likely to have the most advanced level of hip degenerative changes on a 3-point scale: "…almost 70% of…rural patients…compared to only 44.5% of urban patients." No account was taken of potential confounding factors, but data presented demonstrated no large differences in mean age, proportion of males, and mean body mass index between those classified as urban and rural.

Çankaya et al (14) followed 70 patients prospectively with unilateral primary knee OA who were undergoing arthroplasty. Of these, 45 were classified as rural, but no information was given on the classification. Outcome was measured 6 months postsurgery by functional status using the Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form. Quality of life was assessed using the Short Form 36. There was no important or statistically significant difference in change in function or quality of life comparing those resident in rural and urban settings. No effect of potential confounding was considered (all analyses considered only individual variables), but the study did demonstrate that benefit (in terms of quality of life) increased with higher levels of education and absence of other comorbidities.

DISCUSSION

We have identified only a small number of studies examining clinical or patient-reported measures of outcome in RA and OA, and none in PsA and axial SpA. All studies are relatively small, rarely provide a definition of rural status (and specifically do not measure aspects related to rurality, such as access to services), and do not consider factors that could confound any differences observed between urban and rural patients. No eligible studies examined work outcomes between urban and rural patients.

The aim of this systematic review was to determine whether clinical and patient-reported outcomes were poorer for individuals with inflammatory and noninflammatory rheumatic diseases living in rural locations. We excluded papers that did not differentiate between specific rheumatic diseases. Inflammatory and noninflammatory rheumatic diseases have different underlying pathophysiology, service delivery, and resource requirements. To help inform the type of approaches required to reduce any inequalities, comparisons therefore need to be precise. Any observed differences in outcomes are only meaningful if differences between and across conditions can be determined (for example, are differences in outcome driven by specific conditions or common across all conditions?). However, we did identify 2 large studies from the US (15) and Australia (16) that, although they did not differentiate between types of arthritis, were otherwise relevant. Kovac et al (15) examined health-related quality of life among 1,191 individuals with self-reported arthritis (mainly RA and OA). After adjusting for socioeconomic status, rural residency was an independent predictor of poorer physical and mental health. Dowsey et al (16) found that rural patients in Australia presented at a younger age for hip and knee replacement, and with less severe radiographic disease. The authors postulate that this may reflect rural dwellers working in physical jobs and seeking referral earlier due to concerns about disease progression, as well as a lack of multidisciplinary support for self-management in rural communities. In contrast, Rapała et al (13) found that rural patients in Poland presented with more severe degenerative joint disease. This suggests potential differences in illness perceptions, health-seeking behavior, and coping mechanisms between urban and rural dwellers with arthritis across different health care contexts and cultures.

Our systematic literature review has several limitations. Data were scarce, and most studies only sampled rural patients attending urban-based specialist centers (this is likely to underestimate any geographic differences). The term "rurality" can encompass many relevant issues in terms of health and health care, including population density and population size (which are related to rurality) and availability of and travel time to health care services (more related to remoteness). Only 1 included study attempted to measure >1 of these aspects; most relied on a single geographic descriptor or did not provide any definition of rurality. Studies were conducted across diverse health care contexts, with significant variation in access and provision of specialist RMD services and payment systems. However, these important contextual factors and their potential influence on access to services for those living in rural areas were not measured, making it difficult to draw comparisons and explore reasons for any disparities between urban and rural dwellers. Most studies were cross-sectional, which precluded examination of changes over time or response to therapy.

Low socioeconomic status has been associated with worse clinical outcomes, decreased functional ability, and reduced quality of life in RA (17). While rural areas in developed countries are often considered to be less deprived, hidden rural deprivation is increasingly recognized, representing a complex interplay between factors associated with income, social circumstances, access to services, and patient choice that is not captured by existing area-based measurements of deprivation (18). While several studies in this review reported lower education and income levels in rural settings, most eligible studies did not adjust for socioeconomic status. Those that did (8) were underpowered to examine multivariable relationships. It is therefore difficult to determine, based on

existing evidence, whether rurality is an independent predictor of poor outcome in RMD or a surrogate marker for socioeconomic status. We did not undertake a formal quality assessment, as we had already identified that almost all studies failed in terms of 2 key issues: the definition of rural settings, and taking account of confounding factors when examining the relationship between residence and outcome.

In conclusion, we have identified key priorities for future research. Studies using population-level data are necessary to capture the burden of disease and health outcomes in RMD between rural and urban areas. Differentiating between conditions is an important point to be considered when designing future research to examine rural–urban differences in outcomes in rheumatic disease. There is also a need to define rurality consistently to allow comparison across studies and to have valid and measurable indicators of rural deprivation to explore the independent effect of rurality on health outcomes. Mixed-methods approaches provide additional opportunities to explore the complex interplay between rurality and health outcomes in RMD.

ACKNOWLEDGMENTS

We acknowledge the work of Svenja Kleiser from the Faculty of Biology, Albert-Ludwigs University, Freiburg, Germany, who conducted the literature search during an internship at the Epidemiology Group, University of Aberdeen, UK.

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. Dr. Hollick 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. Hollick, Macfarlane.
Acquisition of data. Hollick.
Analysis and interpretation of data. Hollick, Macfarlane.

REFERENCES

1. Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:1–25.
2. Public Health England. Health inequalities in ageing in rural and coastal areas: an evidence summary of health inequalities among older people in coastal and rural areas. 2019. URL: https://www.gov.uk/government/publications/health-inequalities-in-ageing-in-rural-and-coastal-areas.
3. Widdifield J, Paterson JM, Bernatsky S, Tu K, Thorne JC, Ivers N, et al. Access to rheumatologists among patients with newly diagnosed rheumatoid arthritis in a Canadian universal public health-care system. BMJ Open 2014;4:e003888.
4. Puchner R, Brezinschek HP, Herold M, Nothnagl T, Studnicka-Benke A, Fritz J, et al. Quality of care of rural rheumatoid arthritis patients in Austria. Wien Klin Wochenschr 2014;126:360–7.
5. Lekpa FK, Ndongo S, Tiendrebeogo J, Ndao AC, Daher A, Pouye A, et al. Rheumatoid arthritis in Senegal: a comparison between patients coming from rural and urban areas, in an urban tertiary health care center in Senegal. Clin Rheumatol 2012;31:1617–20.
6. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
7. Zhao S, Chen Y, Chen H. Sociodemographic factors associated with functional disability in outpatients with rheumatoid arthritis in Southwest China. Clin Rheumatol 2015;34:845–51.
8. Alarcón AM, Munoz S, Kaufman JS, Martinez C, Riedemann P, Kaliski S. Contribution of ethnic group and socioeconomic status to degree of disability in rheumatoid arthritis in Chilean patients. Rheumatol Int 2015;35:685–9.
9. Ropes MW, Bennett GA, Cobb S, Jacox R, Jessar RA. 1958 revision of diagnostic criteria for rheumatoid arthritis. Bull Rheum Dis 1958;9:175–6.
10. Chikanza IC, Stein M, Lutalo S, Gibson T. The clinical, serologic and radiologic features of rheumatoid arthritis in ethnic black Zimbabwean and British Caucasian patients. J Rheumatol 1994;21:2011–5.
11. Roopsawang I, Aree-Ue S. Knee osteoarthritis in adult and older Thais living in rural and urban areas: a comparative study. Pac Rim Int J Nurs R 2015;19:187–201.
12. 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.
13. Rapała K, Truszczynska A, Tarnowski A. Total hip arthroplasty in the treatment of degenerative disorders in rural and urban patients: a retrospective, randomised and controlled study. Ann Agric Environ Med 2015;22:102–5.
14. Çankaya D, Cakir Y, Yoldas B, Toprak A, Tabak Y. The influence of social and demographic features on functional level and quality of life after total knee arthroplasty. Turk J Phys Med Rehab 2016;62:257–63.
15. Kovac SH, Mikuls TR, Mudano A, Saag KG. Health-related quality of life among self-reported arthritis sufferers: effects of race/ethnicity and residence. Qual Life Res 2006;15:451–60.
16. Dowsey MM, Petterwood J, Lisik JP, Gunn J, Choong PF. Prospective analysis of rural-urban differences in demographic patterns and outcomes following total joint replacement. Aust J Rural Health 2014;22:241–8.
17. Verstappen SM. The impact of socio-economic status in rheumatoid arthritis. Rheumatology (Oxford) 2017;56:1051–2.
18. Fecht D, Jones A, Hill T, Lindfield T, Thomson R, Hansell AL, et al. Inequalities in rural communities: adapting national deprivation indices for rural settings. J Public Health (Oxf) 2018;40:419–25.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 671–679
DOI 10.1002/acr.24172

American College of Rheumatology
Empowering Rheumatology Professionals

"The Financial Impact Is Depressing and Anxiety Inducing": A Qualitative Exploration of the Personal Financial Toll of Arthritis

Danielle Berkovic,[1] Darshini Ayton,[1] Andrew M. Briggs,[2] and Ilana N. Ackerman[1]

Objective. The financial experience faced by working-age people with arthritis includes living below the poverty line for many. Financial distress among people with arthritis is known to contribute to poorer health outcomes, including high psychological distress and more severe pain. Despite the demonstrated societal cost of arthritis care and management, the personal costs borne by the individual are not well understood. The aim of this study was to explore the perceived financial impacts of living with arthritis among younger adults (defined as those ages 18–50 years).

Methods. A qualitative descriptive study design was used. Participants with inflammatory arthritis or osteoarthritis were recruited from the community, including urban and rural settings. An interview schedule was developed, informed by existing literature, which was piloted prior to data collection. Deductive and inductive coding techniques were used to identify financial-related themes arising from the data.

Results. Semistructured interviews were conducted with 21 adults (90% female) with a mix of arthritis conditions, including rheumatoid arthritis, psoriatic arthritis, and osteoarthritis. Four themes were identified: direct arthritis-attributable medical costs, indirect arthritis-attributable costs, insurance and pension costs, and broader financial impacts on the family. Nonsubsidized costs were frequently referenced by participants as burdensome and existed even within the publicly funded Australian health care system.

Conclusion. Adults with arthritis experience significant arthritis-attributable financial burden and related distress. Financial concerns should be actively identified and considered within shared clinical decision-making to provide more patient-centered care for these individuals.

INTRODUCTION

Arthritis is increasingly recognized as a disease that affects people of working age (1). In Australia, the most recent National Health Survey data indicate that 24% of people with arthritis are ages 25–54 years, the peak income-earning years for most (2). Given the breadth of biopsychosocial impacts associated with arthritis, including pain and reduced physical function and higher levels of anxiety and depression, individuals with arthritis are likely to experience career disruption, reduced work productivity, and financial burden sequelae (3–6).

The economic impacts of arthritis in working-age populations are profound because many patients transition into early retirement due to the condition (7–9). At a population level, the sequence of arthritis-attributable early retirement and welfare-related costs in working-age persons cost Australia $7.2 billion (Australian) in 2015. By 2030, this cost is projected to increase to $9.4 billion (10). At a personal level, the median weekly income for an individual with arthritis is $333.13 (Australian) (11). In contrast, the poverty line for a single adult living alone is $433.00 (12). Financial distress is known to contribute to poorer health among people with lifelong illness and pain, including high psychological distress and severe physical pain (13–15).

A limited body of research provides preliminary insights into the personal financial burden borne by working-age people with arthritis. Evidence suggests that individuals with rheumatoid

Ms Berkovic's work was supported by Musculoskeletal Australia. Dr. Briggs' work was supported by a fellowship from the Australian National Health and Medical Research Council (grant 1132548). Dr. Ackerman's work was supported by a Victorian Health and Medical Research fellowship from the Victorian Government.

[1]Danielle Berkovic, BHSc (Hons), Darshini Ayton, BBiomedSci (Hons), MPH, PhD, Ilana N. Ackerman, BPhysio (Hons), PhD: Monash University, Melbourne, Victoria, Australia; [2]Andrew M. Briggs, BSc(Phthy) (Hons), PhD, FACP: Curtin University, Perth, Western Australia, Australia.

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

Address correspondence to Ilana N. Ackerman, BPhysio (Hons), PhD, School of Public Health and Preventive Medicine, Monash University, 553 St Kilda Road, Melbourne, Victoria 3004, Australia. Email: ilana.ackerman@monash.edu.

Submitted for publication October 2, 2019; accepted in revised form February 18, 2020.

SIGNIFICANCE & INNOVATIONS

- Adults with arthritis experience financial hardship associated with their disease, and many live below the poverty line despite access to a publicly funded health care system.
- Financial impacts identified by participants included the costs of clinical care and medication, reduced employment wages, and burden on the family financial situation.
- These financial implications were associated with considerable distress and anxiety, highlighting the wide-ranging impacts of arthritis on adults.
- Discussion of arthritis-related financial concerns should form part of shared clinical decision-making to facilitate patient centered-care.

arthritis (RA) who are age <65 years spend significantly more on self-management measures and prescription medication than people with RA age >65 years to improve their functional capacity and assist with activities of daily living (16). In Australia, women retiring early due to arthritis have an average of 83% less savings to fund their retirement compared to women who work to retirement age (17). The financial burden on individuals with arthritis and musculoskeletal conditions has been estimated to be higher than the burden experienced by individuals with other lifelong conditions (18). Individuals living with arthritis report a high number of general practitioner (GP) appointments for prescription medications, higher psychology-related health care costs (the prevalence of major depression is 1.6 times higher in people with arthritis than in their healthy peers), and additional pain management costs (19,20).

Direct health care costs can include GP and specialist visits (for example, rheumatologists) as well as consultations with allied health professionals (for example, physical therapists) (21). Pharmaceuticals, diagnostic tests, dietary supplements and/or natural therapies, and supervised exercise programs further add to the cost burden (21). In addition to direct costs, indirect costs include a reduced number of hours worked, forced early retirement, home modifications, travel to and from health care appointments, and contributions from family members (for example, unpaid caretaker responsibilities) (22).

The personal financial burden borne by working-age people with arthritis is yet to be explored in depth. The current study sought to examine the perceived financial impacts of living with arthritis among younger adults (defined as those ages 18–50 years).

PATIENTS AND METHODS

Design. A qualitative descriptive study was undertaken in 2019 to explore the perceived work and financial impacts of arthritis on adults. A separate article has previously reported findings around work participation restrictions and workplace impacts (6). This article focuses on the financial impacts. Human research ethics approval was granted from the Monash University Human Research Ethics Committee (Project ID 12657) in May 2018. Reporting of the study was undertaken according to the Consolidated criteria for Reporting Qualitative research (COREQ-32) (23).

Overview of Australian health care system. Australia adopted a taxpayer-funded universal health care scheme (known as Medicare) in 1984 (24), comprising the Medicare Benefits Schedule and the Pharmaceutical Benefits Scheme. The Medicare Benefits Schedule is a list of health services subsidized by the Australian government. There are over 57,000 items that provide benefits for a range of services, including specialist consultations, diagnostic tests, and procedures (25). The Pharmaceutical Benefits Scheme subsidizes the costs of over 5,000 medications. Via the scheme, the Australian government contributes the majority of the medication cost, and the consumer pays the remaining fee, which is termed the out-of-pocket cost (24).

Australia also has a parallel private health system, supported by private health insurance policies, that individuals can choose to purchase alongside access to Medicare (24). Private hospital insurance covers the cost for some (or all, depending on the health care practitioner) treatment in a private hospital. Private ancillary insurance covers other health services not always included as part of Medicare, including dental treatment and other allied health services and programs (24). The most recent data indicate that 45.1% of the Australian population was covered by private health insurance in 2018 (26).

Participants. Men and women ages 18–50 years who reported a diagnosis of inflammatory arthritis (IA) or osteoarthritis (OA) by a registered medical practitioner (GP or rheumatologist) and who were living in Australia were eligible to participate. The study was advertised through arthritis consumer organizations, university staff newsletters, and social media. Individuals with a range of arthritis disease types, sexes, employment status, geographic locations (urban, rural), and socioeconomic status were recruited via a purposive sampling frame. Those who expressed an interest in participating were provided with further information by the lead researcher (DB) and asked to complete a brief screening questionnaire to confirm their eligibility. Women who were pregnant were excluded from the study because they might have additional concerns related to pregnancy (27,28). Those who had an unconfirmed arthritis diagnosis who were unable to communicate in English, or who were unable or unwilling to provide consent, were also ineligible to participate.

Data collection. An interview schedule informed by existing literature and a validated framework was developed by DB and DA (1,29). DB has lived experience with an inflammatory arthritis condition and in this context was able to assess the relevance of the interview schedule (30). Because this is one of

Table 1. Interview guide as mapped to arthritis-attributable financial factors

Topic	Open questions	Probing questions
Current financial factors (direct costs)	What do you currently spend your money on to help manage your arthritis?	What experiences have you had paying for medical and specialist appointments? What experiences have you had paying for medications and other types of tablets? Do you pay for different types of insurance (health, life, travel) because of your arthritis? What level of financial distress do these out-of-pocket costs cause?
Current financial factors (indirect costs)	Do you have costs that are not directly attributable to arthritis, but that you find affect you financially?	Are you still able to work, and if so, have you had to take time off work for medical appointments or sick days? Do you have the level of productivity that you would like to at work? Has this changed since your arthritis diagnosis? Has missing work, or having reduced productivity at work, led to any financial concerns?
Future financial factors (direct and indirect costs)	Do you have financial concerns looking into the future?	What concerns do you have moving forward about continuing to produce an income? What concerns do you have about the progression of your arthritis, and the out-of-pocket costs associated with that? Are you worried about the financial burden that your arthritis may place on the people around you?

the first instances of arthritis-attributable costs for the individual being explored using a qualitative approach, interview questions were intentionally broad. The interview guide incorporated both open-ended questions and probing questions in relation to financial factors (Table 1). Data collection was also iterative, and probing questions were used based on the participants' responses. Responses related to new financial themes captured in early interviews were incorporated as additional questions in later interviews. All individual, semistructured interviews were conducted via telephone by the same researcher (DB), who has experience in qualitative data collection. All interviews were audio-recorded to enable verbatim transcription. Researcher reflections were captured in writing during the data collection process and were used to optimize the conduct of subsequent interviews but are not reported here.

Data analysis. A thematic analysis approach was adopted. Thematic analysis is a method used in qualitative research to determine, analyze, and compile themes from participant-oriented data (31). Thematic analysis is useful for contextualizing similarities and differences across a range of participant perspectives and to highlight unanticipated insights and novel data (31). Because this research was exploratory and included a sample with varying arthritis-related experiences, a thematic approach to data analysis was suitable (32). Data analysis commenced alongside data collection to enable themes identified in earlier interviews (interviews 1–5) to be explored in subsequent interviews. Participant recruitment and data collection ceased when data saturation was evident (33).

NVivo software, version 12, was used to support data management and analysis via a process of inductive and deductive coding methods using open, axial, and thematic coding (34). Open codes were generated by looking for initial concepts from participants about their arthritis-attributable financial experiences. Axial coding was conducted to connect common themes identified by participants. For example, each participants' individual financial concerns were analyzed collectively to identify similar patterns. Using deductive coding, themes that correspond to the 3 interview guide topics were identified. Coding and data analysis were conducted by DB. To ensure construct validity, the emergent framework of codes was continuously presented back to a multidisciplinary research team comprised of qualitative researchers and physical therapists specializing in arthritis care (35). Where participant quotes are cited, these are provided verbatim. All monetary amounts are expressed in Australian dollars ($1 Australian = 0.67 US)

RESULTS

Thirty-nine people expressed interest in participating. Five people could not be contacted, and within our purposive sampling approach, 6 others were not recruited, to prevent over-sampling of specific arthritis conditions. Of the remaining 28 people (71.8%) who were screened for eligibility, 25 participants (64.1%) were eligible. Of the 25 eligible participants, 21 (53.8%) were included in the final sample (the remaining 4 participants declined to participate due to other commitments or illness). The 21 interviews ranged from 30 to 95 minutes. Data saturation was reached in the final 4 interviews when it became apparent that no new themes were emerging.

Participant characteristics are shown in Table 2. The majority of participants were female (90.0%) and age >30 years (62.0%). More than one-third had been diagnosed with RA (38.0%), with psoriatic arthritis being the next most common diagnosis (19.0%). Almost one-third of participants lived with their partner and children (29.0%). Nearly half the participants had an undergraduate university degree (43.0%). Nearly three-fourths of participants had private health insurance (71.5%). Only one-third of participants

Table 2. Participant characteristics (n = 21)

Characteristic	No. (%)
Female	19 (90.0)
Age, years	
18–30	8 (38.0)
31–40	6 (28.5)
41–50	7 (33.5)
Highest education	
High school	2 (9.5)
Certificate/diploma	5 (24.0)
Undergraduate university degree	9 (42.5)
Postgraduate university degree	5 (24.0)
Current living status	
With partner/spouse and children	6 (28.5)
With partner/spouse	4 (19.0)
Alone	3 (14.0)
With parents	3 (14.0)
With other adults (nonfamily members)	3 (14.0)
With own children	2 (9.5)
Current employment	
Full-time, paid work	7 (33.5)
Part-time/casual, paid work	7 (33.5)
Student	3 (14.0)
Unable to work because of arthritis	3 (14.0)
Unemployed or looking for work	1 (5.0)
Arthritis diagnosis	
Rheumatoid arthritis	8 (38.0)
Psoriatic arthritis	4 (19.0)
Osteoarthritis	2 (9.5)
Ankylosing spondylitis	2 (9.5)
Seronegative inflammatory arthritis	2 (9.5)
Combination of arthritis types	2 (9.5)
Juvenile idiopathic arthritis	1 (5.0)
Private health insurance	
Yes (own policy)	12 (57.5)
Yes (parents' policy)	3 (14.0)
No	6 (28.5)

were in full-time paid employment (33.5%), one-third were in part-time or casual paid employment (33.5%), and 15% reported that they were unable to work because of their arthritis.

Four major themes were evident from the interviews (Table 3): 1) the financial burden of direct arthritis-attributable health care costs, 2) the unexpected financial burden of indirect costs of living with arthritis, 3) benefits versus the financial burden of paying for insurance, and 4) the broader financial impacts on the family.

Theme 1: the financial burden of direct arthritis-attributable medical costs. Participants reported that the out-of-pocket or nonsubsidized costs associated with arthritis-attributable medical expenses were "bloody depressing" and "anxiety inducing." In contrast to an acute or short-term illness, participants emphasized the sustained financial burden due to the lifelong nature of arthritis: "it's the rest of your life you're paying for this stuff." The greatest expenditure incurred was for specialist rheumatologist consultations, although the reported figures varied between participants. For some participants, rheumatologist appointments incurred no out-of-pocket costs because they accessed specialist consultants through the public hospital system. One participant stated that they had an initial consultation with a rheumatologist whose fees were $500 (Australian), whereas the majority of participants reported paying approximately $200 (Australian) per appointment. Regardless of the charge, many participants perceived specialist consultation to be costly: "seeing your rheumatologist all the time is expensive." For those with psoriatic arthritis, seeing a dermatologist to manage the psoriasis component of the condition was considered an additional financial burden.

In addition to rheumatologists' fees, participants highlighted the significant expense associated with medications and allied health services. Many expressed gratitude for publicly funded Medicare health care, as illustrated by a quote from 1 participant: "if Medicare didn't cover my etanercept it would be a thousand dollars a fortnight, stupid money." Although participants acknowledged that medications were made more affordable under the Pharmaceutical Benefits Scheme, they noted the substantial expense associated with multiple concurrent medications: "when you're on 2 or 3, that's a monthly cost that adds up." Participants described using allied health to help manage arthritis-attributable symptoms, but "when the physio costs $65.00 and I'm looking at probably the next 10 years of things like physio and acupuncture," the nonsubsidized costs become burdensome. One participant described paying for preventative health services, as "I need to proactively improve my health and arthritis from a nondrug related perspective…particularly being anxious in the workplace about my limitations about being able to pick up things."

Theme 2: the unexpected financial burden of indirect arthritis-attributable costs. Participants stated that arthritis-related physical symptoms caused career disruptions and hindered their ability to work full-time. Many specified that they "weren't able to work for many years after diagnosis," and that even years after diagnosis "it still works better for me to work part-time." As a result, a common sentiment was that "it would be nice to have some extra money." For some participants, having less money was compounded by unanticipated costs associated with the invisible nature of arthritis. For example, participants explained that driving to work and social events was easier than taking public transport, as "standing on the train my legs actually get quite sore." Fellow commuters tend to "look for visual symptoms like crutches or walking sticks" and as a result, "no one's going to give up their seat because obviously they assume nothing's wrong." Participants therefore "often just end up driving to events," which creates additional costs where you "have to pay for parking" and need "extra money for fuel because it's easier to drive places than to walk or take the train."

In addition to transport costs, participants described financial constraints to the extent that they were unable to afford nonmedical assistance with arthritis-related physical limitations. For example, some participants' symptoms inhibited them from completing activities of daily living, yet they were unable to afford professional

Table 3. Arthritis-attributable themes, subthemes, and illustrative quotes*

Theme and subtheme	Illustrative quotes
The financial burden of direct arthritis-attributable health care costs	
Medical specialist costs	I see it [the rheumatologist] as a money-grabbing thing, so I go every 6 months. They feel your joints and they go yeah, see you in 3 months' time. Like, I've just sat in your waiting room for 2 hours, you've just charged me $200 for that 2 hours of sitting for like a 3-minute appointment. Seeing your rheumatologist all the time is expensive. (Participant 7, F, age 41–50 years, RA)
	The rheumatologist that I'm seeing is very expensive and the rebate isn't huge. (Participant 11, F, age 41–50 years, RA)
	The dermatologist who I went to for my psoriasis did not recognize the fact that I had arthritis as well…I've probably paid for his speedboat since then. That's probably where he could have said hey, I can't help you. (Participant 2, M, age 41–50 years, PsA)
Allied health costs	I've been referred to strengthen my core through Pilates because I've got quite a lot of wear and tear in my spine. So just this week I'm going to start Pilates with a physio. The cost is quite shocking and I suppose that's the thing that's really frustrating. (Participant 15, F, age 31–40 years, CA: RA, OA)
	I've seen a Bowen therapist before, she's quite good, she is quite expensive though, so it hasn't been really on my top priority list. (Participant 9, F, age 18–30 years, OA)
	I used to see a physio and we'd do hydrotherapy. I don't know why but I just sort of stopped. You know, it was quite expensive. (Participant 3, F, age 18–30 years, RA)
	I see a professor of physiotherapy who specializes in arthritis but he's very expensive. One of the public [hospital] practicing physios, but expensive, not a run-of-the-mill physio. (Participant 10, F, age 18–30 years, JIA)
Medication costs	I recently did a budget and I added up all my medications. And then there's calcium and fish oil and those sorts of things as well. Like I added all this up and it was like, $1,500. I was in the red and it made me realize that there's actually quite a lot of money attached to having this condition. I actually have to budget for this. (Participant 12, F, age 31–40 years, PsA)
	I was fortunate growing up that my parents sort of paid for the medication. But now I realize wow, this stuff, not exactly the cheapest thing, and I'm a student, it's a little bit more expensive. (Participant 1, F, age 18–30 years, CA: RA, SLE)
The unexpected financial burden of indirect costs of living with arthritis	
Financial impacts of home modifications and household assistance	I remember struggling to mow the lawns and things like that and not being in a financial position to be able to pay someone to do it. (Participant 5, F, age 18–30 years, AS)
	I think if I didn't have chronic illness we would probably have a lot more money. We probably would have paid off the house. (Participant 14, F, RA, age 41–50 years)
	We put a big extension on the back of the house for my arthritis, which we borrowed. We owe a lot, it's not good, it's not manageable. (Participant 12, F, age 31–40 years, PsA)
	I couldn't really get up in the morning, so I went out and bought a new bed thinking that that might fix all the problems. I spent a few thousand dollars on buying a bed. I don't think it helped at all. (Participant 21, M, age 31–40 years, AS)
Transport and parking costs	I pay extra money for fuel because it's easier for me to drive places than to walk. (Participant 10, F, 18–30 years, JIA)
	When you're on drugs that lower your immune system and you catch public transport…one year I got sick 6 times, so now I drive. And of course, I have to pay for parking, which is really expensive too, so that's another added cost. (Participant 15, F, age 31–40 years, CA: RA, OA)
	Parking, like when I was in hospital for 7 months, parking cost us a fortune. We spent heaps on the parking, we didn't save money at all with me being in hospital. Those parking costs just come right out of the budget. (Participant 7, F, age 41–50 years, RA)
Benefits versus the financial burden of paying for insurance	
Private health insurance	It is expensive. My mum says "are you planning any holidays?" and I say "no, we've got private health insurance, we can't afford all that." (Participant 12, F, age 31–40 years, PsA)
	We can't afford not to have private health because if I need an operation, I can get it done tomorrow. It just has a limit and once you reach that limit it's pretty hard. My expenses wouldn't be as high as the sun now, but anything is better than nothing. (Participant 13, F, age 18–30 years, RA)
Travel and life insurance	Things like travel insurance, that tends to be a lot more expensive when I need that. So that's definitely something I need to think about more when planning to travel. (Participant 9, F, age 18–30 years, OA)
	I got life insurance before I got the rheumatoid. It came with our credit card or whatever it was. They don't know I've got rheumatoid. It's so expensive. I got it before I had it, and nobody else will insure me. (Participant 7, F, age 41–50 years, RA)
Disability pension and health care card	I am on a disability pension, like I think the full disability pension, they get about $800 a fortnight, but I get $200 a fortnight. (Participant 12, F, 31–40 years, PsA)
	I couldn't get a health care card because I earn $20 more than I should. Ridiculous. I'm very fortunate that my partner promised to pay for my medical expenses. Otherwise I wouldn't be able to afford it. (Participant 10, F, age 18–30 years, JIA)

(Continued)

Table 3. (*Cont'd*)

Theme and subtheme	Illustrative quotes
Broader financial impacts on the family	
Single income household	Obviously I can't work. We are a single-income family. So that single-income family, that does impact everything. Going away, it impacts where you can go, stuff like that. (Participant 7, F, age 41–50 years, RA)
	Being on a single income we couldn't really afford to put the kids in childcare every day, that sort of thing. (Participant 12, F, age 31–40 years, PsA)
Financial strain on parents and children	Even though it is my parents' role, I do still worry about it. Because it is still very expensive and I'd hate to put a financial burden on my parents and my family. So yeah, it is still definitely a concern, even though I'm not actually paying for it. (Participant 3, F, age 18–30 years, RA)
	My son, he's in year 5 of university now, and I think if I was ordinary, I think he probably would have gone and got a job properly by now. He might have been able to have holiday or something. It would have been nice for him to have some extra money. (Participant 12, F, age 31–40 years, PsA)

* AS = ankylosing spondylitis; CA = combination of arthritis types; F = female; JIA = juvenile idiopathic arthritis; M = male; OA = osteoarthritis; PsA = psoriatic arthritis; RA = rheumatoid arthritis; SLE = systemic lupus erythematosus.

assistance. Several participants described "scrounging pennies" to pay for home-based ergonomic devices, from less expensive aids such as a "basket on wheels so if the washing needs to be done I can carry it," to a more expensive "gadget that lifts the bottom shelf of the dishwasher so you don't have to bend over."

Theme 3: benefits versus the financial burden of paying for insurance. Participants reported that private health insurance was one of their largest health-related expenses. Individuals or families often choose to purchase private health insurance in case of injury or flare of symptoms. However, those living with arthritis perceived private health insurance as an essential expenditure, stating "I can't afford to not have private health." Many confirmed that they "took out private health insurance because of arthritis," and that this was "because if I need an operation I can get it done tomorrow and not wait for 12 months when you're in desperate agony."

Over one-fourth of participants did not have private health insurance, as "contemplating the premiums would be a lot higher for someone like me with arthritis and I already have no money." Participants who stated that they were unable to afford private health insurance were frustrated that others are allowed to access both the public and private health care systems simultaneously. Many expressed sentiments such as "we don't really have the money for health insurance but I probably would like to have it because then I can have my neck fixed straight away" and that as a result "it's frustrating that people can double dip and go public or private; financially, it's abuse."

In addition to private health insurance, other insurance costs were perceived to weigh heavily on people with arthritis. Participants were frustrated that travel and life insurance were more expensive due to the presence of a preexisting medical condition. Many were left uninsured and expressed concern at the potential financial burden placed on their families. Participants fretted over their limited funds and their frustration at minimal government compensation in the form of disability pensions and health care cards to provide those living with lifelong conditions supplementary income and reduced medical costs. However, participants stated that "I was on a disability pension for the arthritis" but that the pension was rendered futile when "not a lot of doctors do a special concession rate for people on a pension."

Theme 4: broader financial impacts on family. Alongside concern for their own finances, participants voiced distress about the broader financial impacts of arthritis on their families. Younger participants (ages 18–21 years) explained that they lived at home with "a supportive family that would help me out in any situation," but that "it's still a bit concerning that I'm not paying for my own appointments and my parents shouldn't have to." Those who were slightly older (ages 25–30 years) acknowledged that their parents noticed when they were having a flare, and that they would "try and pitch in with costs where they can, but I don't like it because they should enjoy their retirement without worrying about my financial state."

In contrast to children placing financial pressure on their parents, participants who were parents expressed similar worry about imposing a financial burden on their own children. For example, 1 participant explained, "I don't want my children to think that they can't have careers because they have to look after me if I'm much worse when I'm older."

The financial consequences of living with arthritis extended into broader implications for the whole family. For example, living on a reduced income for an extended time meant that families were unable to take holidays, mortgage repayments had to be defaulted or extended, and children were forced to enter the workforce earlier than they otherwise would have. As 1 participant explained, "we live like grey nomads [but] in Australia, no overseas travel, with a chronic condition attached to it."

DISCUSSION

Arthritis is clearly associated with profound financial impacts and associated financial distress among adults. This study is one of the first to examine these financial concerns from an in-depth

qualitative approach, involving a community-based sample of people with different arthritis conditions. Our findings indicate that a range of financial impacts and concerns, including direct arthritis-attributable medical costs and other impacts that lie outside of direct health care, characterize people's experiences of living with arthritis.

Study participants highlighted the high fees for access to rheumatologists. The financial burden of paying for specialist appointments is topical in Australia: a recent review found that the average nonsubsidized cost for an initial rheumatology consultation (net of the subsidized Medicare rebate) is $120.00 (Australian) (36). The cost of medical intervention is found to be unrelated to improved health outcomes or superior quality of care (37). Those individuals with lower health literacy levels may be vulnerable to excess health care expenditures and financial burden without receiving best-practice care (38,39). A taskforce has been developed and aims to ensure that all Medicare Benefits Schedule items provide real clinical value or high-value care and do not expose patients to unnecessary expense (36).

In addition, participants expressed their surprise at the expense of nonsubsidized allied health care costs incurred through the outpatient public hospital system, despite access to universal health care in this country. Due to changing health needs, increasing health care costs, health inequities, and complex health conditions, patients are shouldering growing out-of-pocket costs (40). However, within the fee-for-service payment model, health professionals are permitted to set their own fees (which are typically above the schedule fee that is reimbursed), which can lead to high nonsubsidized costs for some patients (41).

Evidence suggests that the current out-of-pocket costs for people living with lifelong illness in Australia are strongly associated with poverty (42). Similar trends are documented in Nordic countries, which also have combination public and private health care systems (43,44). In Australia, growing out-of-pocket costs are partially attributed to increased uptake of private health insurance due to lengthy waiting periods for a rheumatology, pain medicine, or surgery consultation through the public system (45). Participants also highlighted their fiscal concerns extending beyond direct health care costs, including reduced capacity to pay for their mortgage, childcare, and the impacts on travel and life insurance.

There are emerging data on the effects of lifelong illness on financial domains beyond medical expenses. People living with coronary artery disease have outlined challenges relating to driving costs where public transportation or walking are unfeasible (46). People affected by types 1 and 2 diabetes mellitus have explained that only by limiting expenditures on nonmedical-related items were they able to afford medication (47). However, to the best of our knowledge, this study is one of the first dedicated to examining the perceived financial burden of living with arthritis.

Unsurprisingly, adults with arthritis face much broader personal economic challenges beyond their direct medical costs. This population has been documented to have shorter work careers, to be less confident to pursue career progression opportunities, and to earn significantly less throughout their income-earning years than their healthy peers (6,48). Lower work participation rates and financial sequelae present as concurrent challenges to navigate for people with arthritis. Through reporting these fiscal challenges, our findings provide a starting point for understanding the concerns of younger populations with arthritis beyond the health impacts. In particular, education and support from arthritis consumer organizations or other advocacy groups may be provided to, or accessed by, clinicians treating people of working age with arthritis. Clinicians need to be cognizant that their patients may be experiencing financial distress, and that identifying these concerns as part of routine clinical care can help inform shared decision-making, particularly as it relates to accessing interventions or services that are high-value, and identify available services that may be feasible (for example, referring a patient to a community physical therapy program versus a private practice).

Because our study was exploratory in nature, examining broad arthritis-attributable financial experiences was important, and we were able to recruit a heterogeneous participant sample to achieve this goal. Our recruitment strategy spanning arthritis consumer organizations, university networks, and clinical settings generated a sample that was diverse across age and disease characteristics. In-depth semistructured interviews were used to elicit detailed data from participants. However, we did not directly ask about nonmedical-related costs (for example, home modifications or childcare), although these were reported by some participants during the interviews. In this context, we may have underrepresented this theme in the analysis.

Qualitative research is representative of participants' experiences, but our research cannot be generalized to all people's arthritis-attributable finances. Two-thirds of participants were university-educated, which may indicate higher income levels among our sample compared to the broader population with arthritis. We also recognize that a relatively high proportion of our sample had private health insurance, compared with the general population, but this difference does not necessarily reflect the socioeconomic status of our sample, given ongoing government initiatives designed to lower the cost of private health insurance and improve uptake. We also acknowledge the potential for participant bias, where those with a higher financial burden may have been more likely to volunteer to be a part of this research. There was an oversampling of females (reflecting the demographics of arthritis); a potentially important area of future research, therefore, will be to explore these issues among males with IA and OA.

This study highlights the spectrum of ongoing direct and indirect costs borne by adults living with arthritis conditions. The in-depth interviews provided novel insight into the range of financial concerns experienced by younger patient groups and the personal distress associated with these concerns. These findings can be used to raise awareness of key fiscal issues relevant to adults

with arthritis, and to educate clinicians about the wide-ranging impacts of arthritis beyond physical symptoms.

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. Ackerman 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. Berkovic, Ayton, Briggs, Ackerman.
Acquisition of data. Berkovic.
Analysis and interpretation of data. Berkovic, Ayton, Briggs, Ackerman.

REFERENCES

1. Ackerman IN, Kemp JL, Crossley KM, Culvenor AG, Hinman RS. Hip and knee osteoarthritis affects younger people, too. J Orthop Sports Phys Ther 2017;47:67–79.
2. Australian Bureau of Statistics. National health survey: first results, 2017–2018, Australia. Canberra: Australian Bureau of Statistics; 2018.
3. Poh LW, He HG, Chan WC, Lee CS, Lahiri M, Mak A, et al. Experiences of patients with rheumatoid arthritis: a qualitative study. Clin Nurs Res 2017;26:373–93.
4. Jetha A, Badley E, Beaton D, Fortin PR, Shiff NJ, Gignac MA. Unpacking early work experiences of young adults with rheumatic disease: an examination of absenteeism, job disruptions, and productivity loss. Arthritis Care Res (Hoboken) 2015;67:1246–54.
5. Sharif B, Garner R, Sanmartin C, Flanagan WM, Hennessy D, Marshall DA. Risk of work loss due to illness or disability in patients with osteoarthritis: a population-based cohort study. Rheumatology (Oxford) 2016;55:861–8.
6. Berkovic D, Ayton D, Briggs AM, Ackerman IN. "I Would be More of a Liability than an Asset": navigating the workplace as a younger person with arthritis. J Occup Rehabil 2020;30:125–34.
7. Young A, Dixey J, Kulinskaya E, Cox N, Davies P, Devlin J, et al. Which patients stop working because of rheumatoid arthritis? Results of five years' follow up in 732 patients from the Early RA Study (ERAS). Ann Rheum Dis 2002;61:335–40.
8. Laires PA, Canhao H, Rodrigues AM, Eusebio M, Gouveia M, Branco JC. The impact of osteoarthritis on early exit from work: results from a population-based study. BMC Public Health 2018;18:472.
9. Westhoff G, Buttgereit F, Gromnica-Ihle E, Zink A. Morning stiffness and its influence on early retirement in patients with recent onset rheumatoid arthritis. Rheumatology (Oxford) 2008;47:980–4.
10. Australian Commission on Safety and Quality in Health Care. Patient-centred care: improving quality and safety through partnerships with patients and consumers. Sydney: Australian Commission on Safety and Quality in Health Care; 2011.
11. Schofield D, Rupendra S, Cunich C. Counting the cost part 2. Economic costs: the current and future burden of arthritis. University of Sydney: Arthritis Australia; 2016.
12. Davidson P, Saunders P, Bradbury B, Wong M. Poverty in Australia, 2018. Sydney: ACOSS/UNSW Poverty and Inequality Partnership Report; 2018.
13. Jeon YH, Essue B, Jan S, Wells R, Whitworth JA. Economic hardship associated with managing chronic illness: a qualitative inquiry. BMC Health Serv Res 2009;9:182.
14. Sturgeon JA, Arewasikporn A, Okun MA, Davis MC, Ong AD, Zautra AJ. The psychosocial context of financial stress: implications for inflammation and psychological health. Psychosom Med 2016;78:134–43.
15. Skinner M, Zautra A, Reich J. Financial stress predictors and the emotional and physical health of chronic pain patients. Cogn Ther Res 2004;28:695–713.
16. Lapsley H, March L, Tribe K, Cross M, Courtenay B, Brooks P. Living with rheumatoid arthritis: expenditures, health status, and social impact on patients. Ann Rheum Dis 2002;61:818–21.
17. Schofield D, Shrestha R, Percival R, Passey M, Callandar E, Kelly S. The personal and national costs of lost labour force participation due to arthritis: an economic study. BMC Public Health 2013;13:188.
18. Maetzel A, Li L, Pencharz J, Tomlinson G, Bombardier C. The economic burden associated with osteoarthritis, rheumatoid arthritis, and hypertension: a comparative study. Ann Rheum Dis 2004;63:395–401.
19. Deloitte Access Economics. The cost of pain in Australia; 2019. URL: https://www2.deloitte.com/content/dam/Deloitte/au/Documents/Economics/deloitte-au-economics-cost-pain-australia-040419.pdf.
20. Australian Bureau of Statistics. National health survey: first results, Australia, 2014–2015. Canberra: Australian Bureau of Statistics; 2015.
21. Australian Institute of Health and Welfare. A picture of rheumatoid arthritis in Australia. In: National centre for monitoring arthritis and musculoskeletal conditions. Canberra: Australian Institute of Health and Welfare; 2009.
22. Ackerman I, Bohensky M, Pratt C, Gorelik A, Liew D. Counting the cost part 1. Healthcare costs: the current and future burden of arthritis. University of Melbourne: Arthritis Australia; 2016.
23. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19:349–57.
24. Australian Government Department of Health. The Australian health system. 2019. URL: https://beta.health.gov.au/about-us/the-australian-health-system.
25. Australian Government Department of Health. Medicare benefits schedule (MBS) Review. 2019. URL: https://www1.health.gov.au/internet/main/publishing.nsf/Content/MBSReviewTaskforce.
26. Calder R, Dunkin R, Rochford C, Nichols T. Australian health services: too complex to navigate. A review of the national reviews of Australia's health service arrangements. 2019. URL: https://apo.org.au/node/223011.
27. Briggs AM, Jordan JE, Ackerman IN, Van Doornum S. Establishing cross-discipline consensus on contraception, pregnancy and breast feeding-related educational messages and clinical practices to support women with rheumatoid arthritis: an Australian Delphi study. BMJ Open 2016;6:e012139.
28. Ackerman IN, Jordan JE, Van Doornum S, Ricardo M, Briggs AM. Understanding the information needs of women with rheumatoid arthritis concerning pregnancy, post-natal care and early parenting: a mixed-methods study. BMC Musculoskelet Disord 2015;16:194.
29. World Health Organization. International classification of functioning, disability and health (ICF). 2018. URL: https://www.who.int/classifications/icf/icfbeginnersguide.pdf.
30. Berkovic D, Ayton D, Briggs AM, Ackerman IN. The view from the inside: positionality and insider research. Int J Qual Methods 2020;19. URL: https://journals.sagepub.com/doi/full/10.1177/1609406919900828.
31. Nowell S, Norris J, White D, Moules J. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16:1–13.
32. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013;15:398–405.
33. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52:1893–907.

34. Liamputtong P. The science of words and the science of numbers: research methods as foundations for ecidence-based practice in health. In: Liamputtong P, editor. Research methods in health: foundations of evidence-based practice. Melbourne: Oxford University Press; 2010. p. 3–22.
35. Tobin G, Begley C. Methodological rigour within a qualitative framework. J Adv Nurs 2004;48:388–96.
36. Medicare Benefits Schedule Review Taskforce. Report from the Specialist and Consultation Physician Consulation Clinical Committee. 2018. URL: https://www.health.gov.au/sites/default/files/documents/2020/12/medicare-benefits-schedule-review-taskforce-final-report-an-mbs-for-the-21st-century-recommendations-learnings-and-ideas-for-the-future---medicare-benefits-schedule-review-taskforce-final-report-to-the-minister-for-health_0.pdf.
37. Australian Commission on Safety and Quality in Health Care. Vital signs 2016: the state of safety and quality in Australian health care. Sydney (AU); 2016. URL: https://www.safetyandquality.gov.au/sites/default/files/migrated/Vital-Signs-2016-PDF.pdf.
38. Hill CL, Appleton SL, Black J, Hoon E, Rudd RE, Adams RJ, et al. Role of health literacy in self-reported musculoskeletal disorders. Arthritis 2015;2015:607472.
39. Adams RJ, Appleton SL, Hill CL, Dodd M, Findlay C, Wilson DH. Risks associated with low functional health literacy in an Australian population. Med J Aust 2009;191:530–4.
40. Dixit SK, Sambasivan M. A review of the Australian healthcare system: a policy perspective. SAGE Open Med 2018;6:20503121187 69211.
41. Australian Government Department of Health. Medicare benefits schedule: item 10960. 2019. URL: http://www9.health.gov.au/mbs/fullDisplay.cfm?type=item&qt=ItemID&q=10960.
42. Callander EJ, Schofield DJ, Shrestha RN. Chronic health conditions and poverty: a cross-sectional study using a multidimensional poverty measure. BMJ Open 2013;3:e003397.
43. Falk J, Bruce D, Burstrom B, Thielen K, Whitehead M, Nylen L. Trends in poverty risks among people with and without limiting-longstanding illness by employment status in Sweden, Denmark, and the United Kingdom during the current economic recession: a comparative study. BMC Public Health 2013;13:925.
44. Tynkkynen LK, Alexandersen N, Kaarboe O, Anell A, Lehto J, Vrangbk K. Development of voluntary private health insurance in Nordic countries: an exploratory study on country-specific contextual factors. Health Policy 2018;122:485–92.
45. Hogg MN, Gibson S, Helou A, DeGabriele J, Farrell MJ. Waiting in pain: a systematic investigation into the provision of persistent pain services in Australia. Med J Aust 2012;196:386–90.
46. Dhaliwal KK, King-Shier K, Manns BJ, Hemmelgarn BR, Stone JA, Campbell DJ. Exploring the impact of financial barriers on secondary prevention of heart disease. BMC Cardiovasc Disord 2017;17:61.
47. Campbell DJ, Manns BJ, Hemmelgarn BR, Sanmartin C, Edwards A, King-Shier K. Understanding financial barriers to care in patients with diabetes. Diabetes Educ 2017;43:78–86.
48. Arthritis Australia. Arthritis: Providing better care at less cost. NSW; 2018. URL: https://arthritisaustralia.com.au/wordpress/wp-content/uploads/2018/03/171215-Providing-better-care-at-less-cost_Arthritis-Australia-2018-Pre-budget-submission_final.pdf.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 680–686
DOI 10.1002/acr.24146

Health Care Utilization for Musculoskeletal Issues During the Prediagnosis Period in Psoriatic Arthritis: A Population-Based Study

Lihi Eder,[1] Karen Tu,[2] Cheryl F. Rosen,[3] Raed Alhusayen,[4] Stephanie Y. Cheng,[5] Jacqueline Young,[5] Willemina Campbell,[6] Sasha Bernatsky,[7] Dafna D. Gladman,[8] J. Michael Paterson,[9] and Jessica Widdifield[10]

Objective. Information about the prediagnosis period in psoriatic arthritis (PsA) is limited. The present study was undertaken to compare health care utilization related to musculoskeletal issues during a 5-year period prior to the diagnosis of PsA versus that of subjects with no prior inflammatory arthritis within a primary care setting.

Methods. We conducted a population-based, matched cohort study using electronic medical records and administrative data in Ontario, Canada. Age- and sex-matched cohorts of PsA patients and comparators from the same family physicians were assembled. Comparators were not allowed to have prior spondyloarthritis, ankylosing spondylitis, or rheumatoid arthritis billing code diagnoses. The study outcomes included health care utilization and costs related to nonspecific musculoskeletal issues during a 5-year period prior to the index date.

Results. We studied 462 PsA patients and 2,310 matched comparators. The odds ratio (OR) related to visiting a primary care physician for nonspecific musculoskeletal issues in patients with PsA was 2.14 (95% confidence interval 1.73–2.64) in the year immediately preceding the index date and was similarly elevated up to 5 years prior. The OR related to using other musculoskeletal-related health care services, including musculoskeletal specialists visits, joint injections, joint imaging, and emergency department visits, was higher in PsA as early as 5 years preceding the index date. Total and musculoskeletal-related health care costs prior to the index date were higher for patients with PsA versus comparators.

Conclusion. A prodromal PsA phase characterized by nonspecific musculoskeletal symptoms may exist. Further study is needed to determine if this represents a window for earlier diagnosis of PsA.

INTRODUCTION

Psoriatic arthritis (PsA) is a musculoskeletal disease that affects up to 30% of the patients with psoriasis (1). It runs a chronic, progressive course and can lead to severe joint damage and loss of function in the first few years of the disease (2). Tight control of disease activity significantly improves joint outcomes in newly diagnosed PsA patients (3), while diagnostic delays are associated with more radiographic joint damage and worse physical function (4,5).

Current care delivery in PsA falls short of available standards, as a significant proportion of patients with psoriasis have undi-

Parts of this material are based on data and information compiled and provided by the Ontario Ministry of Health and Long-Term Care and the Canadian Institute for Health Information. The analyses, conclusions, opinions, and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources. No endorsement is intended or should be inferred.

Supported by the Canadian Rheumatology Association (Canadian Initiative for Outcome in Rheumatology Care grant) and the Institute for Clinical Evaluative Sciences, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care. Dr. Eder's work was supported by the Arthritis Society (New Investigator Salary award). Dr. Tu's work was supported by the Department of Family and Community Medicine at the University of Toronto (Research Scholar award).

[1]Lihi Eder, MD, PhD: University of Toronto and Women's College Research Institute, Women's College Hospital, Toronto, Ontario, Canada; [2]Karen Tu, MD, MSc: North York General Hospital, University of Toronto and Toronto Western Hospital Family Health Team–University Health Network, Toronto, Ontario, Canada; [3]Cheryl F. Rosen, MD: University of Toronto and Toronto Western Hospital, Toronto, Ontario, Canada; [4]Raed Alhusayen, MBBS, MSCE: University of Toronto and Sunnybrook Research Institute, Sunnybrook Hospital, Toronto, Ontario, Canada; [5]Stephanie Y. Cheng, MSc, Jacqueline Young, MSA: Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada; [6]Willemina Campbell, LLB: Toronto Western Hospital, Toronto, Ontario, Canada; [7]Sasha Bernatsky, MD, PhD: Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada; [8]Dafna D. Gladman, MD, FRCPC: University of Toronto and Krembil Research Institute, University Health Network, Toronto Western Hospital, Toronto, Ontario, Canada; [9]J. Michael Paterson, MSc: University of Toronto and Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada; [10]Jessica Widdifield, PhD: University of Toronto, Sunnybrook Research Institute, and Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada.

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

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

Submitted for publication September 10, 2019; accepted in revised form January 14, 2020.

SIGNIFICANCE & INNOVATIONS

- The study results support the notion that a prodromal psoriatic arthritis (PsA) phase occurs in a significant proportion of patients.
- Health care utilization due to musculoskeletal issues is increased.
- Patients with PsA develop musculoskeletal symptoms that lead to increased health care utilization and costs, including performance of diagnostic tests, visits to family doctors, musculoskeletal specialist consultations, and emergency department visits.
- This pattern reveals some of the underlying causes of diagnosis delays of PsA.

agnosed PsA (6). A recent meta-analysis reported a prevalence of 15.5% of patients with psoriasis with previously undiagnosed PsA (7). Analysis of data from a national, Danish registry showed that only 21% of PsA patients were diagnosed within 3 months of symptoms onset (8). Delayed referrals to rheumatology by primary care physicians may contribute to such findings. In a population-based study in Ontario, Canada, the time from initial primary care visit for a musculoskeletal complaint to rheumatology referral was 513 days, which was substantially longer than for other inflammatory arthritic conditions, such as rheumatoid arthritis (RA) (9).

PsA is a heterogeneous disease that can present in various clinical manifestations, such as synovitis, enthesitis, dactylitis, and spondylitis. Some of these features can present with only minimal findings on physical examination, and the differentiation from other conditions, such as osteoarthritis, can be challenging. Furthermore, unlike RA and lupus, PsA has no reliable diagnostic biomarkers (10). These factors, along with the lack of awareness of PsA among patients and primary care physicians, and limited access to specialty care contribute to delays in the diagnosis of PsA (11).

Better information regarding the prediagnosis phases of PsA is required in order improve our understanding of how to address these gaps in care. Such data, however, are limited particularly in the primary care setting. In addition, studies that have attempted to evaluate aspects of the prediagnosis period (e.g., duration of symptoms) often have been based on self-reported information, frequently from tertiary rheumatology centers (4,8). Here, we aimed to characterize the burden of musculoskeletal symptoms prior to the diagnosis of PsA by assessing musculoskeletal-related health care utilization prior to the diagnosis of PsA in patients from primary care settings compared to matched comparators without inflammatory arthritis.

SUBJECTS AND METHODS

Setting. We conducted a matched cohort study using population-based, linked primary care electronic medical records from the Electronic Medical Records Primary Care (EMRPC) database and administrative health data from Ontario, Canada. All Ontarians are insured by a publicly funded, universal health insurance program, the Ontario Health Insurance Plan (OHIP) (12), that covers all hospital and physician services and procedures. All health care encounters are recorded in administrative health care databases, which are linked using unique encoded identifiers.

The EMRPC was used to create a matched cohort of PsA and noninflammatory arthritis comparators. At the time of the study, EMRPC included electronic clinical data from >350 primary care physicians across Ontario and >400,000 patients, with a data range covering the period of 1998 to 2016 (13). The characteristics and the distribution of EMRPC physicians and patients are generally representative of the population of Ontario. EMRPC data are enriched by linkage to Ontario health administrative data.

The administrative databases included the OHIP Claims History Database to identify physician billing and diagnosis codes. All physicians in Ontario are reimbursed by submitting claims to OHIP for medical services rendered. A single diagnosis code is provided with each claim, which represents the main reason for the visit. The Canadian Institute of Health Information (CIHI) National Ambulatory Care Reporting System and CIHI Discharge Abstract database were used to identify emergency department visits and hospital admissions, respectively. The diagnoses in these databases were coded using the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Canadian version. Information on physicians' specialty (for the OHIP billings) was obtained by linking the Institute for Clinical Evaluative Sciences (ICES) Physician Database with the OHIP Claims History Database. All data sets were linked and analyses performed at the ICES (www.ices.on.ca). The use of data in this study was authorized under section 45 of Ontario's Personal Health Information Protection Act, which does not require review by a Research Ethics Board.

Study subjects. We validated an algorithm to identify patients with PsA registered in the EMRPC by April 2016 using a previously abstracted cohort of PsA patients (12). The algorithm included PsA-related terms found in the EMR medical history fields and performed with a sensitivity of 78%, positive predictive value of 85%, and specificity of 100%, and a negative predictive value of 97% (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24146/abstract). The search was restricted to individuals ages 20 years or older who had a valid health insurance number and whose EMR start date was at least 2 years prior to the study date. The date of diagnosis of PsA (index date) was defined as the earliest entry of 1 of the following: 1) an OHIP physician service claim of an inflammatory arthritis diagnosis code by a rheumatologist; or 2) a hospital discharge abstract with a diagnosis code for PsA listed either as the reason for hospitalization or as a comorbid condition. In Ontario,

the diagnosis of PsA is typically determined by rheumatologists and not by family physicians. Therefore, to minimize inaccuracies in the date of diagnosis, we used the first claim of an inflammatory arthritis diagnosis code by a rheumatologist as the index date. Since Ontario administrative data are available at the individual level from 1991 onward, we included only patients who were diagnosed from January 1996 onward, which allowed a 5-year look-back window. Each patient with PsA was matched with 5 comparators by year of birth, sex, and family practice clinic. We excluded from the comparator group any subjects with diagnosis codes of spondyloarthritis (OHIP diagnosis code 721), ankylosing spondylitis (OHIP diagnosis code 720), and RA (OHIP diagnosis code 714) given by a rheumatologist within the 5-year look-back period. The comparators were assigned the same index date as their corresponding case.

Study outcomes. Health care utilization for musculoskeletal-related issues was evaluated during the 5-year period prior to the index dates in PsA patients and matched comparators. We evaluated a combination of physician services, emergency department visits, and diagnostic and therapeutic procedures for a wide range of noninflammatory musculoskeletal conditions (for a list of codes for diagnoses, procedures, and services, see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24146/abstract online). The following outcomes were evaluated: 1) visits to primary care physicians for musculoskeletal-related issues; 2) visits to a non-rheumatologist specialist; 3) visits to rheumatologists; 4) joint injections; 5) joint imaging, including radiographs, computed tomography, magnetic resonance imaging, and ultrasound; and 6) visits to emergency departments for musculoskeletal-related issues. The following services were not analyzed, as they are not covered by OHIP for the entire population: physical therapy, chiropractor visits, and occupational therapy. In addition, we obtained demographic information including age, sex, comorbidities (using the Johns Hopkins Aggregated Diagnosis Groups [ADGs] System, version 10, for the number of ADGs excluding rheumatologic diseases), socioeconomic status (by census neighborhood income quintiles), and rurality (classified as urban and rural). Annual direct health care costs, overall and for musculoskeletal-related health care services, were estimated using the methods described by Wodchis (14). The costs covered inpatient hospitalizations, visits to emergency departments and outpatient clinics, OHIP physician billing for consultations, procedures and diagnostic services, and prescribed medications (for seniors and social assistance recipients, who are covered by the Ontario Drug Benefit Plan). Musculoskeletal-related costs were calculated as the sum of fees paid by OHIP for the physician claims outlined above. This calculation included only the direct physician fees paid by OHIP and did not include the costs of services such as administrative overhead and equipment.

Statistical analysis. Data were expressed as the mean ± SD or median (25th and 75th percentile) for continuous variables and frequencies (%) for categorical variables. The percentage of patients with health care encounters and frequency of visits for the various categories of service were assessed for each of the 5 calendar years prior to the index date. Trends in the rates and probabilities of health care encounters were compared between patients with PsA and matched comparators using generalized estimating equations (GEEs) models with negative binomial distribution (for rates) and binary distribution (for probabilities). GEEs were used because they accounted for the matched nature of our study design. Mean annual total and musculoskeletal-related health care costs were reported for each of the 5 years prior to the index date. The mean annual costs between patients and controls were compared using *t*-test.

RESULTS

We studied 462 patients with PsA and 2,310 matched comparators without prior inflammatory arthritis with a mean ± SD age of 54.2 ± 13.8 years (55.6% females). As expected, patients with PsA and their comparators were well balanced with respect to the matching variables (age, sex) but also had a similar distribution of socioeconomic status and rurality. However, patients with PsA had a greater number of comorbidities, as indicated by higher ADGs (Table 1).

Table 1. Characteristics of the study population at index date*

Variable	PsA (n = 462)	Matched comparators† (n = 2,310)
Age, mean ± SD years	54.22 ± 13.82	54.22 ± 13.85
Sex, female	257 (55.6)	1,285 (55.6)
Number of ADGs		
0	0 (0)	121 (5.2)
1–5	100 (21.6)	1,236 (53.5)
6–10	238 (51.5)	741 (32.1)
11–15	106 (22.9)	192 (8.3)
16+	18 (3.9)	20 (0.9)
Census-based neighborhood income quintile		
1	77 (16.7)	389 (16.8)
2	91 (19.7)	413 (17.9)
3	93 (20.1)	423 (18.3)
4	86 (18.6)	488 (21.1)
5	112 (24.2)	579 (25.1)
Missing	≤5	18 (0.8)
Rurality		
Urban	387 (83.8)	1,923 (83.3)
Rural	66 (14.3)	313 (13.5)
Psoriasis billing code‡	260 (56.3)	53 (2.3)

* Values are the number (%) unless indicated otherwise. ADG = aggregate diagnosis group; PsA = psoriatic arthritis.
† Age-, sex-, and clinic-matched comparators without inflammatory arthritis.
‡ Within 5 years prior to the index date.

Table 2. The proportion of patients with at least 1 visit to a primary care physician for musculoskeletal causes prior to the index date in psoriatic arthritis (PsA) versus matched comparators*

Years prior to the index date	PsA (n = 462)	Matched comparators (n = 2,310)†	Odds ratio (95% CI)
1	42.2	25.4	2.14 (1.73–2.64)
2	38.5	25.0	1.66 (1.53–2.32)
3	35.3	23.9	1.73 (1.40–2.14)
4	34.8	23.5	1.74 (1.40–2.15)
5	35.5	23.7	1.76 (1.43–2.18)

* Values are the percentage unless indicated otherwise. 95% CI = 95% confidence interval.
† Age-, sex-, and clinic-matched comparators without inflammatory arthritis.

Overall, health care utilization for musculoskeletal-related issues was relatively higher in the years preceding the index date in patients with PsA compared with their matched comparators (Tables 2 and 3). Patients with PsA were more likely to visit their primary care physicians for musculoskeletal-related causes as early as 5 years prior to the diagnosis (odds ratio [OR] 1.76 [95% confidence interval (95% CI) 1.43–2.18]) (Table 2), with a trend toward an increase in the difference as time approached the index date (1 year prior to the index date OR 2.14 [95% CI 1.73–2.64]). The rate of annual visits to primary care physicians for musculoskeletal-related causes was also higher in patients with PsA during the 5-year prediagnosis period (Figure 1).

Similarly, the proportion of patients who required specialty care, diagnostic imaging, and procedures for musculoskeletal issues prior to the index date was higher in patients with PsA than their matched comparators (Table 3 and Supplementary Tables 1–5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24146/abstract online). PsA patients were also more likely to be assessed by non-rheumatologist musculoskeletal specialists (OR ranging from 1.59 to 2.03), visit an emergency department for musculoskeletal-related issues (OR 1.33 to 2.69), have joint imaging (OR 3.20 to 6.26), and joint injections (OR 4.63 to 9.26).

Patients with PsA were 4–5 times more likely to visit a rheumatologist during the 5-year period prior to the index date than their matched comparators (Table 3). The diagnosis codes assigned to these rheumatology visits prior to the index date were not related to an inflammatory arthritic condition and generally included nonspecific musculoskeletal conditions. The most common diagnosis codes administered by the rheumatologists were as follows: other disease of the musculoskeletal systems (OHIP diagnosis code 739); psoriasis (OHIP diagnosis code 696); osteoarthritis (OHIP diagnosis code 715); and leg cramps, leg pain, muscle pain, joint pain, and joint swelling (OHIP diagnosis code 781). Only ~25% of the patients were evaluated by a dermatologist for any reason prior to the index date (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24146/abstract).

Total and musculoskeletal-related health care costs were higher for patients who developed PsA than for their comparators at any point during the 5-year period prior to the index date (Figure 2 and Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24146/abstract). Total health care cost increased gradually in patients with PsA from a mean ± SD of $4,873 ± 8,480 (Canadian) 5 years prior to the index date to a mean ± SD of $6,995 ± 11,270 (Canadian) 1 year prior to the index date. This is compared to a relatively stable total mean health care cost of ~$2,500 (Canadian) observed in the comparator group. Similarly, mean musculoskeletal-related health care costs before the index date were 4 to 7 times higher in patients with PsA than in the comparator group.

DISCUSSION

In this population-based, matched cohort study, we characterized a prediagnosis period in patients with PsA by assessing their health care utilization due to musculoskeletal-related issues and compared it to that of matched comparators without inflammatory arthritis. We found that a significant proportion of patients with PsA experienced a prolonged prediagnosis period of musculoskeletal symptoms that was reflected by an increased number of visits to primary care physicians, musculoskeletal specialists, and emergency departments, as well as the use of diagnostic

Table 3. The proportion of patients with at least 1 visit to a musculoskeletal specialist for nonspecific musculoskeletal causes prior to the index date in psoriatic arthritis (PsA) versus matched comparators*

Years prior to the index date	Non-rheumatologist specialist†			Rheumatologist		
	PsA (n = 462)	Matched comparators (n = 2,310)‡	Odds ratio (95% CI)	PsA (n = 462)	Matched comparators (n = 2,310)‡	Odds ratio (95% CI)
1	14.7	7.8	2.03 (1.51–2.74)	12.1	2.5	5.36 (3.65–7.85)
2	17.5	7.7	2.55 (1.92–3.38)	11.7	2.6	4.96 (3.36–7.28)
3	14.2	8.1	1.88 (1.39–2.54)	10.6	2.0	5.72 (3.78–8.64)
4	13.2	7.1	1.99 (1.46–2.72)	9.8	2.0	5.44 (3.57–8.29)
5	10.4	6.8	1.59 (1.13–2.23)	8.4	2.1	4.34 (2.81–6.71)

* Values are the percentage unless indicated otherwise. 95% CI = 95% confidence interval.
† Orthopedic surgeons and sports physicians.
‡ Age-, sex-, and clinic-matched comparators without inflammatory arthritis.

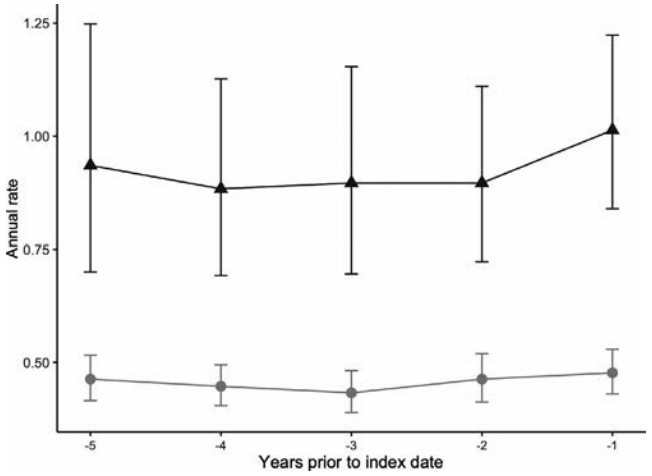


Figure 1. Annual rates of family physician visits for musculoskeletal issues prior to the index date for comparators (gray circles) and patients with psoriatic arthritis (black triangles). Error bars indicate the 95% confidence intervals.

imaging tests and procedures. This pattern was identified as early as 5 years prior to the diagnosis and was significantly different from that observed in the matched comparators. The novelty of the study is the use of musculoskeletal-related health care utilization as a surrogate to characterize the burden of musculoskeletal symptoms prior to the diagnosis of PsA.

It is well documented that the immunologic abnormalities of several rheumatic conditions, such as RA and lupus, start many years prior to the clinical presentation of the disease (15–17). Although there is less information about the prediagnosis phases in PsA, studies using various imaging modalities have documented subclinical inflammation in the joints and entheses in patients with psoriasis without musculoskeletal symptoms (18–21). Abnormal inflammatory biomarkers in the blood have been shown to predict the onset of PsA among psoriasis patients (22). Emerging data support the presence of a prodromal phase in a significant proportion of patients with PsA. This phase is characterized by nonspecific musculoskeletal symptoms such as fatigue, arthralgia, and stiffness in the absence of distinct findings on physical examination (23). The results of our study provide additional support for the presence of such a prodromal PsA phase. We found that more than one-third of the patients with PsA visited their family physician for musculoskeletal-related issues during the 5 years prior to the diagnosis of PsA. This proportion gradually increased and reached >42% of patients in the year immediately preceding the diagnosis, suggesting that PsA can be insidious at onset.

Delays in the diagnosis of PsA are well recognized. However, there is little information about the underlying factors that contribute to these delays in primary care settings. This issue is especially important in health care systems in which patients require a physician referral to access rheumatology care. Additionally, even in systems where unrestricted rheumatology access exists, family physicians play a crucial role in the coordination of specialty care and should raise the initial suspicion of PsA in the appropriate clinical context. Data from Ontario, Canada, suggested that one of the main factors contributing to the delay in the diagnosis of PsA is delayed referral to rheumatology, which is significantly longer than that for other rheumatic conditions, such as RA and ankylosing spondylitis (9). Our study showed that a significant proportion of patients with PsA received care from non-rheumatologist musculoskeletal specialists prior to diagnosis and visited emergency departments for their musculoskeletal symptoms. This pattern of care suggests that there are potential delays in rheumatology referrals, which ultimately lead to delays in diagnosis. Approximately 10% of the patients with PsA were seen by a rheumatologist during the prediagnosis period, which is at least 4 times

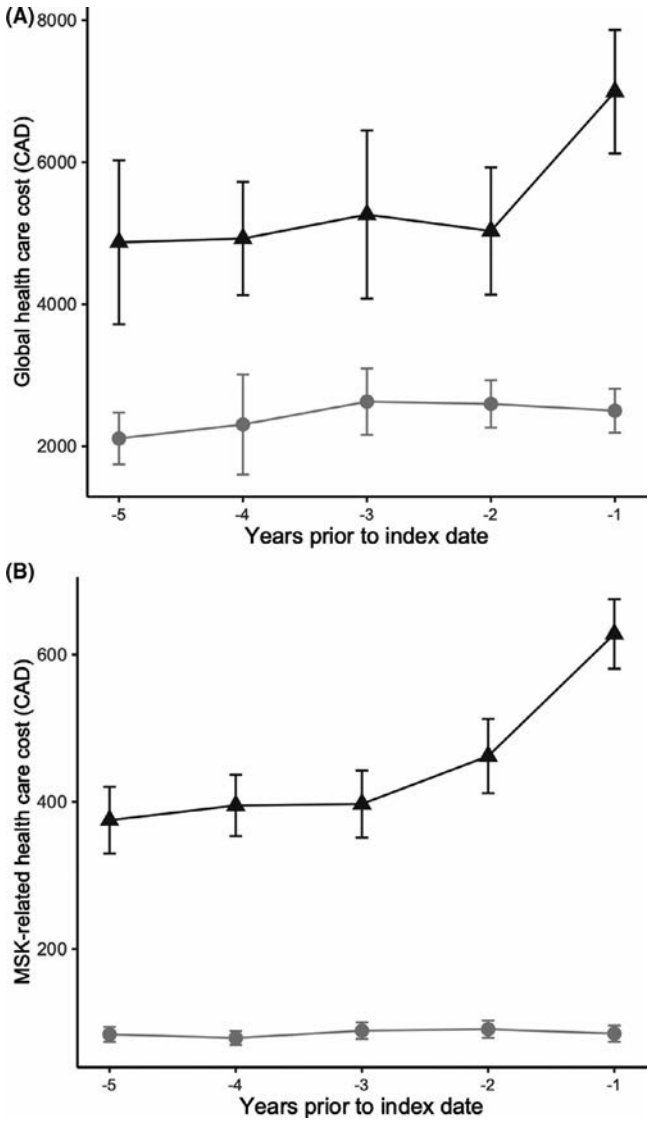


Figure 2. **A**, Mean total health care costs prior to the index date for comparators (gray circles) and patients with psoriatic arthritis (black triangles). **B**, Mean musculoskeletal-related health care costs prior to the index date for comparators (gray circles) and patients with psoriatic arthritis (black triangles). CAD = Canadian dollars. Error bars indicate the 95% confidence intervals.

higher than the rates in the comparators. We cannot rule out the possibility that this difference may be partially explained by an increased surveillance of patients with psoriasis who were referred to rheumatology specialists for suspected PsA. The musculoskeletal symptoms of these patients were largely attributed to nonspecific conditions by the assessing rheumatologists. This finding highlights the difficulties entailed in establishing the diagnosis of PsA in the early phases of the disease. The absence of reliable diagnostic tests, limited findings on musculoskeletal examination, and their overlap with other rheumatic conditions are some of the difficulties that rheumatologists face when examining patients with suspected PsA.

PsA is associated with substantial direct and indirect health care costs. A recent study that used a US claims database showed an incremental total direct health care cost for patients with established PsA of $18,482 compared to matched controls without PsA (24). Limited information exists regarding health care costs prior to the diagnosis of PsA. A nationwide study that used data from Danish registries showed that general health care costs for patients with PsA increased from less than €2,000 5 years prior to diagnosis to more than €5,000 around the time of diagnosis (25), which likely reflected the increased utilization of health care resources associated with reaching the diagnosis. Health care costs among those who developed PsA were higher than those in the general population during the 5-year prediagnosis period, which is similar to the findings of our study. Furthermore, our study also showed that health care costs due to musculoskeletal-related issues showed a similar trend and were higher for patients with PsA as early as 5 years prior to the diagnosis of the disease. It should be noted, however, that these numbers underestimate the true cost, as they do not include the costs of prescription medications among those not covered by the provincial drug plan, nonprescription medications, or other uncovered services received during the study period, such as physical therapy and chiropractic treatment.

Our study has limitations that merit emphasis. The first caveat is that we compared patients with PsA to subjects without inflammatory arthritis. This meant that, by design, the comparators would almost certainly have lower health care use for musculoskeletal conditions than the general unselected population. However, the estimated prevalence of inflammatory arthritis in the general population is <2%; thus, the impact of exclusion of these patients is likely minimal. In addition, some of the health care services related to musculoskeletal issues are not captured in the administrative databases; therefore, the true burden of health care utilization is underestimated. Furthermore, the single diagnosis code allowed for each visit may have led to underestimation of the true burden of musculoskeletal disease. Inaccuracies in the date of diagnosis may have occurred as well. In order to minimize this risk, we used a combined approach using a validated algorithm with high accuracy to first assemble a PsA cohort in the EMR as well as a number of diagnosis codes for various types of inflammatory arthritides to determine the date of diagnosis. Using this approach, we attempted to minimize the possibility that the actual date of diagnosis occurred prior to the index date, as rheumatologists are presumably more likely to assign a diagnosis of an inflammatory arthritis than a noninflammatory code when clearly indicated. Last, the increased total health care costs in patients with PsA were driven in part by other nonmusculoskeletal diseases, as the burden of comorbidities was elevated in patients with PsA compared to their matched comparators. This, however, does not explain the elevated health care costs due to musculoskeletal issues.

In conclusion, our study characterized a prediagnosis period in PsA and supports the notion that a prodromal PsA phase occurs in a significant proportion of patients. This phase is characterized by musculoskeletal symptoms and leads to increased health care utilization and costs, including performance of diagnostic tests, visits to family doctors, musculoskeletal specialist consultations, and emergency department visits. This pattern reveals some of the underlying causes of delays in the diagnosis of PsA and highlights the need for diagnostic strategies and novel reliable biomarkers to aid in the early diagnosis of 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. Eder had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Eder, Tu, Rosen, Alhusayen, Campbell, Bernatsky, Gladman, Paterson, Widdifield.

Acquisition of data. Cheng, Young.

Analysis and interpretation of data. Eder, Cheng, Young, Paterson, Widdifield.

REFERENCES

1. Ritchlin CT, Colbert RA, Gladman DD. Psoriatic arthritis. N Engl J Med 2017;376:2095–6.
2. Geijer M, Lindqvist U, Husmark T, Alenius GM, Larsson PT, Teleman A, et al. The Swedish early psoriatic arthritis registry 5-year followup: substantial radiographic progression mainly in men with high disease activity and development of dactylitis. J Rheumatol 2015;42:2110–7.
3. Coates LC, Moverley AR, McParland L, Brown S, Navarro-Coy N, O'Dwyer JL, et al. Effect of tight control of inflammation in early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. Lancet 2015;386:2489–98.
4. Haroon M, Gallagher P, FitzGerald O. Diagnostic delay of more than 6 months contributes to poor radiographic and functional outcome in psoriatic arthritis. Ann Rheum Dis 2015;74:1045–50.
5. Gladman DD, Thavaneswaran A, Chandran V, Cook RJ. Do patients with psoriatic arthritis who present early fare better than those presenting later in the disease? Ann Rheum Dis 2011;70:2152–4.
6. McHugh NJ. Early psoriatic arthritis. Rheum Dis Clin North Am 2015;41:615–22.
7. Villani AP, Rouzaud M, Sevrain M, Barnetche T, Paul C, Richard MA, et al. Prevalence of undiagnosed psoriatic arthritis among psoriasis patients: systematic review and meta-analysis. J Am Acad Dermatol 2015;73:242–8.

8. Sorensen J, Hetland ML, all departments of rheumatology in Denmark. Diagnostic delay in patients with rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis: results from the Danish nationwide DANBIO registry. Ann Rheum Dis 2015;74:e12.
9. Widdifield J, Bernatsky S, Thorne JC, Bombardier C, Jaakkimainen RL, Wing L, et al. Wait times to rheumatology care for patients with rheumatic diseases: a data linkage study of primary care electronic medical records and administrative data. CMAJ Open 2016;4:E205–12.
10. McArdle A, Pennington S, FitzGerald O. Clinical features of psoriatic arthritis: a comprehensive review of unmet clinical needs. Clin Rev Allergy Immunol 2018;55:271–94.
11. Favier G, Gladman DD, Merola JF, Armstrong AW, Boehncke WH, Helliwell PS. Benchmarking care in psoriatic arthritis - the QUANTUM report: a report from the GRAPPA 2016 Annual Meeting. J Rheumatol 2017;44:674–8.
12. Eder L, Widdifield J, Rosen CF, Cook R, Lee KA, Alhusayen R, et al. Trends in the prevalence and incidence of psoriasis and psoriatic arthritis in Ontario, Canada: a population-based study. Arthritis Care Res (Hoboken) 2018;71:1084–91.
13. Tu K, Mitiku TF, Ivers NM, Guo H, Lu H, Jaakkimainen L, et al. Evaluation of electronic medical record administrative data linked database (EMRALD). Am J Manag Care 2014;20:e15–21.
14. Wodchis W, Bushmeneva K, Nikitovic M, McKillop I. Guidelines on person-level costing using administrative databases in Ontario. Working Paper Series. Vol. 1. Toronto: Health System Performance Research Network; 2013.
15. Rantapää-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;48:2741–9.
16. Nielen MM, van Schaardenburg D, Reesink HW, Twisk JW, van de Stadt RJ, van der Horst-Bruinsma IE, et al. Simultaneous development of acute phase response and autoantibodies in preclinical rheumatoid arthritis. Ann Rheum Dis 2006;65:535–7.
17. Hu ZD, Deng AM. Autoantibodies in pre-clinical autoimmune disease. Clin Chim Acta 2014;437:14–8.
18. Mathew AJ, Bird P, Gupta A, George R, Danda D. Magnetic resonance imaging (MRI) of feet demonstrates subclinical inflammatory joint disease in cutaneous psoriasis patients without clinical arthritis. Clin Rheumatol 2018;37:383–8.
19. Eder L, Jayakar J, Thavaneswaran A, Haddad A, Chandran V, Salonen D, et al. Is the MAdrid Sonographic Enthesitis Index useful for differentiating psoriatic arthritis from psoriasis alone and healthy controls? J Rheumatol 2014;41:466–72.
20. Ash ZR, Tinazzi I, Gallego CC, Kwok C, Wilson C, Goodfield M, et al. Psoriasis patients with nail disease have a greater magnitude of underlying systemic subclinical enthesopathy than those with normal nails. Ann Rheum Dis 2012;71:553–6.
21. Ogdie A, Weiss P. The epidemiology of psoriatic arthritis. Rheum Dis Clin North Am 2015;41:545–68.
22. Abji F, Pollock RA, Liang K, Chandran V, Gladman DD. CXCL10 is a possible biomarker for the development of psoriatic arthritis among patients with psoriasis. Arthritis Rheumatol 2016;68:2911–6.
23. Eder L, Polachek A, Rosen CF, Chandran V, Cook R, Gladman DD. The development of psoriatic arthritis in patients with psoriasis is preceded by a period of nonspecific musculoskeletal symptoms: a prospective cohort study. Arthritis Rheumatol 2017;69:622–9.
24. Merola JF, Herrera V, Palmer JB. Direct healthcare costs and comorbidity burden among patients with psoriatic arthritis in the USA. Clin Rheumatol 2018;37:2751–61.
25. Kristensen LE, Jorgensen TS, Christensen R, Gudbergsen H, Dreyer L, Ballegaard C, et al. Societal costs and patients' experience of health inequities before and after diagnosis of psoriatic arthritis: a Danish cohort study. Ann Rheum Dis 2017;76:1495–501.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 687–692
DOI 10.1002/acr.23581

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Relationship Between Self-Reported Restless Sleep and Objectively Measured Physical Activity in Adults With Knee Osteoarthritis

Abigail L. Gilbert,[1] Jungwha Lee,[2] Jing Song,[2] Pamela A. Semanik,[3] Linda S. Ehrlich-Jones,[4] C. Kent Kwoh,[5] Dorothy D. Dunlop,[2] and Rowland W. Chang[2]

Objective. Despite the numerous health benefits of physical activity, inactivity is endemic among adults with knee osteoarthritis (OA). Because sleep quality may be a target in order to improve physical activity behavior, we investigated the cross-sectional relationship between restless sleep and physical activity in participants with or at risk for knee OA.

Methods. We analyzed accelerometer-measured physical activity and clinical data of participants included in the Osteoarthritis Initiative (OAI). We used multiple regression analysis to evaluate physical activity for participants, who were grouped by the reported frequency of restless sleep, and adjusted for demographic and medical confounders.

Results. Of the 1,892 OAI participants for whom complete data were available, 300 participants (16%) reported restless sleep ≥3 days in the past week. Participants who reported restless sleep for much of the time (3–4 days/week) and most of the time (5–7 days/week) had 11.9% and 23.7% less weekly minutes of moderately vigorous activity, respectively, compared to participants who reported rarely restless sleep (<1 day/week) (*P* for trend 0.021). These differences persisted after accounting for age, sex, race, body mass index, medical comorbidity, and knee OA severity and pain (*P* for trend 0.023). Differences related to restless sleep were largely attenuated by the presence of high depressive symptoms and low energy levels.

Conclusion. Poor sleep quality is associated with less physical activity in persons with or at risk for knee OA. Future studies are needed to determine the mechanisms of how poor sleep and physical activity are related, how energy and depression mediate these relationships, and whether interventions that improve sleep quality might result in increased physical activity.

Introduction

One in 10 adults ages ≥65 years experiences knee osteoarthritis (OA) and the associated progressive pain and disability (1). Physical activity is well known to improve long-term functional status and to preserve independence in adults with knee OA (2). Despite current recommendations to promote physical activity as a primary treatment, almost half of adults with knee OA are physically inactive, not even performing a single 10-minute session of moderately intense physical activity in a week (2–4). Over half of individuals with knee OA report sleep disturbance, including difficulty initiating or maintaining night-time sleep (5). This is in contrast to one-third of all adults reporting sleep complaints (6). Poor sleep quality is a modifiable risk factor (7) but is under-investigated as a

This article was prepared using an Osteoarthritis Initiative (OAI) public-use data set, and its contents do not necessarily reflect the opinions or views of the OAI Study Investigators, the NIH, or the private funding partners of the OAI. The OAI is a public–private partnership between the NIH (contracts N01-AR-2-2258, N01-AR-2-2259, N01-AR-2-2260, N01-AR-2-2261, and N01-AR-2-2262) and private funding partners (Merck Research Laboratories, Novartis Pharmaceuticals, GlaxoSmithKline, and Pfizer, Inc.) and is conducted by the OAI Study Investigators. Private sector funding for the OAI is managed by the Foundation for the NIH. The authors of this article are not part of the OAI investigative team.

Supported in part by the NIH (grants R01-AR-054155, R21-AR068500, P60-AR-064464, and T32-AR-007611 from the National Institute for Arthritis and Musculoskeletal Diseases) and the National Center for Advancing Translational Sciences (UL1-TR-001422).

[1]Abigail L. Gilbert, MD, MSCI: University of North Carolina, Chapel Hill; [2]Jungwha Lee, PhD, MPH, Jing Song, MS, Dorothy D. Dunlop, PhD, Rowland W. Chang, MD, MPH: Northwestern University Feinberg School of Medicine, Chicago, Illinois; [3]Pamela A. Semanik, PhD: Rush University, Chicago, Illinois; [4]Linda S. Ehrlich-Jones, PhD, RN: Shirley Ryan Ability Lab, Chicago, Illinois; [5]C. Kent Kwoh, MD: University of Arizona, Tucson.

Address correspondence to Abigail L. Gilbert, MD, MSCI, Thurston Arthritis Research Center, 3300 Thurston Building, CB #7280, Chapel Hill, NC 27599. E-mail: abigail.gilbert@unc.edu.

Submitted for publication August 10, 2017; accepted in revised form April 10, 2018.

SIGNIFICANCE & INNOVATIONS

- Physical inactivity is endemic in individuals with knee osteoarthritis (OA), and increasing physical activity has been shown to improve long-term functional status.
- The relationship between sleep and physical activity in individuals with knee OA remains unknown.
- This study, which examined the cross-sectional relationship between self-reported restless sleep and objectively measured physical activity in adults with or at risk for knee OA, demonstrated a significant trend between greater frequency of restless sleep and less time engaged in moderately vigorous physical activity.
- Future research should characterize the mechanisms of how sleep disturbance and physical activity are related, and should evaluate whether a focus on sleep as part of a multitargeted intervention results in more sustainable increases in physical activity.

determinant of physical activity. Improving sleep quality may be a novel target to increase physical activity behavior; the relationship between sleep quality and physical activity may be bidirectional, and improving physical activity may also benefit sleep.

Good sleep quality is important for good health, yet chronic pain, including pain from OA, interferes with sleep. In general populations, sleep and physical activity have been shown to have an association (6,8); however, the relationship between sleep quality and physical activity in adults with knee OA remains unknown. This is likely a complex relationship, because sleep disturbance is associated with pain, depression, and fatigue and it exacerbates depression among individuals with high levels of pain (5,9). Individuals with major depressive disorder are less physically active than individuals without depression (10). If restless sleep increases depressive symptoms and fatigue, and depression and fatigue lead to decreased physical activity, depressive symptoms and low energy may serve as mediators to explain the mechanism by which restless sleep results in less physical activity. The goal of this study was to investigate the relationship between sleep quality (measured by self-reported restless sleep) and physical activity (objectively measured) in adults with or at high risk for knee OA. We hypothesized that there was an inverse relationship between restless sleep and physical activity (after controlling for confounders), and that energy and depressive symptoms from sleep disturbance may explain the pathway between restless sleep and physical activity.

Materials and methods

Study population. The Osteoarthritis Initiative (OAI) is a multicenter prospective study that investigates risk factors and biomarkers for the progression and/or onset of knee OA (http://www.oai.ucsf.edu/datarelease/About.asp). The study design and eligibility criteria of the OAI have been described in detail elsewhere (11). Participants were a subcohort of the OAI and were enrolled into an accelerometer ancillary study conducted at the OAI 2008–2010 clinic visit (2008–2010) (OAI 4-year follow-up) (4). At enrollment, the OAI recruited 4,796 men and women ages 45–79 years from 4 clinical sites, who had or were at high risk for developing symptomatic, radiographic knee OA. High risk was defined as the presence of ≥2 eligibility risk factors (e.g., age >70 years, sex- and age-specific overweight criteria, prior knee injury, prior knee surgery, family history of total knee replacement for OA, Heberden's nodes, and/or frequent knee symptoms).

The study population was drawn from 2,127 people who were enrolled in an OAI accelerometer monitoring substudy at the OAI 48-month follow-up visit (2008–2010). A total of 1,705 participants were not invited to participate because their 48-month visit was not during the study dates, 585 participants declined to participate in the accelerometer study, 70 were deceased, and 309 were not available because they did not attend the 48-month visit. Approval was obtained from the institutional review board at each OAI site and at Northwestern University. Each participant provided written informed consent.

Restless sleep. The frequency of restless sleep was evaluated by participants' responses to a question on the Center for Epidemiologic Studies Depression Scale (CES-D), which asked how often in the past week their sleep was restless, with the following response options: rarely (rarely or none of the time; <1 day), some (some or a little of the time; 1–2 days), moderate/much (occasionally or a moderate amount of time; 3–4 days), and most (most or all of the time; 5–7 days) (12). Accelerometer monitoring and the CES-D results were both from the substudy baseline (i.e., the OAI 48-month study visit).

Covariates. Demographic factors included age, sex, and race (4). Medical covariates included body mass index (BMI), medical comorbidity, knee OA severity, and pain (4). BMI was calculated as kg/m^2, and height was measured using calibrated, wall-mounted stadiometers. Body weight was measured using calibrated standard balance beam scales. Medical comorbidity was ascertained using the Charlson comorbidity index, a validated self-administrated questionnaire evaluating comorbid chronic conditions. Knee OA severity was identified by the worse Kellgren/Lawrence grade score of both knees, assessed from fixed-flexion knee radiography protocol. Self-reported knee pain was measured on a 5-point Likert scale from the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), modified to ask separately about the right and left knee symptoms in the past 7 days. The WOMAC pain score range is 0–20, and higher numbers represent worse symptoms. Energy level was based on a Likert scale response to the question "during the past 4 weeks, did you have a lot of energy?", which was ascertained from an item on the Short Form 12 (SF-12) Health Survey. CES-D scores were calculated excluding the question of restless sleep and

included the remaining 19 questions from the full 20-item scale. Patients were considered to have evidence of a high level of depressive symptoms if they had a score of ≥16 on the modified CES-D (12).

Physical activity assessment. Physical activity was monitored using the GT1M uniaxial accelerometer (ActiGraph). Trained research personnel gave uniform scripted in-person instructions to wear the accelerometer for 7 consecutive days on a belt at the natural waistline in line with the right axilla upon arising in the morning until bedtime, except during water activities. Accelerometer data were analytically filtered using validated methodology (13). Periods of non-wear were defined as ≥90 minutes with zero activity counts (allowing for 2 consecutive interrupted minutes with counts <100) (14). We identified participants with 4–7 valid monitoring days (i.e., ≥10 wear hours per day) needed for reliable physical activity estimates (14). Thresholds used by the National Cancer Institute on a minute-by-minute basis were applied to identify moderate-to-vigorous (≥2,020 counts/minute) intensity activity. Minutes of weekly, moderately vigorous activity were summed from the daily totals over the monitored hours and averaged across valid monitored days; for individuals for whom 4, 5, or 6 valid days of activity

Table 1. Participant characteristics by self-reported restless sleep during the past week (n = 1,892)*

	Rarely/none of the time (<1 day)	Some of the time (1–2 days)	Moderate/much of the time (3–4 days)	Most/all of the time (5–7 days)	*P* for trend†
No. (%)	751 (39.7)	841 (44.5)	168 (8.9)	132 (7.0)	–
Age, years	65.4 ± 9.0	65.5 ± 9.0	63.7 ± 9.4	63.0 ± 9.3	0.003
Female, %	53.9	55.1	60.7	60.6	0.255
White, %	81.5	86.0	84.5	82.6	0.108
BMI, kg/m^2	28.2 ± 4.7	28.4 ± 4.8	28.5 ± 5.4	29.7 ± 5.1	0.011
Normal (<25), %	25.6	26.3	28.6	15.9	0.059
Overweight (25–30), %	41.4	38.4	33.3	43.9	
Obese (≥30), %	33.0	35.3	38.1	40.2	
K/L grade, %‡					
0–1	38.4	38.8	44.1	41.7	0.477
2–3	52.7	52.9	44.1	50.0	
4	8.9	8.3	11.9	8.3	
WOMAC pain§	2.1 ± 2.8	2.8 ± 3.4	3.5 ± 3.8	4.6 ± 4.3	<0.001
Charlson comorbidity index score, %					
0	71.9	71.3	63.1	65.9	0.004
1–2	24.9	23.2	27.4	27.3	
3+	3.2	5.5	9.5	6.8	
Depression, %	3.6	7.3	22.6	40.2	<0.001
Energy level (had a lot of energy), %					
All of the time	11.5	3.6	1.8	0.0	<0.001
Most of the time	61.7	58.7	35.1	25.0	
Some of the time	21.4	27.9	39.3	36.4	
A little of the time	4.5	7.7	19.1	23.5	
None of the time	0.9	2.0	4.8	15.2	
Weekly, moderately vigorous activity minutes, median (IQR)	79 (28–188)	83 (27–192)	77 (22–200)	58 (19.4–142)	0.189

* Values are the mean ± SD unless otherwise indicated. IQR = interquartile range.
† Mantel-Haenszel chi-square test used to test for trend (1 df) except for race and sex comparisons, which used chi-square test for overall differences, analysis of variance for continuous factors age, body mass index (BMI), and Western Ontario McMaster Universities Osteoarthritis Index (WOMAC) pain, and quantile regression for physical activity minutes/week.
‡ Kellgren/Lawrence (K/L) grade for severity of knee osteoarthritis.
§ WOMAC pain score modified to determine right and left knee symptoms separately, range 0–20, with worse knee reported.
¶ High depressive symptoms as defined by modified Center for Epidemiological Studies Depression Scale ≥16.
# Item from Short Form 12 Health Survey.

Table 2. Difference in weekly, moderately vigorous physical activity minutes compared to restless sleep rarely or none of the time (n = 1, 892)*

	Rarely/none of the time (<1 day)	Some of the time (1–2 days)	Much of the time (3–4 days)	Most/all of the time (5–7 days)	*P* for trend
Unadjusted model, %	ref.	−1.1 (−14.5, 14.4)	−14.9 (−33.5, 9.0)	−25.2 (−43.1, −1.7)	0.019
Adjusted model 1, %†	ref.	−1.5 (−13.2, 11.7)	−23.8 (−38.5, −5.6)	−36.3 (−49.7, −19.3)	<0.001
Adjusted model 2, % ‡	ref.	4.9 (−7.0, 18.4)	−12.4 (−28.7, 7.7)	−20.4 (−36.8, 0.3)	0.020

* Except where indicated otherwise, values are the percentage difference (95% confidence interval) for each restless sleep category compared to rarely or none of the time. ref. = reference.
† Adjusted for age, sex, and race.
‡ Adjusted for model 1 and body mass index, Kellgren/Lawrence grade, Charlson comorbidity index.

were monitored, weekly activity minutes were estimated as 7 average daily activity minutes spent in moderately vigorous activity.

Statistical analysis. We compared moderately vigorous physical activity minutes per week across 4 self-reported restless sleep categories. The outcome of weekly minutes of moderately vigorous physical activity was log-transformed due to the skewed distribution. For 13 participants with 0 weekly moderate-vigorous physical activity minutes, we added 0.1 and then log-transformed. Multiple linear regression was used to estimate the difference between restless sleep groups in outcome using the reference group that reported rarely having restless sleep. Regression findings on log-transformed outcomes can be validly translated as the percentage of difference in the weekly minutes of moderately vigorous physical activity among each restless sleep group and reference group by using the equation $(e^{\wedge(\text{coefficient})}-1) \times 100\%$, where the coefficient is estimated for each restless sleep group from the log-transformed regression model (15). Hierarchical multiple regression adjusted for potential confounders, first adjusting for demographic factors (age, gender, and race, because prior studies have demonstrated that women, older individuals, and nonwhite individuals are less likely to be physically active) (4), then adjusting for medical characteristics (BMI, knee OA severity, WOMAC pain, and medical comorbidity, because individuals who have high BMI, more knee pain, and more medical comorbidities are more likely to be physically inactive) (16). To explore if depressive symptoms and low energy mediate the relationship between restless sleep and physical activity, we added the presence of high depressive symptoms and low energy to the final model to see if the magnitude of the relationship between restless sleep and physical activity was diminished. Sex, race, and presence of high depressive symptoms were entered as categorical variables, and all others were treated as continuous. All statistical analyses were completed using SAS software (version 9.4).

Results

Of the 2,127 accelerometer study participants, 1,927 individuals (91%) had complete physical activity outcomes (4–7 valid days of monitoring). Incomplete covariate data were available for <2% of the participants. The analysis sample included 1,892 participants. As shown in Table 1, 168 participants (8.9%) reported moderately restless sleep much of the time (3–4 days in the past week), and 132 participants (7.0%) reported restless sleep most of the time (5–7 days in the past week). Participants who reported restless sleep most or all of the time had high depressive symptoms (40.2%) compared to only 3.6% of participants who reported rarely restless sleep. Table 2 shows that, compared to participants reporting rarely restless sleep (*P* for trend 0.019), participants who reported restless sleep for a moderate amount of the time and those who reported restless sleep most of the time had 14.9% and 25.2% less weekly minutes of moderately vigorous physical activity, respectively. After adjusting for potential demographic confounders (age, sex, and race), the trend became stronger for spending less time in moderately vigorous physical activity with more frequently reported restless sleep (*P* for trend < 0.001). Compared to the reference group, which reported restless sleep rarely/none of the time, participants who reported restless sleep for much of the time and most/all of the time had 23.8% and 36.3% less likely weekly, moderately vigorous physical activity, respectively. This translates to 19 and 29 fewer minutes of weekly, moderately vigorous physical activity for each restless sleep group compared to the reference group (reporting restless sleep rarely/none of the time), who participated in an average of 80 minutes of weekly, moderately vigorous physical activity. The trend toward significance persisted with additional adjustment for potential confounders relating to general health, pain, and knee OA severity (*P* for trend 0.020). When we evaluated if energy and high depressive symptoms mediate this relationship, we demonstrated that adjusting for these variables largely attenuated the relationship between increased frequency of restless sleep and less time spent in moderately vigorous physical activity with individuals with restless sleep much of the time and individuals with restless sleep most of the time, respectively, performing 2.0% (95% confidence interval [95% CI] −17.3%, 25.7%) and 1.8% (95% CI −20.0%, 29.7%) more physical activity (*P* for trend 0.950).

Discussion

In this study, we evaluated the relationship between the frequency of self-reported restless sleep and objectively measured physical activity in adults with or at high risk for knee OA. We demonstrated a significant trend between greater frequency of restless sleep and less time engaged in moderately vigorous physical activity after adjusting for potential demographic and medical confounders. The difference in physical activity was attenuated by differences in how often participants reported having a lot of energy and the frequency of participants reporting high depressive symptoms.

Previous studies in general adult populations have examined the relationship between sleep quality and physical activity and found a bidirectional relationship: participants with higher levels of physical activity are less likely to report sleep complaints and, conversely, better sleep efficiency is associated with more daily physical activity (6,17). A study by Mesci et al is the only study that we are aware of that has examined the relationship between sleep quality and physical activity in individuals with knee OA; however, in contrast to our study, the study by Mesci et al did not find an association between self-reported sleep quality (as measured by the Pittsburgh Sleep Quality Index) and physical activity (classified by meeting physical activity guidelines by self-report) (18). The Pittsburgh Sleep Quality Index evaluates several different sleep components, whereas our study focused solely on restless sleep. In addition, our study objectively measured physical activity and was analyzed on a continuous scale, which most likely yields more powerful data than the self-reported, categorical physical activity classification used in the study by Mesci and colleagues.

Our study demonstrated that depression and energy were important mediators in the relationship between restless sleep and physical activity. Prior studies have shown that increased fatigue is associated with reduced physical activity in individuals with knee or hip osteoarthritis and fatigue (19). It is likely that participants who reported restless sleep experience significant fatigue. Fatigue and depression may be part of the causal pathway explaining the mechanism by which restless sleep resulted in decreased physical activity. Individuals with major depressive disorder are known to have decreased physical activity compared to individuals without depression (10). Among individuals with knee OA, sleep disturbance is known to exacerbate depression in those with high levels of pain (5). It is likely that increased depressive symptoms and low energy are both part of the causal pathway partially explaining the relationship between restless sleep and decreased physical activity; however, individuals who are physically inactive may experience more restless sleep and low energy and be more vulnerable to depression as a result of low levels of physical activity. Thus the relationships are likely to be bidirectional. Future research should explore if improving sleep quality may help individuals experience less depressive symptoms and have more energy, resulting in more participation in physical activity. Future studies could also more fully evaluate how fatigue and depression mediate the relationship between sleep and physical activity.

Our study had some limitations. Due to log adjustments for skewed data, we were unable to directly transform the results of the regression analyses back to absolute minutes, limiting our results to percent difference between restless groups. CES-D has not previously been examined as a measure of sleep quality, and information regarding sleep duration was unavailable. We are not aware of any studies validating using the SF-12 question as a measure of energy, or any studies validating the CES-D using 19 of the 20 questions. Although we adjusted for known confounders (e.g., demographic, medical covariates including depression), some of the differences in physical activity may be explained by residual confounding between the groups. Finally, we were unable to assess the temporal relationship between restless sleep and physical activity, given the cross-sectional nature of this analysis. Future studies are needed to better characterize the bidirectional relationship between sleep and physical activity. Strengths of our study include a large cohort with objectively measured physical activity.

In conclusion, we demonstrated a significant relationship between greater frequency of restless sleep and less time engaged in moderately vigorous physical activity among participants with or at high risk for knee OA. Future research is needed to characterize mechanisms of how poor sleep quality might result in less physical activity or vice versa, and to determine if improvement of sleep quality alone or in conjunction with other interventions can increase physical activity in those with restless sleep.

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. Gilbert 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. Gilbert, Kwoh, Dunlop, Chang.
Acquisition of data. Song, Semanik, Kwoh, Dunlop.
Analysis and interpretation of data. Gilbert, Lee, Song, Semanik, Ehrlich-Jones, Kwoh, Dunlop, Chang.

REFERENCES

1. Deshpande BR, Katz JN, Solomon DH, Yelin EH, Hunter DJ, Messier SP, et al. The number of persons with symptomatic knee osteoarthritis in the United States: impact of race/ethnicity, age, sex, and obesity. Arthritis Care Res (Hoboken) 2016;68:1743–50.
2. Kretzschmar M, Lin W, Nardo L, Joseph GB, Dunlop DD, Heilmeier U, et al. Association of physical activity measured by accelerometer, knee joint abnormalities, and cartilage T2 measurements obtained from 3T magnetic resonance imaging: data from the Osteoarthrtits Initiative. Arthritis Care Res (Hoboken) 2015;67:1272–80.
3. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic

therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012;64:465–74.

4. Dunlop DD, Song J, Semanik PA, Chang RW, Sharma L, Bathon JM, et al. Objective physical activity measurement in the osteoarthritis initiative: are guidelines being met? Arthritis Rheum 2011;63: 3372–82.
5. Parmelee PA, Tighe CA, Dautovich ND. Sleep disturbance in osteoarthritis: linkages with pain, disability, and depressive symptoms. Arthritis Care Res (Hoboken) 2015;67:358–65.
6. Kline CE. The bidirectional relationship between exercise and sleep: implications for exercise adherence and sleep improvement. Am J Lifestyle Med 2014;8:375–9.
7. Vitiello MV, Rybarczyk B, Von Korff M, Stepanski EJ. Cognitive behavioural therapy for insomnia improves sleep and decreases pain in older adults with co-morbid insomnia and osteoarthritis. J Clin Sleep Med 2009;5:355–62.
8. Holfeld B, Ruthig JC. A longitudinal examination of sleep quality and physical activity in older adults. J Appl Gerontol 2014;33: 791–807.
9. Vitiello MV, McCurry SM, Shortreed SM, Baker LD, Rybarczyk BD, Keefe FJ, et al. Short-term improvement in insomnia symptoms predicts long-term improvements in sleep, pain, and fatigue in older adults with comorbid osteoarthritis and insomnia. Pain 2014;155: 1547–54.
10. Schuch F, Vancampfort D, Firth J, Rosenbaum S, Ward P, Reichert T, et al. Physical activity and sedentary behavior in people with major depressive disorder: a systematic review and meta-analysis. J Affect Disord 2017;210:139–50.
11. Nevitt MC, Felson DT, Lester G. The Osteoarthritis Initiative Protocol for the Cohort Study. URL: https://oai.epi-ucsf.org/datarelease/docs/StudyDesignProtocol.pdf.
12. Smarr KL, Keefer AL. Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). Arthritis Care Res (Hoboken) 2011;63 Suppl:S454–66.
13. Song J, Semanik P, Sharma L, Chang RW, Hochberg MC, Mysiw WJ, et al. Assessing physical activity in persons with knee osteoarthritis using accelerometers: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2010;62:1724–32.
14. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci in Sports Exec 2008;40:181–8.
15. Vittinghoff E. Regression methods in biostatistics : linear, logistic, survival, and repeated measures models. 2nd ed. New York: Springer; 2012.
16. Lee J, Song J, Hootman JM, Semanik PA, Chang RW, Sharma L, et al. Obesity and other modifiable factors for physical inactivity measured by accelerometer in adults with knee osteoarthritis. Arthritis Care Res (Hoboken) 2013;65:53–61.
17. Hartescu I, Morgan K, Stevinson CD. Sleep quality and recommended levels of physical activity in older people. J Aging Phys Act 2016;24:201–6.
18. Mesci E, Icagasioglu A, Mesci N, Turgut ST. Relation of physical activity level with quality of life, sleep and depression in patients with knee osteoarthritis. North Clin Istanb 2015;2:215–21.
19. Murphy SL, Alexander NB, Levoska M, Smith DM. Relationship between fatigue and subsequent physical activity among older adults with symptomatic osteoarthritis. Arthritis Care Res (Hoboken) 2013;65:1617–24.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 693–701
DOI 10.1002/acr.24179

American College of Rheumatology
Empowering Rheumatology Professionals

Precision Medicine Approach to Develop and Internally Validate Optimal Exercise and Weight-Loss Treatments for Overweight and Obese Adults With Knee Osteoarthritis: Data From a Single-Center Randomized Trial

Xiaotong Jiang,[1] Amanda E. Nelson,[1] Rebecca J. Cleveland,[1] Daniel P. Beavers,[2] Todd A. Schwartz,[1] Liubov Arbeeva,[1] Carolina Alvarez,[1] Leigh F. Callahan,[1] Stephen Messier,[3] Richard Loeser,[1] and Michael R. Kosorok[1]

Objective. To apply a precision medicine approach to determine the optimal treatment regime for participants in an exercise (E), dietary weight loss (D), and D + E trial for knee osteoarthritis that would maximize their expected outcomes.

Methods. Using data from 343 participants of the Intensive Diet and Exercise for Arthritis (IDEA) trial, we applied 24 machine-learning models to develop individualized treatment rules on 7 outcomes: Short Form 36 physical component score, weight loss, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain/function/stiffness scores, compressive force, and interleukin-6 level. The optimal model was selected based on jackknife value function estimates that indicate improvement in the outcomes if future participants follow the estimated decision rule compared to the optimal single, fixed treatment model.

Results. Multiple outcome random forest was the optimal model for the WOMAC outcomes. For the other outcomes, list-based models were optimal. For example, the estimated optimal decision rule for weight loss indicated assigning the D + E intervention to participants with baseline weight not exceeding 109.35 kg and waist circumference above 90.25 cm, and assigning D to all other participants except those with a history of a heart attack. If applied to future participants, the optimal rule for weight loss is estimated to increase average weight loss to 11.2 kg at 18 months, contrasted with 9.8 kg if all participants received D + E (P = 0.01).

Conclusion. The precision medicine models supported the overall findings from IDEA that the D + E intervention was optimal for most participants, but there was evidence that a subgroup of participants would likely benefit more from diet alone for 2 outcomes.

INTRODUCTION

Knee osteoarthritis (OA) is one of the most common forms of arthritis worldwide, accounting for a significant proportion of pain and disability in the adult population (1). Known risk factors for knee OA include older age (especially ≥55 years), increased body weight, previous joint injury, and genetics (2). Clinical trials in overweight and obese adults with symptomatic knee OA have shown that weight loss and exercise interventions can improve pain and function, although not all individuals achieve a similar amount of benefit (3–5). Overweight and obese patients with knee OA will want to know whether they need to diet and exercise, or whether exercise or diet alone would be sufficient. Likewise, clinicians would value additional insights into which specific therapies are most likely to benefit particular patients in a given situation. To address these questions, we used machine learning tools to

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grants R01-AR-052528 [for the IDEA trial] and P60-AR-064166, P30-AR-072580, P01-CA-142538, and UL1-TR-002489 [for precision medicine analysis]).

[1]Xiaotong Jiang, PhD, Amanda E. Nelson, MD, MSCR, Rebecca J. Cleveland, PhD, Todd A. Schwartz, DrPH, Liubov Arbeeva, MS, Carolina Alvarez, MS, Leigh F. Callahan, PhD, Richard Loeser, MD, Michael R. Kosorok, PhD: University of North Carolina, Chapel Hill; [2]Daniel P. Beavers, PhD: Wake Forest School of Medicine, Winston-Salem, North Carolina; [3]Stephen Messier, PhD: Wake Forest University, Winston-Salem, North Carolina.

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

Address correspondence to Richard Loeser, MD, Division of Rheumatology, Allergy and Immunology, Thurston Arthritis Research Center, 3300 Thurston Building, CB#7280, University of North Carolina, Chapel Hill, NC 27599. Email: richard_loeser@med.unc.edu.

Submitted for publication May 7, 2019; accepted in revised form February 25, 2020.

SIGNIFICANCE & INNOVATIONS

- Diet and exercise or diet alone can benefit overweight or obese individuals with knee osteoarthritis (OA), although the response varies, suggesting that there may be subgroups who would achieve more benefit from a specific intervention.
- This study is the first to apply precision medicine–based machine learning approaches to clinical trial data in knee OA.
- These approaches identified subgroups of patients for whom a precision medicine decision rule would lead to improved outcomes over assignment of all individuals to the combined exercise and weight-loss intervention.

develop and internally validate the optimal precision medicine treatment regime using OA clinical trial data and simulations that would maximize expected clinical outcomes.

A precision medicine approach incorporates patient heterogeneity to inform clinical decisions (see Supplementary Table 1 in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract, for brief explanations of statistical terms and abbreviations used throughout this article). In many routine clinical settings, all patients with a given condition commonly receive the same treatment, despite the fact that treatment effectiveness differs by individual. Precision medicine is able to leverage the abundant patient information collected in the clinical setting (e.g., demographic and social economic characteristics, clinical history and physical examination findings, laboratory results, and in some cases even medical imaging and genetic traits) in the decision-making process about who should receive what treatment at what time. The precision medicine approach involves a function called a decision rule that maps individual characteristics to a recommended intervention. The decision rules are estimated by machine learning models, which have been recommended to aid clinical decision-making (6). Although many decision rules could potentially map patient information to a treatment, an optimal treatment rule (or optimal treatment regime) can be identified that maximizes the expected clinical outcomes of interest, thus serving to provide the optimal treatment recommendation to a patient population of interest (7).

We used the precision medicine approach to develop and internally validate the optimal exercise and weight-loss regimen for individuals with knee OA using data collected during the Intensive Diet and Exercise for Arthritis (IDEA) trial. The IDEA trial compared 3 randomized interventions over 18 months: 1) exercise (E) alone, considered the standard of care as a control group, 2) diet (D) with the goal of a 10% reduction in body weight, and 3) diet plus exercise (D + E), in overweight or obese adults with knee OA (3). IDEA results showed that, compared to exercise alone, participants randomized to the D and D + E groups had greater weight loss and greater reductions in interleukin-6 (IL-6). The other primary outcome, knee compressive force, was significantly reduced in the D group but not the D + E group. Self-reported pain and function scores improved more in the D + E group. Not unexpectedly, there was a variable response to each intervention among study participants, and those who lost more weight demonstrated more improvements in function, pain, knee compressive force, and IL-6 levels (3,5), independent of group assignment. We hypothesized that 1 or more of these variables could be used to determine an optimal treatment regime that would indicate which individuals would benefit the most (in terms of specific outcomes) from a given intervention when compared to assigning all individuals to just 1 of the 3 interventions.

MATERIALS AND METHODS

Participants. IDEA was an assessor-blinded, single-center randomized trial conducted during 2006–2011 at Wake Forest University and Wake Forest School of Medicine. Details of the study design and the results for the main outcomes have been previously published (3,8). Briefly, IDEA included 454 overweight and obese individuals (body mass index between 27 and 41) with mild or moderate symptomatic knee OA in 1 or both knees. Participants were ambulatory, sedentary, community-dwelling individuals ages ≥55 years with pain on most days due to knee OA. Measures (76 covariates) relevant to participant demographics, knee OA, and its effects on pain and function were collected at baseline, with selected outcome measures also obtained at 6 months (not used in this study) and 18 months.

Data preprocessing. The initial precision medicine analysis used 5 of the 7 clinical outcomes at 18 months that would be easiest for a clinician to obtain in a practice setting: weight loss since baseline, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain, function, and stiffness scores, and the Short Form 36 physical component score (PCS). Of the 454 participants who entered the trial, 399 completed the 18-month study. Because observed outcomes provide important information that drives the decision rule, we excluded participants missing 1 or more outcomes at 18 months, leaving 343 participants (Table 1). Dimension reduction was applied to control overfitting and extract the important features of the original 76 covariates at baseline, from which 15 covariates (Table 1) were chosen based on 3 criteria: 1) <15% missing data, 2) data that were clinically important and potentially measurable in clinical practice, and 3) statistically important data as determined by the variable importance measure from random forests (9). Selected covariates were then imputed via a nonparametric random forests method called missForest (10), which does not require assumptions about the data distribution, avoids cross validation, and can be applied to high-dimensional mixed-type data of unequal scales. Last, categorical variables were conformed and dichotomized, and all outcomes were transformed such that higher values represented

Table 1. Description of input data sets used in the analyses*

	Input data 1 (n = 343)	Input data 2 (n = 293)
Outcomes at 18 months	(n = 5) PCS, weight loss since baseline, WOMAC pain score, WOMAC function score, WOMAC stiffness score	(n = 7) PCS, weight loss since baseline, WOMAC pain score, WOMAC function score, WOMAC stiffness score, compressive force, plasma IL-6
Baseline covariates	(n = 15) ABC, BMI, walking distance, WOMAC function score, gait, heart attack, hip circumference, WOMAC pain score, PCS, average walking speed, WOMAC stiffness score, waist circumference, whole body lean DXA, whole body fat DXA, weight, randomization group	(n = 17) ABC, BMI, walking distance, WOMAC function score, gait, heart attack, hip circumference, IL-6, WOMAC pain score, PCS, average walking speed, WOMAC stiffness score, waist circumference, whole body lean DXA, whole body fat DXA, whole body percentage fat DXA, weight, randomization group

* ABC = activities-specific balance confidence scale; BMI = body mass index; DXA = dual-energy X-ray absorptiometry; IL = interleukin; PCS = physical component score; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

improvements in the outcomes. All baseline covariates were standardized to the standard normal distribution to avoid artifacts from differences in scaling due to the potential for varying scales to create misleading values of coefficients in models such as penalized regression. Missing data were investigated in the original IDEA study with multiple imputation analysis, which "revealed minimal differences from [the] original intention-to-treat analysis" (3). Further details on data cleaning, dimension reduction, and missing data and imputation are provided in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract.

A second analysis used all 7 outcomes, which included the 2 mechanistic outcomes (knee compressive force and plasma IL-6) at 18 months. This analysis was considered so as not to overlook any potentially valuable information from the 2 mechanistic outcomes, although such outcomes are not patient-reported or as easily obtainable in clinical practice as the other outcomes. We cleaned and imputed the second input data set (Table 1) and applied the same preprocessing procedure as described above. Values for IL-6 at 18 months were log-transformed in the analyses due to right-skewness and exponentiated back to original values during testing and optimal estimation.

Training process and performance. After the input data were cleaned and preprocessed, a total of 24 machine learning models were implemented (Table 2). They were selected specifically to suit the IDEA data, which represent a single-decision setting. The candidate models can be summarized in the following categories: penalized linear regression (models 1–4), ensemble learning of decision trees (models 5–7), tree-based dynamic treatment regime (models 8–20), support vector machine-based learning (models 21–23), and Bayesian model (model 24). Our selection of models covered both conventional and emerging concepts in the statistical literature; the rationale for each model choice is included in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract. In addition to the precision medicine models, we investigated 3 zero-order models (ZOMs), which assigned just q of the treatments (E, D, and D + E) to all participants (models 25–27). ZOMs are named after zero-order processes, which are fixed decision rules that do not change by individual.

Twenty-four machine-learning models and the 3 ZOMs, for a total of 27 models, provided estimated individualized treatment rules (ITRs), which were compared based on estimated value functions separately for each outcome. The value function is a scalar measure of performance for each ITR that evaluates the expectation of an outcome if future patients followed the estimated decision rule that is derived from training input data. A higher value function indicates a higher quality of the estimated ITR and more benefit to future patients in terms of that outcome. Hence, a learning model that maximizes the value estimate with small variation would be preferred. Mathematical definitions of the true and estimated value functions can be found in the subsection Value Function of the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract. The value estimates are usually derived from model evaluation techniques, such as cross validation. A simple cross validation procedure that splits the sample data into 1 training and 1 testing set would usually generate biased, ungeneralizable results. An alternative is *K*-fold cross validation, which refers to training ITRs on $K - 1$ of the randomly divided folds and testing the performance and generalization of ITRs on the 1 remaining fold. This process is repeated until every fold has been tested.

We used the jackknife method to estimate the bias and SE of the estimated value function used for model selection. The jackknife is a leave-one-out cross validation or n-fold cross validation method, where each individual serves as a fold, so the training sample leaves 1 observation out at a time (11). We chose the jackknife estimator because it requires weak assumptions (i.e., unrestricted shape of the probability distribution as long as the observations are independently and identically distributed) and is approximately unbiased for the true prediction error (12). In addition, stratified 10-fold cross validations were also performed to check the stability of jackknife value function estimates and to compare test results. Such validation methods (jackknife and cross validation) as well as simulation experiments accommodate for internal validation to prevent overfitting. More details on the jackknife and cross validation estimators as well

Table 2. Listing of precision medicine–based machine-learning models and zero-order models used in the analyses*

Model	Parameters	Model no.
Penalized regression (refs. 19,20)	Lasso, α = 1	1
	Ridge, α = 0	2
	Elastic net, α = 0.5	3
Kernel ridge regression (ref. 21)	Gaussian kernel	4
Random forests (ref. 9)	Rules based on each individual outcome	5
Number of trees = 500	Rules based on a weighted outcome of weight loss, pain, and function	6
Reinforcement learning trees (ref. 22)	Number of trees = 50	7
List-based dynamic treatment regime (ref. 21)	Embedded with Kernel ridge regression; no. of nodes = 2, 3, 5, 10	8, 9, 10, 11
	Embedded with random forests; no. of nodes = 2, 3, 5, 10	12, 13, 14, 15
	Embedded with super learning; no. of nodes = 2, 3, 5, 10	16, 17, 18, 19
	Embedded with elastic net; no. of nodes = 10	20
Residual weighted learning (ref. 23)	Linear kernel	21
	Polynomial kernel with 2nd order	22
	Polynomial kernel with 3rd order	23
Bayesian additive regression trees (ref. 24)	No. of trees = 500; no. of draws = 5,500 (including 500 burn-ins)	24
Zero-order model	Always assign to E	25
	Always assign to D	26
	Always assign to D + E	27

* ref. = reference.

as simulations on their theoretical properties may be found in the subsections The Jackknife and Stratified Cross Validation in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract.

Testing process and model selection. We applied all 27 candidate models to each outcome for training, recorded their estimated decision rules, and compared the jackknife value function estimators and their SEs in the testing process. For each outcome separately, the optimal precision medicine model (PMM) was

Table 3. Descriptive characteristics of baseline input data sets*

Characteristic	Input data 1 (n = 343)	Input data 2 (n = 293)	Overall (n = 399)
Randomization group			
Exercise	111 (32)	99 (34)	135 (34)
Diet	116 (34)	95 (32)	129 (32)
Diet and exercise	116 (34)	99 (34)	136 (34)
Age, mean ± SD years	65.6 ± 6.1	65.9 ± 6.2	65.9 ± 6.2
Weight, mean ± SD kg	92.1 ± 14.5	92.0 ± 14.8	92.4 ± 14.6
Body mass index, mean ± SD kg/m^2	33.4 ± 3.8	33.3 ± 3.8	33.5 ± 3.7
Female	251 (73)	211 (72)	291 (73)
Race			
African American	57 (17)	47 (16)	68 (17)
White	286 (83)	246 (84)	332 (83)
Education			
High school	100 (29)	84 (29)	117 (29)
College	164 (48)	142 (48)	194 (49)
Post college	77 (22)	65 (22)	87 (22)
Missing	2 (1)	2 (1)	2 (<1)
Smoking			
Never	196 (57)	169 (58)	229 (57)
Former	132 (38)	112 (38)	153 (38)
Current	10 (3)	8 (3)	12 (3)
Missing	5 (1)	4 (1)	6 (2)
Alcohol			
Never	66 (19)	60 (20)	77 (19)
Former	69 (20)	51 (17)	83 (21)
Current	199 (58)	174 (59)	229 (57)
Missing	9 (3)	8 (3)	11 (3)
Marital status			
Presently married or in a marriage-like relationship	239 (70)	208 (71)	276 (69)
Never married, divorced, separated, widowed	103 (30)	85 (29)	123 (31)
Missing	1 (<0.5)	–	1 (<0.5)

* Values are the number (%) unless indicated otherwise.

the model with the highest estimated value function with a smaller SE among the 24 machine learning models, i.e., its decision rule would bring the highest reward to future patients with small uncertainty in the value estimate. We found that, in general, SEs of the value estimators on the same outcome did not differ substantially across candidate models, so we focused on the value estimators. Similarly, the optimal ZOM is the model with the highest value estimate and relatively small SE from the 3 ZOMs. We performed a 2-sample Z test to compare the optimal PMM with the optimal ZOM (details in the subsection Model Selection in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract). After outcomes with statistically significant results were found, we estimated the decision rule of the optimal PMM trained on the entire data set (without jackknife validation), which served as the final data-driven, precision medicine–based treatment recommendation.

Multiple outcomes. To account for potential correlations among outcomes, we derived optimal treatment rules based on a weighted sum across multiple outcomes. A minimax algorithm was proposed to optimize data-driven weights for the 3 outcomes of greatest interest: weight loss since baseline, WOMAC pain subscore, and WOMAC function subscore at 18 months. To reduce computational time, we used a coarse-to-fine grid search with random forests models to determine the weight combination that maximized the lowest jackknife value function estimates among the 3 outcomes, hence the name "minimax" (details in the subsection Multiple Outcomes in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract). The selected minimax weights were then used to create a composite outcome, i.e., the weighted sum of weight loss, pain, and function scores, to train a random forests model (model 6) and estimate the optimal treatment rule. The random forest model contrasts with the other models discussed above where the precision medicine treatment rule was trained on a single outcome, while all models were tested on a single outcome.

All analyses were performed with R software, version 3.4.4 (13). Information on specific packages can be found in the subsection Choice of Models in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract. Because these analyses were exploratory in nature, the significance level was relaxed to 0.10. A complete outline of the entire precision medicine approach is shown in Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract.

RESULTS

Participant characteristics. Descriptive characteristics for the 2 input data sets, as well as the full data set with all 399 participants who finished 18 months of follow-up, are summarized in Table 3. In general, baseline characteristics of participants with

Table 4. Comparison between the optimal zero-order model (ZOM) and the optimal precision medicine model (PMM) for each outcome*

Data set and outcomes (18 months)	Optimal ZOM	Optimal PMM†	Estimated value (optimal PMM)	Relative increment‡	*P*§
Input data 1 (n = 343)					
Physical component score	D + E	Model 10	45.47	0.10	0.88
Weight loss since baseline	D + E¶	Model 10¶	11.21¶	1.45¶	0.01¶
WOMAC pain score	D + E	Model 15	3.25	0.02	0.38
WOMAC function score	D + E	Models 1 and 12	12.63	0.00	1.00
WOMAC stiffness score	D + E	Model 23	2.12	0.03	0.86
Input data 2 (n = 293)					
Compressive force	D	Model 9	2,336.21	21.74	0.73
IL-6 level	D	Model 12¶	2.29¶	0.26¶	0.09¶
Physical component score	D + E	Model 7	46.46	0.96	0.24
Weight loss since baseline	D + E	Model 7¶	11.76¶	1.31¶	0.06¶
WOMAC pain score	D + E	Model 9	3.24	0.08	0.59
WOMAC function score	D + E	Models 1 and 12–15	12.58	0.00	1.00
WOMAC stiffness	D + E	Model 8	2.08	0.04	0.31

* IL = interleukin; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† This table focuses on modeling on individual outcomes. The precision medicine model selection was among 23 models in Table 2, excluding model 6 (the results of which are shown in Table 5). Models 1–4 are penalized regression models. Models 5 and 7 are random forests and reinforcement learning trees. Models 8–11, models 12–15, and models 16–19 are, respectively, kernel ridge regression, random forests, and super learning list-based dynamic treatment regimes (DTRs) with 2, 3, 5, and 10 nodes. Model 20 is an elastic net list-based DTR with 10 nodes. Models 21–23 are residual weighted learning of different kernels, and model 24 is a Bayesian regression model.
‡ Relative increment is the jackknife estimated value function of the optimal PMM minus the jackknife estimated value function of the optimal ZOM, the increment in future expected outcome based on the optimal PMM relative to the optimal ZOM.
§ *P* value from the Z test (details in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract).
¶ Statistically significant.

available data were evenly distributed across the 3 intervention groups, as would be expected from a randomized clinical trial. There were no differences in selected baseline characteristics for participants with or without missing outcome data.

The optimal ZOM. Considering the 3 ZOMs, we found that the optimal ZOM model assigned every individual to D + E for all 5 clinical outcomes: weight loss since baseline, WOMAC pain, function, and stiffness scores, and PCS at 18 months (Table 4). Treatment D was the optimal ZOM for the 2 mechanistic outcomes: knee compressive force and plasma IL-6 level at 18 months.

The optimal PMM. The random forests model with minimax weights (model 6) was the optimal PMM for each of the 3 WOMAC subscores, regardless of input data (Table 5). For the rest of the outcomes (Table 4), list-based models (models 9–13) and reinforcement learning trees (model 7) were optimal among the 24 PMMs.

The optimal ZOM versus the optimal PMM. The relative increments between the estimated value functions of the optimal PMM and those of the optimal ZOM were positive (Table 4), indicating that the optimal PMM outperformed the optimal ZOM for all outcomes. According to the Z test, such improvement of the optimal PMMs compared to the optimal ZOMs was significant both for weight loss since baseline and for IL-6 levels (Table 4). We investigated these 2 outcomes further.

For weight loss between baseline and 18 months, the application of our precision medicine approach showed that future patients are estimated to lose 11.2 kg of weight on average between baseline and 18 months, according to the optimal PMM (list-based dynamic treatment regime with at most 5 nodes). This is an average of 1.4 kg more weight loss than if all patients had received D + E, the optimal ZOM (significant improvement, $P = 0.01$). Trained on input data 1 as a whole, the estimated optimal decision regime for weight loss would recommend intervention D + E to patients who meet either of the following 2 conditions: 1) if, at baseline, weight does not exceed 109.35 kg and waist circumference is above 90.25 cm, or 2) if, at baseline, weight is >109.35 kg or waist circumference does not exceed 90.25 cm, and they have reported a prior heart attack. If neither of these conditions are met, the recommendation is treatment D. The visualization of this optimal rule can be found in Figure 1A.

For IL-6, the application of our precision medicine approach showed that future patients are estimated to decrease IL-6 to 2.29 pg/ml on average at 18 months, according to the optimal PMM (list-based dynamic treatment regime with at most 2 nodes). This is an average of 0.26 pg/ml more reduction than if all patients had received D, the optimal ZOM (significant improvement, $P = 0.09$). Trained on input data 2, the estimated optimal treatment rule for IL-6 assigned D + E to patients who meet the following condition: if, at baseline, IL-6 does not exceed 4.5 pg/ml and WOMAC function score is >12.5. If this condition is not met, patients would be assigned to treatment D (Figure 1B). As evidence of stability, we found similar patterns and similar conclusions for weight loss and IL-6 using the stratified 10-fold cross validation (see the subsection Stratified Cross Validation in the Supplementary Materials,

Table 5. Comparison between the optimal zero-order model (ZOM) and the random forest model for weighted sum of selected outcomes (model 6)*

Data set and outcomes (18 months)	Optimal ZOM	Optimal PMM	Estimated value (optimal PMM)	Relative increment†	*P*‡
Input data 1 (n = 343)					
Physical component score	D + E	Model 6	45.43	0.05	0.86
Weight loss since baseline	D + E§	Model 6§	10.10§	0.34§	0.05§
WOMAC pain score	D + E	Model 6	3.24	0.03	0.71
WOMAC function score	D + E	Model 6	12.42	0.21	0.54
WOMAC stiffness score	D + E	Model 6	2.11	0.03	0.44
Input data 2 (n = 293)					
Compressive force	D	Model 6	2,446.46	−88.50	0.41
IL-6 level	D	Model 6	2.55	0.01	0.98
Physical component score	D + E	Model 6	45.70	0.20	0.58
Weight loss since baseline	D + E	Model 6	10.76	0.32	0.29
WOMAC pain score	D + E	Model 6	3.23	0.10	0.41
WOMAC function score	D + E	Model 6	12.26	0.32	0.47
WOMAC stiffness	D + E	Model 6	2.03	0.09	0.13

* IL = interleukin; PMM = precision medicine model; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† Relative increment is the jackknife estimated value function of the model 6 optimal PMM minus the jackknife estimated value function of the D + E, the increment in future expected outcome based on model 6 relative to the optimal ZOM.
‡ *P* value from the Z test (details in the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract).
§ Statistically significant.

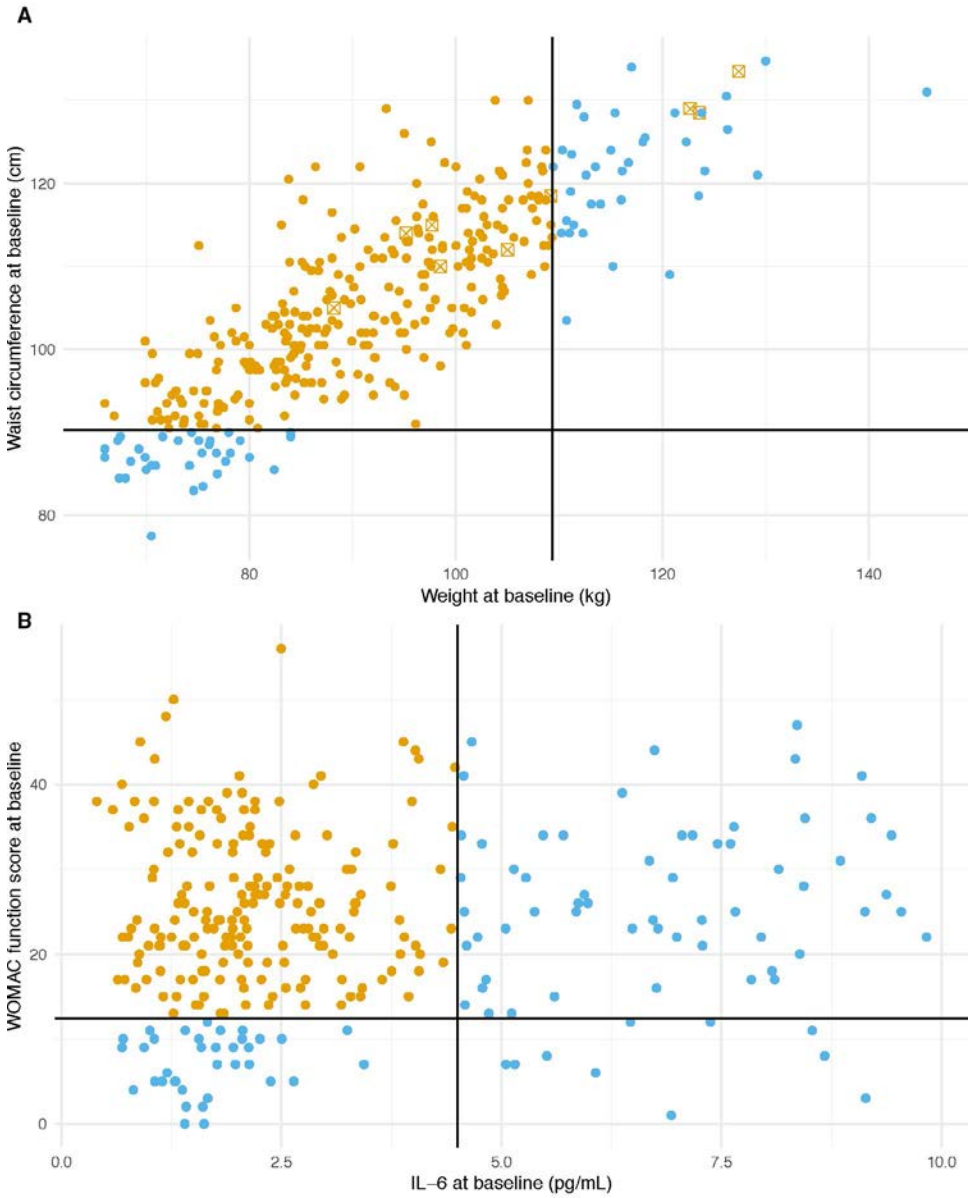


Figure 1. Visualization of the estimated optimal decision regimes for outcomes: **A**, weight loss since baseline; **B**, interleukin-6 (IL-6) at 18 months. Scatter plots of data for each individual are color-coded to indicate the optimal treatment group assignment of all individuals in the input data (input data 1 for outcome weight loss since baseline, and input data 2 for outcome IL-6 at 18 months). Blue indicates individuals who would be assigned to diet only (D) and orange to those assigned to diet plus exercise (D + E). For weight loss since baseline, previous heart attack (yes or no) also determined the group assignment and is shown as a checked box for those individuals who met that criterion. The horizontal and vertical reference lines indicate the cutoff levels for the variables shown on the horizontal and vertical axis, which determined group assignment. WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract).

Multiple outcomes results. The outcomes were positively correlated to some extent. The highest correlations were found among WOMAC scores (pain, function, stiffness) and PCS values (Pearson's correlation coefficients between 0.52 and 0.87). For input data 1, the minimax rule selected 0.1, 0.6, and 0.3 as data-driven weights for the 3 selected outcomes (weight loss, pain, and function, respectively) and 0, 0.32, and 0.68 for input data 2, respectively. We did not scale the outcomes but allowed the weights to adjust for different scales in the outcomes. Similar to the Z test comparison in the previous subsection (Table 4), we compared the optimal ZOM with 1 PMM: the random forests model trained on the weighted composite outcome (model 6) (Table 5). There was evidence of significant improvement of model 6 relative to the optimal ZOM (D + E) for weight loss since baseline for input data 1 ($P = 0.05$). Although not statistically significant, the

remaining outcomes (except compressive force) also expressed positive relative improvement in both input data sets. In particular, model 6 outperformed other PMMs in terms of the estimated value function for the 3 WOMAC scores, but not for outcomes uncorrelated to the 3 weighed outcomes, which are compressive force and IL-6.

DISCUSSION

In this study, we investigated optimal treatment recommendations for older and overweight or obese individuals with knee OA using precision medicine and machine learning tools applied to data obtained from the IDEA trial. The individual treatment decisions obtained from our precision medicine approach are data-driven (requiring no strong assumptions), reproducible (with careful reporting of the analysis process) (7), and generalizable and extendable to other clinical settings (because of rich heterogeneity in the clinical input data).

The results of the optimal ZOM, where everyone would be assigned to a single intervention, match with those from the published IDEA trial (3,8). The assignment of patients to the D + E intervention would be expected to result in the optimal improvement in the majority of patients in the clinical outcomes of weight loss since baseline, WOMAC pain, function, and stiffness scores, and PCS, and so should remain the recommendation of choice. In individuals where the primary goal is to reduce systemic inflammation as measured by plasma IL-6 levels and/or reduce the knee compressive force, then D alone would be the treatment of choice.

The treatment rules of the optimal PMMs suggested that not everyone benefits from D + E, even though patients are expected to be assigned to this group based on the ZOM. Further improvements in weight loss could be obtained in certain patients selected by measures of high baseline weight (>109.35 kg) or low waist circumference (≤90.25 cm) accompanied by lack of a previous heart attack that would result in assigning them to D rather than D + E. This treatment would only be a consideration if weight loss alone was more important to the patient than the level of improvement in pain and function. We can only speculate why people of higher weight or relatively lower waist circumference and no history of heart attack would benefit more from D than D + E. First, following the suggested exercise program may likely be more difficult for patients with a higher weight. Second, higher weight with lower waist circumference could be seen in individuals who have more peripheral adiposity rather than central adiposity. In these cases, D could be more effective in losing weight. The finding that our results were modified by a history of a heart attack may be that the cardiac status of these individuals encourages optimal compliance and improves more with the combined D + E than D alone, which allows for greater activity levels, resulting in greater weight loss.

The finding that the IL-6 outcome improves more with D than D + E in certain individuals is not easily explained. We noted that individuals with high baseline IL-6 levels (i.e., >4.5 pg/ml) or those with low baseline function scores (≤12.5; range 0–68) reduced their IL-6 more from diet only. Individuals whose IL-6 is not high but who have poorer function are recommended to receive both diet and exercise. A decrease in IL-6 suggests less systemic inflammation, but there is no solid evidence to suggest that exercise would modulate the reduction in IL-6 that occurs with dietary weight loss. Because all 3 groups received an intervention, the significant differences in outcomes noted among the groups at 18 months would be unlikely to be due to regression to the mean. Our findings that specific subgroups of individuals received more benefit from specific interventions argues against the premise that the response was simply due to patient perception rather than to the intervention itself.

As for the multiple outcomes, comparison between Tables 4 and 5 suggested that our minimax rule together with the coarse-to-fine grid search for parameter optimization can be a useful way to incorporate multiple outcomes, and combining correlated outcomes has the potential for bringing more benefits to patients than single outcomes. However, uncorrelated outcomes do not benefit from the composite outcome.

Potential limitations of this study include the following. First, we were not able to use the information of ~100 of the trial participants due to missing outcome data. Although a larger sample size should lead to higher power, our 2 input data sets remain representative of the overall data, as Table 3 shows. Second, the analyses did not include intermediate follow-up data at 6 months. Although longitudinal analysis methods could be applied to the IDEA data, we were more interested in the final improvements of each outcome between the start and the end of the trial and less on the intermediate progress. In addition, an additional time point shortly after the trial started would not be likely to be influential because we expect that the interventions take time to have an effect.

Third, there were some covariates with a large proportion of missing data excluded from this analysis. The majority of these were measures that would not be routinely collected in the clinical setting, such as full-length lower extremity radiographs for alignment, computed tomography for abdominal and thigh fat, knee magnetic resonance imaging, and isokinetic strength testing. Finally, our results are from a single clinical trial of patients with mild-to-moderate symptomatic knee OA (3) and may not be generalizable to populations with more severe knee OA.

We expect that the following future studies would be useful: 1) Exploring more robust models that directly determine the optimal treatment rules. We have observed from the results that, in general, machine learning models that predict well do not necessarily find the optimal treatment rule. The objectives are different: better prediction aims to lower mean squared errors, whereas optimal treatment rules aim to increase value functions.

For example, there have been recent advances in super learning that directly learn the optimal treatment regime (14–16), and we believe that further such robust models are worth investigating. 2) Finding optimal treatment regimes in the setting of multiple decision time points, where data can vary regularly by time and treatment plans, regimes that can be adjusted periodically rather than fixed for the entire intervention period. This process would make the individualized treatment recommendations more up-to-date and adaptive. There are many dynamic models that can be applied to this setting, and we recommend reinforcement learning (17) and Gaussian processes (18). 3) Although external validation was investigated via simulations (see the Supplementary Materials, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24179/abstract), a new randomized trial on a similar population would be needed for external validation of our findings.

ACKNOWLEDGMENTS

The authors thank all the members of the IDEA study team for their contributions to the original 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. Loeser 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. Jiang, Callahan, Loeser, Kosorok.
Acquisition of data. Beavers, Messier, Loeser.
Analysis and interpretation of data. Jiang, Nelson, Cleveland, Schwartz, Arbeeva, Alvarez, Callahan, Loeser, Kosorok.

REFERENCES

1. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014;73:1323–30.
2. Vina ER, Kwoh CK. Epidemiology of osteoarthritis: literature update. Curr Opin Rheumatol 2018;30:160–7.
3. Messier SP, Mihalko SL, Legault C, Miller GD, Nicklas BJ, DeVita P, et al. Effects of intensive diet and exercise on knee joint loads, inflammation, and clinical outcomes among overweight and obese adults with knee osteoarthritis: the IDEA randomized clinical trial. JAMA 2013;310:1263–73.
4. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: the Chronic Osteoarthritis Management Initiative of the U.S. Bone and Joint Initiative. Semin Arthritis Rheum 2014;43:701–12.
5. Messier SP, Resnik AE, Beavers DP, Mihalko SL, Miller GD, Nicklas BJ, et al. Intentional weight loss in overweight and obese patients with knee osteoarthritis: is more better? Arthritis Care Res (Hoboken) 2018;70:1569–75.
6. Jamshidi A, Pelletier JP, Martel-Pelletier J. Machine-learning-based patient-specific prediction models for knee osteoarthritis. Nat Rev Rheumatol 2019;15:49–60.
7. Kosorok MR, Laber EB. Precision medicine. Annu Rev Stat Appl 2019;6:263–86.
8. Messier SP, Legault C, Mihalko S, Miller GD, Loeser RF, DeVita P, et al. The Intensive Diet and Exercise for Arthritis (IDEA) trial: design and rationale. BMC Musculoskelet Disord 2009;10:93.
9. Breiman L. Random forests. Machine Learning 2001;45:5–32.
10. Stekhoven DJ, Bühlmann P. MissForest: non-parametric missing value imputation for mixed-type data. Bioinformatics 2012;28:112–8.
11. Efron B, Tibshirani RJ. An introduction to the bootstrap. Boca Raton (LA): Chapman and Hall/CRC; 1994.
12. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. 2nd ed. New York: Springer New York; 2009.
13. R Foundation. The R project for statistical computing. URL: https://www.r-project.org/.
14. Luedtke AR, van der Laan MJ. Super-learning of an optimal dynamic treatment rule. Int J Biostat 2016;12:305–32.
15. Van der Laan MJ, Luedtke AR. Targeted learning of the mean outcome under an optimal dynamic treatment rule. J Causal Inference 2015;3:61–95.
16. Van der Laan MJ, Luedtke AR. Targeted learning of an optimal dynamic treatment, and statistical inference for its mean outcome. U.C. Berkeley Division of Biostatistics Working Paper Series, 2014, Paper 329. University of California, Berkeley.
17. Schulte PJ, Tsiatis AA, Laber EB, Davidian M. Q- and A-learning methods for estimating optimal dynamic treatment regimes. Stat Sci 2014;29:640–61.
18. Wilson AG, Adams RP. Gaussian process kernels for pattern discovery and extrapolation. J Proceedings of the 30th International Conference on Machine Learning, PMLR 2013; 28:1067–75.
19. Tibshirani R. Regression shrinkage and selection via the lasso. J Royal Stat Soc B 1996;58:267–88.
20. Zou H, Hastie T. Regularization and variable selection via the elastic net. J Royal Stat Soc B 2005; 67:301–20.
21. Zhang Y, Laber EB, Davidian M, Tsiatis AA. Estimation of optimal treatment regimes using lists. J Am Stat Assoc 2018.113:1541–9.
22. Zhu R, Zeng D, Kosorok MR. Reinforcement learning trees. J Am Stat Assoc 2015;110:1770–84.
23. Zhou X, Mayer-Hamblett N, Khan U, Kosorok MR. Residual weighted learning for estimating individualized treatment rules. J Am Stat Assoc 2017;112:169–87.
24. Chipman HA, George EI, McCulloch RE. BART: Bayesian additive regression trees. Ann Appl Stat 2010;4:266–98.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 702–706
DOI 10.1002/acr.24147

Adherence to Antimalarial Therapy and Risk of Type 2 Diabetes Mellitus Among Patients With Systemic Lupus Erythematosus: A Population-Based Study

Shahrzad Salmasi,[1] Eric C. Sayre,[2] J. Antonio Aviña-Zubieta,[3] John M. Esdaile,[3] and Mary A. De Vera[1]

Objective. To evaluate the association between adherence to antimalarials and type 2 diabetes mellitus (DM) in patients with systemic lupus erythematosus (SLE).

Methods. Using administrative health databases in British Columbia, Canada, we conducted a retrospective, longitudinal cohort study of patients with incident SLE and incident antimalarial use. We established antimalarial drug courses by defining a new course when a 90-day gap is exceeded between refills and we calculated proportion of days covered (PDC) for each course. We categorized medication taking as: 1) adherent (PDC ≥0.90), 2) nonadherent (0 < PDC < 0.90), and 3) discontinuer (no drug). Type 2 DM outcomes were based on outpatient or inpatient visits, or antidiabetic medication use. We used multivariable Cox proportional hazards models with time-dependent variables.

Results. Over a median of 4.62 years of follow-up in our incident cohort of 1,498 patients with SLE (90.8% women), we recorded 140 incident cases of type 2 DM. Multivariable hazard ratios were 0.61 (95% confidence interval [95% CI] 0.40–0.93) for adherent and 0.78 (95% CI 0.50–1.22) for nonadherent, respectively, as compared to discontinuers.

Conclusion. Our findings of a protective effect of adherence to antimalarials in preventing type 2 DM provides further support for the importance of adherence to antimalarials to obtain the benefits of therapy.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic systemic autoimmune disease affecting most body systems. Antimalarials (e.g., hydroxychloroquine and chloroquine) are the cornerstone of SLE management with proven effects on improving survival, reducing disease activity and flares, and reducing the risks of irreversible organ damage, venous thromboembolism, and dyslipidemia (1). In addition, antimalarials have been shown to have beneficial effects, including prevention of type 2 diabetes mellitus (DM), which is both a complication of SLE and a side effect of medications (e.g., glucocorticoids) used to manage SLE (2). Indeed, a 2015 cohort study of 8,628 incident SLE cases reported 74% lower risk of incident type 2 DM associated with hydroxychloroquine use (hazard ratio [HR] 0.26, 95% confidence interval [95% CI] 0.18–0.37) (2). The protective effect of antimalarials is, however, cumulative, highlighting the importance of adherence, which is low in SLE. In a 2017 systematic review, we reported antimalarial adherence rates as low as 25% among patients (3). In this study we aimed to evaluate the association between antimalarial adherence and incidence of type 2 DM among patients with SLE.

MATERIALS AND METHODS

Data sources. We used Population Data British Columbia (4), which has captured data on outpatient visits (5), hospitalizations (6), demographic information (7), and vital statistics (8), since 1990 on the entire population of British Columbia (~5 million residents). We linked these to PharmaNet, which contains complete information on all drug prescriptions dispensed outside of the

Supported by the Canadian Initiative for Outcomes in Rheumatology Care (operating grant) and the Canadian Institutes of Health Research (team grant PRECISION: Preventing Complications from Inflammatory Skin, Joint and Bowel Conditions [FRN#THC-135235]). Dr. De Vera holds a Canada Research Chair in Medication Adherence, Utilization, and Outcomes and is a Michael Smith Foundation for Health Research Scholar. Dr. Aviña-Zubieta is a Michael Smith Foundation for Health Research Scholar and the British Columbia Lupus Society Research Scholar.

[1]Shahrzad Salmasi, B.Pharmacy (Hons), MSc, Mary A. De Vera, PhD: University of British Columbia, Vancouver, British Columbia, Canada, Arthritis Research Canada, Richmond, British Columbia, Canada, and Collaboration for Outcomes Research and Evaluation, Vancouver, British Columbia, Canada; [2]Eric C. Sayre, PhD: Arthritis Research Canada, Richmond, British Columbia, Canada; [3]J. Antonio Aviña-Zubieta, MD, PhD, John M. Esdaile, MD, MPH: Arthritis Research Canada, Richmond, British Columbia, Canada, and University of British Columbia, Vancouver, British Columbia.

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

Address correspondence to Mary A. De Vera, PhD, University of British Columbia Faculty of Pharmaceutical Sciences, 2405 Wesbrook Mall, Vancouver, British Columbia, Canada, V6T 1Z3. Email: mdevera@mail.ubc.ca.

Submitted for publication August 27, 2019; accepted in revised form January 14, 2020.

SIGNIFICANCE & INNOVATIONS

- Patients with incident systemic lupus erythematosus (SLE) who followed their prescribed antimalarial had a 39% lower risk of developing type 2 diabetes mellitus (DM), compared to those who discontinued therapy.
- Our findings also suggest that when taking less than 90% of the prescribed doses of antimalarials, the protective effect against type 2 DM is lost.
- Our study provides support for the importance of adherence to antimalarials in SLE by demonstrating protective impacts on type 2 DM, a serious complication in SLE.

hospital in British Columbia since 1996 (e.g., drug identification number, dispensing date, quantity, and duration) (9).

Study design and cohort definition. We conducted a retrospective, longitudinal cohort study of adults with incident SLE and incident antimalarial use between January 1996 and December 2012. To define incident SLE, we used a 7-year run-in period and applied a case definition of the International Classification of Diseases, Ninth Revision (ICD-9) code 710.0 at least 2 months apart and within a 2-year period by a nonrheumatologist, or 1 ICD-9 code by a rheumatologist, or 1 ICD-10 code (M32.1, M32.8, or M32.9) from hospitalization data. The second date in the pair of codes or date of rheumatologist visit/hospitalization was considered the index SLE date. This SLE definition has been shown to have a sensitivity rate of 85%, specificity of 73%, and positive predictive value of 90% (10). As with prior studies, we applied further exclusions to improve specificity. We defined the first antimalarial prescription in PharmaNet after the SLE diagnosis as the incident prescription and considered both hydroxychloroquine and chloroquine. Since it is possible to start therapy as an SLE diagnosis is being confirmed, we permitted any antimalarial prescriptions that occurred within a 90-day period prior to index SLE date.

Exposure and outcome assessment. To define our study exposure of antimalarial adherence, we used PharmaNet data on prescription dates and days' supply and established antimalarial drug courses and gaps. A "course" comprised subsequent prescriptions allowing overlaps and permissible gaps of up to 90 days (length of 1 prescription cycle in British Columbia) between each refill. For each course, we calculated proportion days covered (PDC) minus the total number of days with antimalarials divided by the length of course (11) and categorized medication taken during each as adherent (PDC ≥ 0.90), nonadherent (0 < PDC < 0.90), and discontinuer (PDC = 0 during gaps). This approach of operationalizing antimalarial adherence exposure captures the dynamic nature of medication taking and accounts for both components of implementation (or execution) of the dosing regimen and persistence with therapy (Figure 1). Our study outcome of interest was incident type 2 DM during the follow-up period, which was defined using outpatient or hospitalizations (ICD 9/10 codes 250.X, E11.X, E14.X) or antidiabetic medication use.

Statistical analysis. We used multivariable Cox proportional hazards models to estimate the association between adherence to antimalarial therapy and time to incident type 2 DM. We modeled antimalarial adherence as a categorical

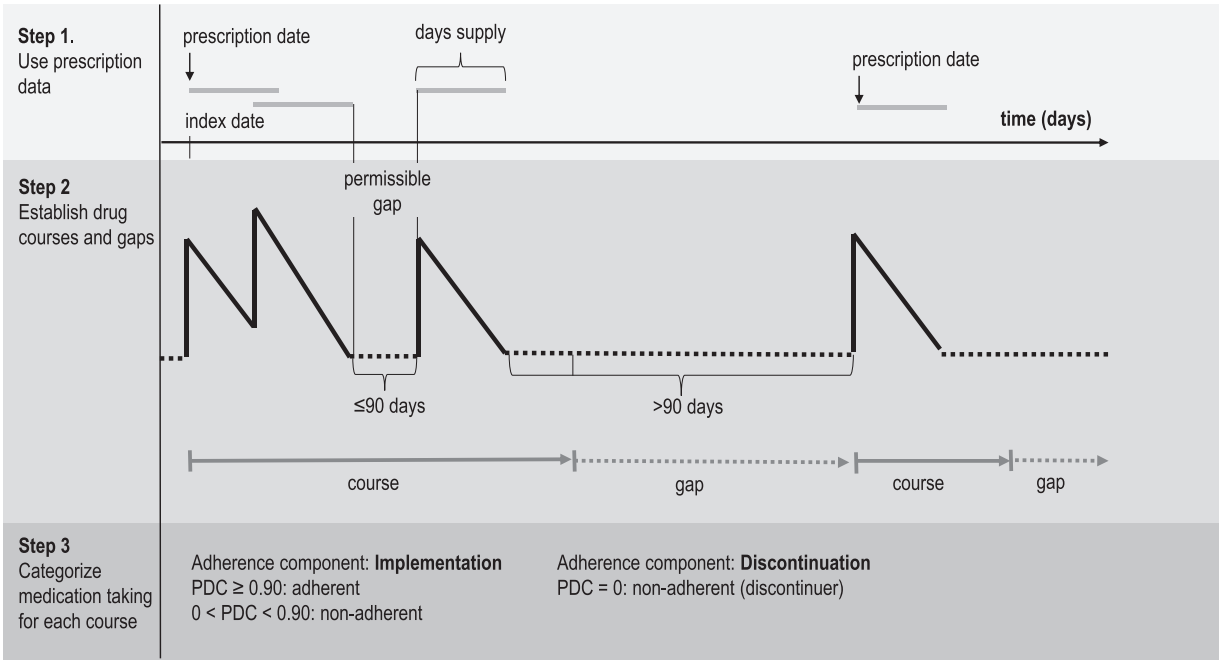


Figure 1. Schema of operationalization of antimalarial adherence. PDC = proportion of days covered.

time-dependent covariate over follow-up, with the HR representing the risk of type 2 DM associated with antimalarial adherence in the current drug course (as compared to discontinuers). Covariates considered in multivariable models included age, sex, socioeconomic status (using a proxy measure based on neighborhood income quintile), and residence (rural versus urban, as determined by using Census Metropolitan Area/Census Agglomeration from geographic census data). We also considered fixed-in-time binary variables measured in the year before the index date, such as comorbid conditions (e.g., depression, chronic kidney disease, hypertension) and modified Charlson Comorbidity Index, as well as time-dependent covariates, such as use of other medications (e.g., other disease-modifying antirheumatic drugs [DMARDs] and glucocorticoids) and health care utilization (e.g., hospitalizations and visits to specialists, including a rheumatologist, nephrologist, dermatologist, and psychiatrist). Polypharmacy (defined as being treated with at least 2 different medications in the past year out of 11 DMARDS, excluding antimalarials, 11 cardiovascular drugs, hormone replacement therapy, and oral contraceptives) was included both at baseline and as a time-dependent covariate. We conducted sensitivity analyses, which involved varying the permissible gap (e.g., 120 days, 180 days) as well as the adherence cutoff (e.g., PDC ≥ 0.70 and ≥ 0.80). All analyses were conducted using SAS software, version 9.4.

Study conduct. We obtained ethics approval from the University of British Columbia. All data were deidentified and no personal information was available at any point of the study. All inferences, opinions, and conclusions drawn in this manuscript are those of the authors, and do not reflect the opinions or policies of the data stewards.

RESULTS

Our study cohort comprised 1,498 patients with SLE (90.8% female) with a mean ± SD age of 44.4 ± 14.8 years (Table 1). The mean ± SD number of antimalarial prescriptions and drug courses over the follow-up period were 23.2 ± 37.7 and 2.1 ± 1.8, respectively, with a mean ± SD drug course duration of 553.9 ± 820.8 days. Over the drug course of a median of 4.6 years of follow-up, we recorded 140 incident cases of type 2 DM.

After adjusting for age, sex, comorbidities, and concomitant medications, the adjusted HR for developing type 2 DM among those who were adherent to antimalarials was 0.61 (95% CI 0.40–0.93) as compared to discontinuers. In contrast, the adjusted HR for those who were nonadherent was 0.78 (95% CI 0.50–1.22) as compared to discontinuers (Table 2). Sensitivity analyses involving permutations of permissible gaps (Table 2) and PDC cutoff (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24147/abstract) did not materially change our results.

Table 1. Characteristics of the incident SLE cohort with incident antimalarial use (n = 1,498)*

Characteristic	Values
Demographic information	
Age, mean ± SD years	44.4 ± 14.8
Women	1,360 (90.8)
Socioeconomic status	
High (4th and 5th quintile)	573 (38.3)
Middle (3rd quintile)	335 (22.4)
Low (1st and 2nd quintile)	590 (39.4)
Rural residence	183 (12.2)
Comorbid conditions†	
Depression	353 (23.6)
Chronic kidney disease	313 (20.9)
Hypertension	226 (15.1)
Chronic obstructive pulmonary disease	155 (10.4)
Angina	64 (4.3)
Charlson Comorbidity Index score, mean ± SD	1.3 ± 0.9
Medication use	
Glucocorticoids†	664 (44.3)
Traditional NSAIDs†	612 (40.9)
Other DMARDs†	297 (19.8)
COX-2 inhibitors†	191 (12.8)
Polypharmacy‡	206 (13.8)
Pharmacy loyalty§	275 (18.4)
Health care utilization	
Hospitalized†	419 (28.0)
Outpatinet visits, mean ± SD§	20.3 ± 14.9
No. (range) of rheumatologist visits§	1.90 (0–41)
No. (range) of nephrologist visits§	0.15 (0–62)

* Values are the number (%) unless indicated otherwise. COX-2 = cyclooxygenase 2; DMARDs = disease-modifying antirheumatic drugs; NSAIDs = nonsteroidal antiinflammatory drugs; SLE = systemic lupus erythematosus.
† Evaluated over the year preceding index date.
‡ Evaluated over follow-up period; filling of >75% of prescriptions by the patient in the same pharmacy.
§ Evaluated over follow-up.

DISCUSSION

Using population-based administrative health data with complete information on dispensed prescriptions, we conducted a longitudinal cohort study to evaluate the association between adherence to antimalarials and risk of developing type 2 DM in a Canadian cohort with a new diagnosis of SLE. Adherent patients were 39% less likely to develop type 2 DM compared to those who discontinued their therapy. Our findings also suggest that when taking <90% of the prescribed doses of antimalarials, the protective effect against type 2 DM is lost. Altogether, our study provides further support for the importance of adherence to antimalarials in SLE by demonstrating protective impacts on type 2 DM, a serious complication in SLE.

While antimalarials have been shown to be protective against type 2 DM in SLE patients, their effectiveness is dose dependent

Table 2. Univariable and multivariable models on the association between antimalarial adherence and risk of type 2 diabetes mellitus*

	Model 1 (90-day permissible gap)	Model 2 (120-day permissible gap)	Model 3 (180-day permissible gap)
Antimalarial adherence (unadjusted HR [95% CI])			
Adherent (PDC ≥0.9) (vs. discontinuer PDC = 0)	0.70 (0.47–1.05)	0.63 (0.40–0.98)	0.64 (0.39–1.03)
Nonadherent (0 < PDC < 0.9)	0.75 (0.49–1.16)	0.71 (0.46–1.09)	0.67 (0.44–1.04)
Antimalarial adherence			
Adherent (PDC ≥0.9) (vs. discontinuer PDC = 0)	0.61 (0.40–0.93)	0.54 (0.34–0.86)	0.56 (0.34–0.92)
Nonadherent (0 < PDC < 0.9)	0.78 (0.50–1.22)	0.74 (0.47–1.15)	0.68 (0.44–1.06)
Demographic characteristics			
Age	1.01 (1.00–1.03)	1.02 (1.00–1.03)	1.02 (1.00–1.03)
Sex (women vs. men)	0.68 (0.40–1.14)	0.79 (0.45–1.39)	0.82 (0.45–1.47)
Residence (rural vs. urban)	0.71 (0.40–1.28)	0.74 (0.40–1.36)	0.75 (0.39–1.41)
Socioeconomic status			
Quintile 1 (versus quintile 3)	0.99 (0.62–1.60)	1.01 (0.62–1.66)	1.07 (0.64–1.79)
Quintile 2 (versus quintile 3)	0.78 (0.47–1.30)	0.82 (0.49–1.39)	0.92 (0.54–1.57)
Quintile 4 (versus quintile 3)	0.56 (0.32–0.99)	0.59 (0.33–1.06)	0.60 (0.33–1.09)
Quintile 5 (versus quintile 3)	0.59 (0.34–1.02)	0.56 (0.31–1.01)	0.54 (0.29–1.00)
Comorbid conditions			
Depression	0.80 (0.52–1.23)	0.76 (0.48–1.20)	0.76 (0.48–1.21)
Chronic kidney disease	0.88 (0.53–1.44)	0.84 (0.49–1.41)	0.86 (0.50–1.48)
Hypertension	1.11 (0.66–1.86)	0.84 (0.48–1.48)	0.88 (0.48–1.58)
Chronic obstructive pulmonary disease	1.04 (0.61–1.78)	1.15 (0.67–1.98)	1.23 (0.71–2.13)
Angina	0.67 (0.31–1.46)	0.86 (0.39–1.89)	0.85 (0.37–1.96)
Charlson Comorbidity Index score	1.10 (0.90–1.36)	1.12 (0.91–1.37)	1.12 (0.91–1.38)
Medications			
Glucocorticoids†	1.27 (0.88–1.83)	1.33 (0.90–1.95)	1.23 (0.82–1.83)
Traditional NSAIDs†	1.29 (0.91–1.85)	1.39 (0.96–2.00)	1.49 (1.02–2.17)
COX-2 inhibitors†	1.38 (0.82–2.32)	1.53 (0.91–2.59)	1.51 (0.88–2.59)
Other DMARDs†	0.88 (0.53–1.44)	0.76 (0.44–1.31)	1.03 (0.61–1.75)
Polypharmacy‡	1.19 (0.68–2.06)	1.09 (0.60–1.97)	1.12 (0.61–2.03)
Polypharmacy§	1.19 (0.77–1.85)	1.18 (0.74–1.89)	1.09 (0.68–1.77)
Health care utilization			
No. of psychiatrist visits†	1.01 (0.98–1.04)	1.01 (0.98–1.04)	1.00 (0.98–1.03)
No. of rheumatologist visits†	1.02 (0.95–1.10)	1.02 (0.94–1.10)	1.01 (0.93–1.10)
No. of dermatologist visits†	0.78 (0.58–1.05)	0.77 (0.57–1.06)	0.78 (0.58–1.06)
No. of nephrologist visits†	1.00 (0.98–1.03)	1.00 (0.98–1.03)	1.00 (0.98–1.03)

* Values are the adjusted hazard ratio (HR) (95% confidence interval [95% CI]) unless indicated otherwise. COX-2 = cyclooxygenase 2; DMARDs = disease-modifying antirheumatic drugs; NSAIDs = nonsteroidal antiinflammatory drugs; PDC = proportion of days covered.
† Time-dependent variable evaluated over follow-up period.
‡ Filling of >75% of prescriptions by the patient in the same pharmacy.
§ Time-dependent variable evaluated over follow-up period; filling of >75% of prescriptions by the patient in the same pharmacy.

(2). Patients on a higher cumulative dose (≥129 gm) have been shown to have a significantly lower hazard of developing type 2 DM compared to patients with a lower cumulative dose; this demonstrates that therapy effectiveness is critically dependent on patients' medication-taking behavior (2). Our study provides support for this relationship as the adjusted HR for adherent SLE patients, defined as those taking 90% of prescribed doses, was 0.61 (95% CI 0.40–0.93) as compared to discontinuers. In contrast, the adjusted HR for nonadherent SLE patients of 0.78 (95% CI 0.50–1.22) suggests that taking any fewer than 90% of the prescribed doses of antimalarials compromises their protective effect against type 2 DM.

Indeed, prevention of type 2 DM is of paramount importance in SLE. Type 2 DM can lead to many other complications such as neuropathy, cardiovascular disease, and renal insufficiency, further complicating management of the patient (12). Our findings are concerning, given that currently SLE patients' adherence to antimalarials is as low as 25% and that there are very few adherence interventions designed for these patients (3). For example, in our 2015 systematic review on adherence intervention studies in rheumatic diseases, of the 23 identified studies, only 3 were among patients with SLE (13). Given the direct clinical implications of nonadherence in SLE, future research should focus on development and evaluation of interventions to improve antimalarial adherence.

Strengths of the study include use of population data, including all medications dispensed (public or private payee) to the entire SLE population in British Columbia (high external validity). Furthermore, availability of complete prescription data allowed us to apply an approach of defining antimalarial adherence that captures implementation (or execution) of the dosing regimen as well as persistence with therapy. Administrative data, however, are vulnerable to diagnostic uncertainty and hence misclassification. We addressed this by using the strictest published case definition for SLE with additional exclusions, as described in the methods. Administrative data are also limited to prescriptions dispensed and lack information on the actual consumption of medications by patients (14).

This population-based study highlighted that taking <90% of the prescribed antimalarials compromises their effect in preventing type 2 DM in patients with SLE. Our findings should be used to emphasize the importance of medication adherence in not only treating SLE, but also preventing its complications.

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. De Vera 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. Sayre, De Vera.
Acquisition of data. Aviña-Zubieta, De Vera.
Analysis and interpretation of data. Salmasi, Sayre, Aviña-Zubieta, Esdaile, De Vera.

REFERENCES

1. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. BMJ 2019;78:736–45.
2. Chen YM, Lin CH, Lan TH, Chen HH, Chang SN, Chen YH, et al. Hydroxychloroquine reduces risk of incident diabetes mellitus in lupus patients in a dose-dependent manner: a population-based cohort study. Rheumatology (Oxford) 2015;54:1244–9.
3. Mehat P, Atiquzzaman M, Esdaile JM, Avina-Zubieta A, De Vera MA. Medication nonadherence in systemic lupus erythematosus: a systematic review. Arthritis Care Res (Hoboken) 2017; 69:1706–13.
4. British Columbia Ministry of Health. About PopData. 2018. URL: https://www.popdata.bc.ca/.
5. British Columbia Ministry of Health. Medical Services Plan (MSP) Payment Information File V2. 2014. URL: http://www.popdata.bc.ca/data.
6. Canadian Institute for Health Information. Discharge Abstract Database (Hospital Separations).V2. Population Data BC, Data Extract. MOH. 2014. URL: http://www.popdata.bc.ca/data.
7. BCMoHc. Consolidation File (MSP Registration & Premium Billing). V2. Population Data British Colombia. Data Extract. MOH. 2014. URL: http://www.popdata.bc.ca/data.
8. British Colombia Vital Statistics Agency. Vital Statistics Deaths. V2. Population Data British Colombia. Data Extract. BC Vital Statistics Agency. 2014. URL: http://www.popdata.bc.ca/data.
9. British Columbia Ministry of Health. PharmaNet. V2. British Colombia Ministry of Health. Data Extract. Data Stewardship Committee. 2014. URL: http://www.popdata.bc.ca/data.
10. Katz JN, Barrett J, Liang MH, Bacon AM, Kaplan H, Kieval RI, et al. Sensitivity and positive predictive value of medicare part B physician claims for rheumatologic diagnoses and procedures. Arthritis Rheum 1997;40:1594–600.
11. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005;353:487–97.
12. Cortes S, Chambers S, Jeronimo A, Isenberg D. Diabetes mellitus complicating systemic lupus erythematosus: analysis of the UCL lupus cohort and review of the literature. Lupus 2008; 17:977–80.
13. Galo JS, Mehat P, Rai SK, Avina-Zubieta A, De Vera MA. What are the effects of medication adherence interventions in rheumatic diseases: a systematic review. Ann Rheum Dis 2016;75:667–73.
14. De Vera MA, Choi H, Abrahamowicz M, Kopec J, Goycochea-Robles MV, Lacaille D. Statin discontinuation and risk of acute myocardial infarction in patients with rheumatoid arthritis: a population-based cohort study. Ann Rheum Dis 2011;70:1020–4.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 707–716
DOI 10.1002/acr.24155

American College of Rheumatology
Empowering Rheumatology Professionals

Clinical Significance of Monitoring Hydroxychloroquine Levels in Patients With Systemic Lupus Erythematosus: A Systematic Review and Meta-Analysis

Shivani Garg,[1] Rachna Unnithan,[1] Karen E. Hansen,[1] Nathalie Costedoat-Chalumeau,[2] and Christie M. Bartels[1]

Objective. Despite the pivotal role that hydroxychloroquine (HCQ) plays in treating systemic lupus erythematosus (SLE), less than 50% of patients take HCQ as prescribed. Measurement of HCQ blood levels can help clinicians distinguish nonadherence versus lack of efficacy of HCQ. Our objective was to systematically review publications and perform a meta-analysis to examine the correlation between HCQ levels and 1) nonadherence and 2) Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) scores, in SLE.

Methods. A comprehensive search was performed. We included observational and interventional studies that measured HCQ levels and assessed adherence or SLEDAI scores in adults with SLE. Forest plots compared pooled estimates of correlations between HCQ levels and reported nonadherence or SLEDAI scores.

Results. Among 604 studies screened, 17 were reviewed. We found 3-times higher odds of reported nonadherence in patients with low HCQ levels (odds ratio 2.95 [95% confidence interval (95% CI) 1.63, 5.35], $P < 0.001$). The mean SLEDAI score was 3.14 points higher in groups with below-threshold HCQ levels on a priori analysis ($\delta = 3.14$ [95% CI –0.05, 6.23], $P = 0.053$), and 1.4 points higher in groups with HCQ levels of <500 ng/ml ($\delta = 1.42$ [95% CI 0.07, 2.76], $P = 0.039$). Among 1,223 patients, those with HCQ levels ≥750 ng/ml had a 58% lower risk of active disease, and their SLEDAI score was 3.2 points lower.

Conclusion. We found a strong association between low HCQ levels and reported nonadherence. Our results suggest that HCQ levels of ≥750 ng/ml might be a potential therapeutic target.

INTRODUCTION

Hydroxychloroquine (HCQ) is recommended for all patients with systemic lupus erythematosus (SLE or lupus) to reduce disease activity and improve damage-free survival (1–8). Nevertheless, up to 83% of lupus patients are nonadherent to HCQ (9,10) commonly because of poor understanding of the benefits of HCQ, lack of motivation to continue therapy, and inflated concerns regarding side effects from HCQ use. Further, a Medicaid lupus study showed that only 17% of patients with lupus were adherent to HCQ therapy (10). Although nonadherence is high and addressable, over two-thirds of rheumatologists are unaware of HCQ nonadherence (11,12). When asked why, clinicians report that the validated gold-standard adherence tools lack feasibility for use in routine follow-up visits (13). Therefore, nonadherence is not assessed or addressed and can result in 37% higher hospital admissions, 37% higher risk of end-stage renal disease, and an 8-fold higher risk of death (6,14,15).

Conversely, some patients have refractory lupus, despite taking HCQ regularly, and require treatment escalation (11,16–18). Diagnosing nonadherence is therefore a critical step in the care of patients who have uncontrolled lupus. To identify nonadherence, some researchers recommend routine testing of HCQ levels (11,12,17–19). Studies underscore a significant role for routine monitoring of HCQ levels as a measure of HCQ nonadherence, disease activity, and ability to predict lupus flares (12,19–21). Furthermore, studies report that regularly measuring HCQ levels during clinic visits can improve subsequent adherence to HCQ (12,19). Despite this finding, there is insufficient information on the overall clinical impact of incorporating routine testing of HCQ blood levels. Therefore, the objective of this study was to systematically review and analyze the published literature to examine

[1]Shivani Garg, MD, MS, Rachna Unnithan, MD, Karen E. Hansen, MD, MS, Christie M. Bartels, MD, MS: University of Wisconsin School of Medicine and Public Health, Madison; [2]Nathalie Costedoat-Chalumeau, MD, PhD: AP-HP, Cochin University Hospital and Université de Paris, CRESS, INSERM, INRA, F-75004 Paris, France.

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

Address correspondence to Shivani Garg, MD, MS, University of Wisconsin–Madison, 1685 Highland Avenue, MFCB #4122, Madison, WI 53705. Email: sgarg@medicine.wisc.edu.

Submitted for publication July 24, 2019; accepted in revised form January 21, 2020.

SIGNIFICANCE & INNOVATIONS

- Our unique meta-analysis highlights a strong association between blood hydroxychloroquine (HCQ) levels and patient- or physician-reported HCQ nonadherence in systemic lupus erythematosus.
- Routine monitoring of HCQ levels showed a significant improvement of HCQ levels on subsequent follow-up.
- All studies measuring HCQ blood levels demonstrated that flares strongly correlated with low HCQ levels, and the levels predicted Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) score change. Groups with HCQ blood levels ≥500 ng/ml showed a significant reduction in SLEDAI score by 1.4 points. Individuals with HCQ levels ≥750 ng/ml had significantly lower SLEDAI scores and a lower risk of having active disease.
- We propose further study of HCQ blood levels to test using the 750 ng/ml threshold to improve adherence and disease activity and to reduce flares.

the clinical significance of measuring HCQ levels among lupus patients, including both SLE and cutaneous SLE. We hypothesized that low HCQ levels would correlate with reported nonadherence, higher disease activity, and lupus flares.

MATERIALS AND METHODS

A comprehensive search was performed in Medline, Embase, CINHL, and Web of Science databases in August 2018 and in June 2019. The results were limited to articles published between January 1, 1997, and June 1, 2019. The search question included the Medical Subject Heading terms and keywords "systemic lupus erythematosus" OR "lupus" OR "SLE" OR "lupus nephritis" OR "lupus vasculitis" OR "CNS lupus" OR "discoid lupus" OR "cutaneous lupus" AND "concentrations" OR "weight" OR "measures" OR "levels" OR "quantification" AND "antimalarial" OR "hydroxychloroquine" OR "Plaquenil" OR "HCQ." Additional studies were identified through checking the references and other publications of the articles selected for full text review using Medline and Web of Science.

During the initial screening, 2 independent reviewers (SG and RU) screened the retrieved articles by reading the title and abstract to identify the studies that met the a priori list of inclusion criteria (Table 1). Any discrepancy on the decision to include or exclude a study was resolved by a third reviewer (CMB). We selected observational and interventional studies on human subjects measuring HCQ levels and reporting adherence or disease activity or flares. Case reports, expert opinion, reviews, and abstracts were excluded. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis guidelines and checklist were used for the literature identification process (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24155/abstract) (22). The study protocol was registered with PROSPERO (CRD42018107151).

Quality assessment of the observational studies was performed using the Newcastle Ottawa Scale (23), and for the interventional studies, it was assessed with the Cochrane Collaboration Risk Assessment (24) tool. An overall score was generated using the standard guidelines on quality assessment from the Agency for Healthcare Research and Quality (AHRQ) (25) and the Cochrane handbook. Both reviewers performed the quality assessment independently, and any discrepancies were resolved by the third reviewer. The data extraction Excel sheet was developed for extracting an a priori list of variables, including patient demographics, disease characteristics, threshold HCQ levels, reported adherence, and disease activity from the studies included in our review. When appropriate, the original authors of the studies were contacted for further information. Data were extracted by 1 reviewer and verified by the other. Disagreements were settled by the third reviewer.

The primary outcome of the study was reported HCQ nonadherence. Patient-reported nonadherence was defined as <80% medication adherence reported. Physician-reported adherence was estimated based on physicians' interpretations of the previous month's adherence, as reported by patients during clinic visits. Analogously, adherence was reported using a 0–10 scale, with <8 considered to be nonadherent to HCQ (21). The secondary outcome of the study was SLE disease activity, defined as >3-point increase in the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) score (26).

Table 1. Inclusion and exclusion criteria for screening and full text review*

Inclusion criteria
Studies among lupus patients taking HCQ or antimalarial
Studies done in the past 20 years
Studies of human subjects
Full text available
Studies done on measuring HCQ levels
Studies reporting a correlation or change in adherence OR disease activity (e.g., change in prednisone use, disease activity score, clinical assessment, physician global assessment, or flare frequency)
Exclusion criteria
Animal studies (nonhumans)
Studies done before 1997 (no HCQ blood level testing done before this time)
Studies done on other autoimmune diseases (no lupus)
Studies not measuring HCQ levels
Studies measuring HCQ levels but not examining correlation or change in adherence or disease activity
Gray literature: abstracts (incomplete data reported), case reports (no comparison group, hence inability to calculate effect size), conference proceedings or reviews or expert opinions (as we searched extensively and included primary studies/referenced studies from review/expert opinions/abstracts to be included in review)

* HCQ = hydroxychloroquine.

For the primary outcome assessment, we extracted data from patient- or physician-reported HCQ nonadherence in groups with low and high HCQ blood levels (between subjects) or at baseline and subsequent visits (within subjects). For the secondary outcome, we extracted the mean SLEDAI scores in groups with low and high HCQ levels or at baseline and subsequent visits. We used forest plots to compare the pooled estimates (with 95% confidence interval [95% CI]) of the correlation between HCQ levels and the patient- or the physician-reported nonadherence and the SLEDAI scores. Heterogeneity was assessed using I^2. Publication bias was evaluated by generating funnel plots. We systematically reviewed the literature for outcomes, lupus flare, and improvement in adherence, because the data were insufficient to perform pooled analysis.

Further, noting heterogeneity in HCQ threshold levels, we requested individual patient data for SLEDAI and HCQ levels from 5 studies that examined the correlation between SLEDAI score and HCQ blood levels (11,12,19,20,27), and the parent study (21) of 1 of the included nested studies (28). We received de-identified, limited data sets from 4 of these 6 studies. Based on recommendations from published studies of HCQ blood level thresholds, we categorized HCQ levels into 4 categories by ng/ml: <250, 250–499, 500–749, 750–999, and >1,000 (19,21,27–29). We used linear regression models to analyze the association between HCQ blood level categories and change in SLEDAI scores. Further, we used logistic regression to analyze the odds of active SLE (defined as SLEDAI score ≥6) (26) in patients with HCQ blood levels of ≥500 ng/ml and of ≥750 ng/ml compared to those with lower levels. Finally, we grouped HCQ blood level data on all 1,223 patients into 2 groups, with HCQ levels less than or greater than 500 ng/ml and calculated the mean SLEDAI score in each group. Including the mean SLEDAI scores from the pooled individual patient data analysis and the 2 other studies that used the ≥500 ng/ml threshold, we then compared the pooled estimate correlation between HCQ levels of ≥500 ng/ml and SLEDAI scores.

RESULTS

Study selection. Our initial literature search yielded 671 articles. After removing duplicates, 604 articles were included in the initial screening (details in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24155/abstract) (22). Twenty-seven studies met inclusion criteria for second-level review. Ultimately, 17 studies were selected for final data extraction and quality assessment (11,12,16,19–21,27–35) (details in Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24155/abstract) (22). There were 13 observational studies (prospective = 9, retrospective = 2, cross-sectional = 2) and 4 interventional studies. Most were single-center studies performed in Europe (Europe = 10, Asia = 3, US = 3, Australia = 1), and most were conducted after 2005 (n = 15). The overall AHRQ risk of bias assessment showed 8 "fair" to "good quality" observational studies and 3 "unclear" to "low risk" interventional studies (Table 2).

SLE study population and disease characteristics. Patient demographics are summarized in Table 2. In included studies, an SLE and/or cutaneous SLE diagnosis was validated using the American College of Rheumatology 1997 criteria (n = 10) (36), SLICC 2012 criteria (n = 3), cases from registry-validated board-certified rheumatologist diagnosis (n = 1), or histopathologic findings on cutaneous biopsy with physician validation (n = 3). Overall, 14 studies included patients with SLE, 2 studies included both SLE and cutaneous SLE, and 1 included exclusively patients with cutaneous SLE. Most studies included patients with active disease, and 1 included patients with clinically quiescent SLE (30).

Measuring HCQ levels and threshold levels. HCQ levels were measured using 3 methods. High-performance liquid chromatography (HPLC) was used in 14 studies, mass spectrometry (MS) in 2 studies, and both HPLC and MS in 1 study (Table 2). HCQ levels were measured in blood in 14 studies, both in serum and blood in 2 studies (34,37), and only in serum in 1 (33).

The targeted threshold HCQ levels differed between studies. Six studies used <205 ng/ml (129–205 ng/ml) as the threshold HCQ blood level to identify severe nonadherence, and other studies used higher HCQ levels of 500–1,000 ng/ml as a therapeutic threshold (Table 2). The studies also evaluated the correlation between HCQ levels and the glomerular filtration rate (GFR) (n = 6), body mass index (BMI) (n = 8), and smoking (n = 8), as summarized in Table 2. One study reported a correlation between high estimated GFR and low HCQ levels (28). Another reported lower HCQ levels in patients with renal impairment due to routine underdosing in renal insufficiency (200 mg versus 400 mg daily) (19), and the remaining studies reported no correlation with renal function. Three studies reported a correlation between high BMI and low HCQ levels (28,29,34). One study with only 9 smokers reported a borderline correlation between active smoking and higher HCQ levels on unadjusted analysis (P = 0.08) (32), but this finding was not confirmed by another study (28).

Correlation between patient- or physician-reported nonadherence and HCQ levels. We found that 7 studies measured patient- or physician-reported nonadherence, using self-report questionnaires or a physician reporting tool (11,12,19,21,32,38) or pharmacy refill information (35), in addition to HCQ levels. We obtained data on 4 of these studies; the remaining 3 studies had limited comparative data due to study design (12,19,32). Three studies recorded patient-reported nonadherence or patient-reported missed doses (11,35,38), and 1 study used physician-reported nonadherence (11,21). The pooled

Table 2. Study and patient demographics*

Author, year (ref.)	Study population	HCQ-level threshold (sample, method)	Outcomes reported	Bias assessment tool: risk assessed
Francès et al, 2012 (38)	(n = 300) age 43.6 years (12–85 years), F 84%, smoker 41%, SLE 33%, CLE 77%	>200 ng/ml (blood, HPLC)	1) Patient-reported nonadherence and correlation with HCQ levels; 2) cutaneous disease activity and HCQ levels	NOS: good
Carmichael et al, 2013 (37)	(n = 60) mean ± SD age 42.8 ± 15.2 (18–74 years), F 95%, SLE 100%, CrCl 100.8 ± 36.3 ml/min	>750 ng/ml (blood and plasma, HPLC)	SLAM correlation with HCQ levels	NOS: poor
Chasset et al, 2016 (16)	(n = 34) age 45 years (28–72 years), F 78%, smoker 47%, CLE 100%, SLE 53%	>750 ng/ml (blood, HPLC)	Improvement in CLE (>4 points increase or 20% change)	CCRA: unclear
Costedoat et al, 2006 (20)	(n = 143) mean ± SD age 35 ± 11 years, F 93%, smoker 23%, SLE, GFR 90 ± 2 ml/min, BMI 23 ± 4 kg/m^2	>1,000 ng/ml (blood, HPLC)	1) SLEDAI correlation with HCQ levels; 2) flare correlation with HCQ levels	NOS: good
Costedoat-Chalumeau et al, 2007 (12)	(n = 203) mean ± SD age 35 ± 11 years, F 91%, White 66%, smoker 22%, SLE, BMI 23 ± 4 kg/m^2	>205 ng/ml (blood, HPLC)	1) Patient-reported nonadherence and correlation with HCQ levels; 2) disease activity and HCQ levels; 3) number of patients with flare and HCQ levels	NOS: poor
Costedoat-Chalumeau et al, 2013 (21)	(n = 171) mean ± SD age 40 ± 11 years, F 87%, smoker 24%, SLE, BMI 25 ± 5 kg/m^2	>1,000 ng/ml (blood, HPLC)	1) Physician-reported nonadherence in correlation with HCQ levels; 2) SLEDAI correlation with HCQ levels; 3) number of patients with flare and HCQ levels	CCRA: unclear
Costedoat-Chalumeau et al, 2018 (11)	(n = 305) mean ± SD age 37.7 ± 11.6 years, F 94.4%, White 50%, smoker 14.5%, SLE, BMI 23.2 (21.2–28) kg/m^2, CrCl 111 (87–132) ml/min	>200 ng/ml (blood, HPLC)	1) Patient-reported nonadherence in correlation with HCQ levels; 2) SLEDAI and HCQ levels; 3) number of patients with flare and HCQ levels	NOS: good
Cunha et al, 2017 (29)	(n = 171) mean ± SD age 39.8 ± 15.6 years, F 86%, White 16%, Indian subcontinent ancestry 43%, LN	>200 ng/ml (blood, HPLC)	1) LN activity correlation with HCQ levels; 2) flare frequency correlation with HCQ levels	NOS: fair
Durcan et al, 2015 (19)	(n = 686) age <45 years 45%, F 92%, White 49%, African American 42%, SLE, GFR >90 ml/min 90%, BMI >30 kg/m^2 31%	>500 ng/ml (blood, HPLC)	1) Improvement in adherence with routine HCQ-level monitoring; 2) SLEDAI and HCQ levels	NOS: poor
Geraldino-Pardilla et al, 2019 (27)	(n = 108) age 38 years, F 91%, SLE	>500 ng/ml (blood, HPLC)	SLEDAI score correlation with HCQ levels	NOS: good
Iudici et al, 2018 (30)	(n = 83) mean ± SD age 41 ± 11 years, F 95%, smoker 37%, SLE, BMI 25 ± 5 kg/m^2, CrCl 89 ± 24 ml/min	>100 ng/ml (blood, UPLC-MS)	Number of flares in correlation with HCQ levels	NOS: good
Jallouli et al, 2015 (28)	(n = 509) mean ± SD age 30 ± 11.5 years, F 91%, White 55%, smoker 23%, SLE, BMI 23.9 ± 2.4 kg/m^2, GFR 103.5 ± 31.5 ml/min	>917 ng/ml (blood, HPLC)	Correlation of SLEDAI with HCQ levels	NOS: good
Jolly et al, 2016 (31)	(n = 171) mean ± SD age 44.4 ± 10.7 years, F 87%, White 73%; SLE	>750 ng/ml (blood, HPLC)	Correlation of SLEDAI with HCQ levels	NOS: fair
Lee et al, 2017 (32)	(n = 189) mean ± SD age 39.1 ± 11.6 years, F 93.7%, smoker 4.8%, SLE, BMI 22.0 ± 3.4 kg/m^2, GFR >90 ml/min 72.5%	>100 ng/ml (blood, HPLC)	1) SLEDAI in correlation with HCQ levels; 2) initial patient-reported nonadherence	NOS: fair
Mok et al, 2016 (33)	(n = 276) mean ± SD age 41 ± 14 years, F 93%, SLE, GFR 95.6 ± 31.3 ml/min	>500 ng/ml (serum, MS)	1) SLEDAI correlation with serum HCQ levels; 2) flare frequency correlating with serum HCQ levels	NOS: poor
Morita et al, 2016 (34)	(n = 103) mean ± SD age 42.5 ± 12.2 years, F 64%, CLE 100%, SLE 55%, BMI 22.8 ± 4 kg/m^2	Ctrough 462 ng/ml, Cmax 942.7 ng/ml (blood and serum, HPLC and MS)	HCQ levels in Japanese patients with CLE/SLE and clinical efficacy of HCQ without retinopathy	CCRA: unclear
Ting et al, 2012 (35)	(n = 41) mean ± SD age 18.6 ± 2.5 years, F 93%, SLE	>900 ng/ml (blood, HPLC)	Patient-reported nonadherence correlation with HCQ levels and pharmacy refills	CCRA: high

* BMI = body mass index; CCRA = Cochrane Collaboration risk assessment; CLE = cutaneous lupus erythematosus; CrCl = creatinine clearance; F = female; GFR = glomerular filtration rate; HCQ = hydroxychloroquine; HPLC = high-performance liquid chromatography; LN = lupus nephritis; ml/min = milliliters/minute; MS = mass spectrometry; NOS = Newcastle Ottawa Scale; ref. = reference; SLAM = Systemic Lupus Activity Measure; SLE = systemic lupus erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; UPLC = ultraperformance liquid chromatography.

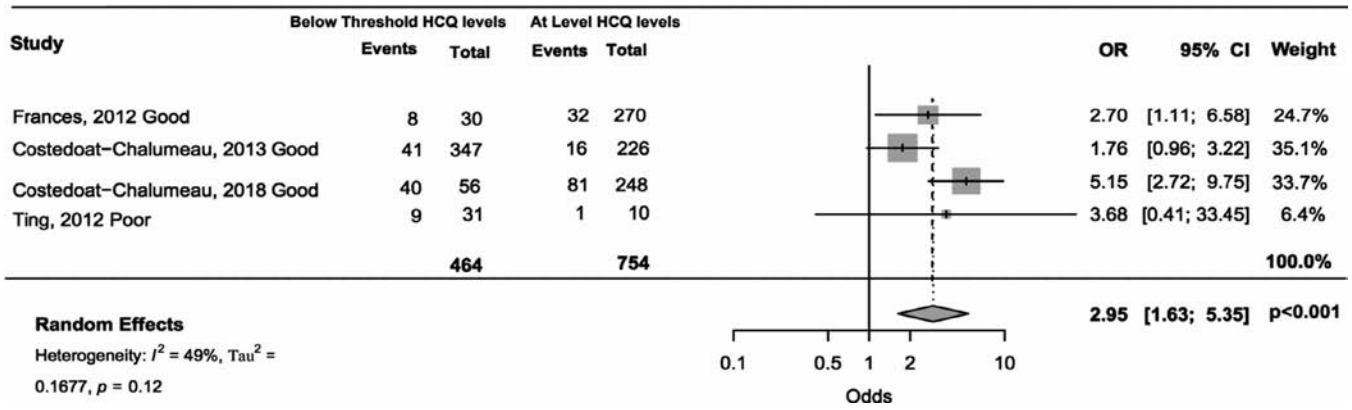


Figure 1. Forest plot showing the association between low hydroxychloroquine (HCQ) blood levels and nonadherence. 95% CI = 95% confidence interval; OR = odds ratio.

odds of nonadherence reported by patient or physician was 3-fold higher in patients with low HCQ levels, compared to patients with higher HCQ levels (odds ratio [OR] 2.95 [95% CI 1.63, 5.3], $P < 0.001$, $I^2 = 49\%$) (Figure 1). The threshold for HCQ blood level indicating nonadherence was variable in different studies. The funnel plot showed asymmetry, indicating moderate publication bias.

Improvement in nonadherence with regular HCQ-level monitoring. Three studies reported that measurement of HCQ levels improved subsequent adherence (12,19,30). These studies reported that the recorded improvement could have resulted from increased physician awareness, followed by discussing adherence strategies, or from patient behavior changes because of regular monitoring. Furthermore, 2 studies reported that an increase in adherence discussions and counseling sessions by rheumatologists led to an increase in HCQ adherence in subsequent visits (12,19). The data obtained from these studies were qualitative, which limited meta-analysis.

Correlation of disease activity and HCQ levels. We included 10 studies that reported measuring disease activity and measured HCQ levels. We excluded 1 study that measured serum HCQ levels because HCQ serum levels lack reliability compared to blood levels (33), and we excluded 1 nested study derived from another included study (31). One study examined the correlation between HCQ level and high versus low disease activity using the Systemic Lupus Activity Measure (SLAM) instrument but did not report actual disease activity scores; therefore, this study was not included in the pooled analysis (37). In the pooled analysis, the mean SLEDAI score was 3.14 points higher in groups with HCQ levels below therapeutic threshold ($\delta = 3.14$ [95% CI –0.05, 6.23], $P = 0.053$) (Figure 2A). A sensitivity analysis did not change the results. Very high study heterogeneity was observed ($I^2 = 98\%$). Upon including all 8 studies in the meta-analysis, we found a statistically significant increase in the mean SLEDAI score by 1.23 points in groups with low HCQ levels ($\delta = 1.23$ [95% CI 0.26, 2.23], $P < 0.01$. (data not shown).

Individual patient data and pooled data analysis to examine correlations between SLEDAI score and HCQ levels. We found that SLEDAI scores decreased by 1.7 with each stepwise 250 ng/ml increase in HCQ blood levels from <250 to 749 ng/ml ($P < 0.001$) (Table 3). Furthermore, we found a statistically significant decrease in SLEDAI scores by 3.2 points with an increase in HCQ blood levels in ng/ml from <250 to ≥750 ($P < 0.0001$) (Table 3). Using logistic regression, we found that patients with HCQ levels of ≥500 ng/ml had a 56% lower risk of active SLE (OR 0.44 [95% CI 0.33, 0.59], $P < 0.001$), and that HCQ levels of ≥750 ng/ml predicted a 58% lower risk of active SLE compared to patients with lower levels (OR 0.42 [95% CI 0.2, 0.55], $P < 0.0001$). Finally, after pooling the mean SLEDAI scores from 4 studies that shared individual patient data (n = 1,223) and 2 other studies that used the ≥500-ng/ml threshold, we found a significantly lower SLEDAI score in patients with levels of ≥500 ng/ml (mean SLEDAI score difference = 1.42 [95% CI 0.07, 2.8], $P = 0.04$) (Figure 2B).

Correlations of flares with HCQ levels. Seven studies reported an increase in the number of SLE flares or patients with flares in groups with low HCQ levels (Table 4). Most studies reported a higher number of lupus flares in groups with low HCQ levels (11,12,20,21,30). Only 1 study, which used HCQ serum levels, reported no significant link between HCQ levels and flares (33). One study reported a 6-fold higher odds of flare in patients with low compared to patients with high HCQ levels (OR 5.89 [95% CI 1.38, 25.08]) (20). Cunha et al reported a higher flare frequency in patients with HCQ levels of <620 ng/ml ($P = 0.041$) (29). Meta-analysis was not performed, given the limited number of studies with data on patient-level or event-level flare.

Improvement in cutaneous SLE with high HCQ levels. Three studies documented within-subject improvements in Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) scores after these subjects achieved high HCQ levels (16,34,38). One study reported that the relative risk of

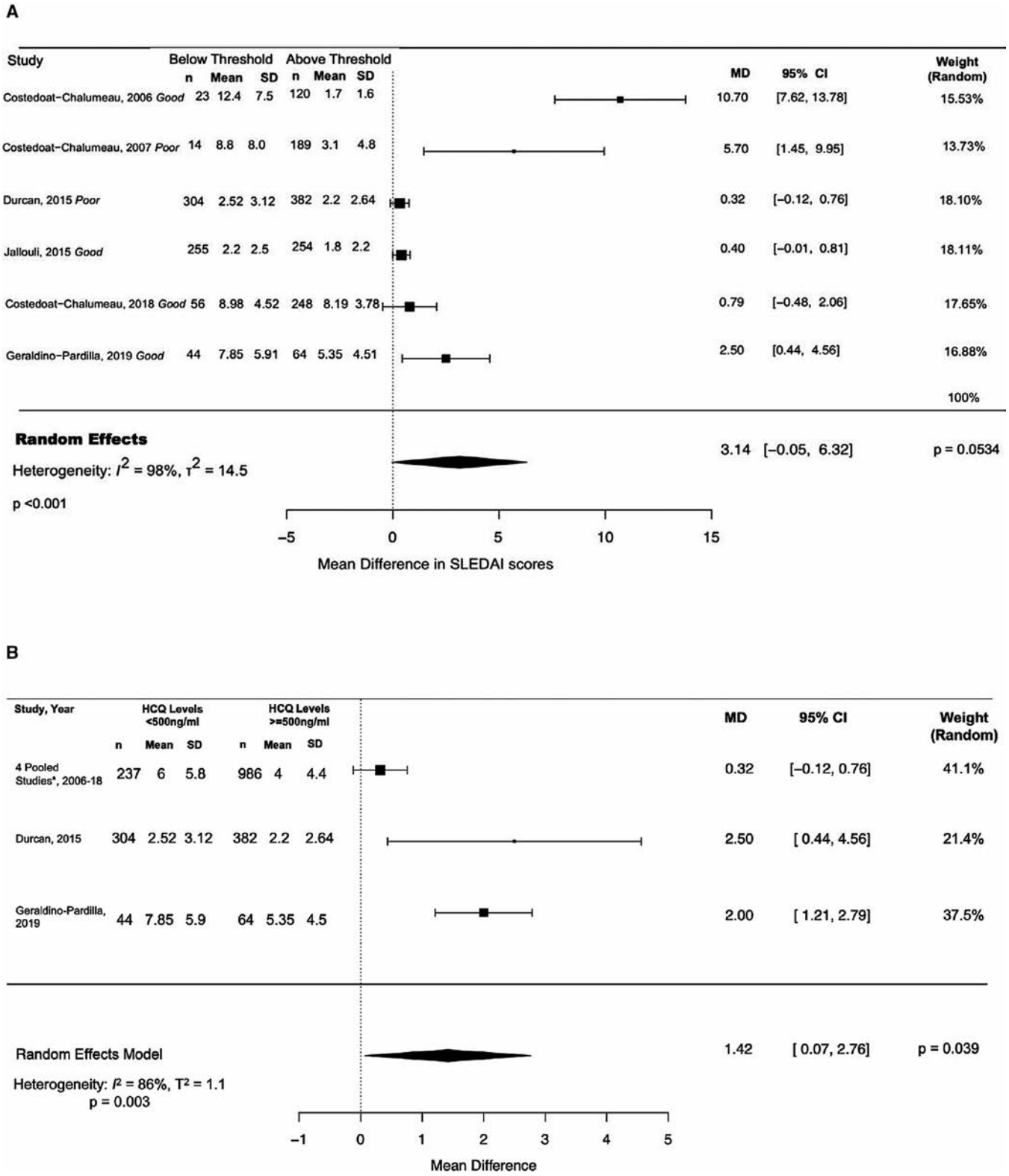


Figure 2. **A**, Forest plot correlating hydroxychloroquine (HCQ) blood levels below or above threshold and the mean Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) score using a priori analysis. **B**, Forest plot correlating HCQ levels below or above 500 ng/ml and the mean SLEDAI score. 95% CI = 95% confidence interval; MD = mean difference; * = using individual patient HCQ levels from these 4 studies.

Table 3. Change in SLEDAI score by increase in HCQ levels, using individual patient data (n = 1,223)*

HCQ levels, categories	Values	*P*
Mean change in SLEDAI (95% CI)		
<250 ng/ml	Ref.	Ref.
250–499 ng/ml	–1.7 (–0.06, –2.8)	0.004
500–749 ng/ml	–1.8 (–0.8, –2.7)	<0.001
≥750 ng/ml	–3.2 (–2.2, –4.2)	<0.001
Odds ratio of active SLE (95% CI)†		
<500 ng/ml	Ref.	Ref.
≥500 ng/ml	0.44 (0.33, 0.58)	<0.001
<750 ng/ml	Ref.	Ref.
≥750 ng/ml	0.42 (0.22, 0.49)	<0.001

* 95% CI = 95% confidence interval; HCQ = hydroxychloroquine; Ref. = reference; SLE = systemic lupus erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index.
† Odds of active SLE (SLEDAI ≥6) by threshold HCQ levels (n = 1,223).

complete remission increased by 1.00073 (95% CI 1.0002, 1.156; *P* = 0.005) for every 1 ng/ml increase in HCQ level (38). Two other studies reported a significant improvement in the median CLASI score in patients with high HCQ levels (change from 8 to 1.5; *P* < 0.001; –8.0 with mean HCQ level 853 ng/ml) (16,34).

DISCUSSION

We performed a systematic review and meta-analysis to evaluate the clinical significance of routinely measuring HCQ levels. We found strong association between HCQ levels and adherence. Additionally, measurement of HCQ blood levels improved adherence. While SLEDAI scores were numerically 3 points higher in patients with below-HCQ threshold levels, results were not statistically significant in the a priori analysis (*P* = 0.0534). However, when pooling individual patient data from 4 studies that dichotomized for levels less than or greater than/equal to 500 ng/ml and data from 2 other studies that used the ≥500-ng/ml threshold, we found a significant reduction in the SLEDAI score by 1.4 points in patients with ≥500 ng/ml (*P* = 0.04). Further, using individual patient data, we found that patients with levels of ≥750 ng/ml had both a clinically meaningful and statistically significant decrease in SLEDAI scores by 3.2 points, and such patients had a 58% lower risk of active SLE compared to those with <750 ng/ml. Our findings support the use of levels of <750 ng/ml as a clinical threshold to predict active lupus and lupus disease flares. Finally, we found good evidence (6 studies) that low HCQ levels predicted lupus flares and higher CLASI scores.

Our findings are particularly relevant amid reports regarding the serious magnitude and consequences of HCQ nonadherence. Feldman et al reported that 83% of the lupus patients in a Medicaid cohort were nonadherent with HCQ (10). They reported that younger age, Black race, and poor health literacy were common predictors of HCQ nonadherence (10). Researchers state that despite the alarming rate of HCQ nonadherence, it remains unaddressed during routine lupus visits (39) due to a lack of standard clinical assessment of adherence.

Suggesting a possible future clinical standard, our meta-analysis supports previously published studies reporting that the

Table 4. Studies reporting a correlation in lupus flares and HCQ levels*

Author, year (ref.)	Comparator groups	Flares in groups, with *P* value	Outcome
Costedoat-Chalumeau et al, 2006 (20)	Low HCQ levels <1,000 ng/ml (n = 14) vs. high HCQ levels >1,000 ng/ml (n = 106)	14 flares (low HCQ level) vs. 0 flares (high HCQ level), OR 5.89 (95% CI 1.38, 25.08), *P* = 0.01†	6-fold higher odds of flare with <1,000 ng/ml HCQ levels
Costedoat-Chalumeau et al, 2007 (12)	Low HCQ levels (<200 ng/ml; n = 14) vs. high HCQ levels (>200 ng/ml; n = 189)	7 flares (low HCQ level) vs. 18 flares (high HCQ level), *P* = 0.004	Higher flares in group with lower HCQ level
Costedoat-Chalumeau et al, 2013 (21)	1–7 months follow-up: low HCQ levels (<1,000 ng/ml; n = 55) vs. high HCQ levels (>1,000 ng/ml; n = 36)	36 flares (low HCQ level) vs. 5 flares (high HCQ level), *P* = 0.04; OR for flare = 3.82 (95% CI 1.16, 12.58), *P* = 0.027†	Negative correlation of flares with higher HCQ levels after 1st month; 4-fold higher odds of flare in group with low HCQ level
Costedoat-Chalumeau et al, 2018 (11)	Low HCQ levels (<200 ng/ml; n = 56) vs. high HCQ levels (>200 ng/ml; n = 248)	20 flares (low HCQ level) vs. 112 flares (high HCQ level), *P* = 0.20	No correlation reported
Cunha et al, 2017 (29)	Patients with flare vs. patients with no flare	HCQ level = 590 ng/ml (group with flare) vs. HCQ level = 810 ng/ml (group with no flare), *P* = 0.005	Higher flare frequency in patients with HCQ level <620 ng/ml
Iudici et al, 2018 (30)	77 follow-ups, no groups	5 patients with flares (HCQ level 284 ng/ml) vs. 72 without flares (HCQ level 435 ng/ml), *P* = 0.225	More flares in patients with lower HCQ level, although statistical significance not achieved
Mok et al, 2016 (33)‡	<10 ng/ml HCQ level (n = 31) vs. 10–500 ng/ml HCQ level (n = 212) vs. >500 ng/ml HCQ level (n = 33)	Mean ± SD 0.14 ± 0.42 flares/year (<10 ng/ml HCQ level); 0.12 ± 0.29 flares/year (10–500 ng/ml HCQ level); 0.19 ± 0.57 flares/year (>500 ng/ml HCQ level), *P* = 0.82	No correlation reported between serum HCQ levels and annual flare frequency

* 95% CI = 95% confidence interval; HCQ = hydroxychloroquine; OR = odds ratio; ref. = reference.
† Adjusted analysis.
‡ Serum HCQ level.

HCQ blood level is a good objective measure of severe nonadherence in lupus. Previous studies have reported a moderate correlation between HCQ blood levels and reported nonadherence (r = 0.37) and lower HCQ levels in patients reporting nonadherence (37). We reported strong association between reported nonadherence and low HCQ levels (OR 2.95 [95% CI 1.63, 5.35]). Finally, researchers report particularly strong association between very low HCQ levels (<205 ng/ml) and severe nonadherence (<15%), recommending 200 ng/ml as a minimum diagnostic threshold for severe nonadherence (11).

We and others have also reported a significant improvement in HCQ adherence over time in lupus patients with routine longitudinal monitoring of HCQ levels (7% to 2%; 56% to 80%; $P < 0.001$) (11,19). Researchers have reported that routine HCQ-level monitoring led to improvement in adherence because of increased physician awareness, more in-clinic discussions to address barriers to adherence, and patient behavior changes (12,19,21).

Conversely, some physicians challenge the routine measurement of HCQ levels because few US laboratories measure HCQ levels, the cost ranges from $70 to $250 per test, and the cost might not be covered by some US insurance plans (40). However, the cost of measuring HCQ levels is not charged for patients in the hospital or in standard-of-care in France and some European countries (41). Moreover, there is no consensus on the HCQ blood-level thresholds to identify nonadherence. Using the proposed blood level of <200 ng/ml identifies severe nonadherence, but this level can be quickly achieved, even with sporadic HCQ use the week before the HCQ-level measurement. When measuring objective HCQ levels is not feasible, our meta-analysis underscores the clinical utility of asking patients directly or using questionnaires to identify nonadherence during routine clinic visits. However, physicians identify the need for training to effectively discuss and address nonadherence with patients (42).

Lupus flares are the leading cause of hospital admission in SLE (15,43), and researchers recommend that measuring HCQ levels could help (12,19,20). Two other included studies reported a clinically significant change in the SLEDAI score (defined as >3-point increase in SLEDAI score) in patients with low HCQ levels compared to those with high HCQ levels (12,20). One study suggested that HCQ blood levels of <500 ng/ml strongly predicted active lupus (SLEDAI score 7.9 [<500 ng/ml] versus SLEDAI score 5.9 [>500 ng/ml]; $P < 0.001$) (32). One study on clinically stable patients with SLE (mean SLEDAI score 2.2) showed a significant trend for a decrease in SLEDAI scores with increasing HCQ levels ($P = 0.04$) (19). Finally, unlike other studies, 1 cross-sectional study reported no correlation between SLAM categories and high versus low HCQ blood-level group ($r^2 = 0.21$, $P = 0.12$) (36). Our study underscores using the idea of HCQ levels of <750 ng/ml as a clinical threshold to identify the risk for active lupus and disease flares. Future prospective clinical trials could confirm a level of ≥750 ng/ml as a clinically relevant threshold for interpreting and targeting HCQ levels.

Most studies have reported that high HCQ levels negatively correlated with flare frequency and severity. The 2 included studies reported a 6-fold higher risk of flare in patients (OR 5.89 [95% CI 1.38, 25.08]) and a 4-fold higher flare frequency within 1–7 months of diagnosis (65% versus 14%; $P = 0.04$) (20,21). Overall, all studies monitoring HCQ blood levels consistently noted that low HCQ levels predicted lupus flares. Similarly, the 3 cutaneous SLE studies included in our systematic review reported a better clinical response with higher HCQ levels.

The HCQ blood-level threshold varied in most studies. Four studies reported <200 ng/ml as a minimum threshold to identify nonadherence (11,12,29,37). Other studies reported 96% negative predictive value of active SLE in patients with HCQ levels of >1,000 ng/ml due to complete inhibition of Toll-like receptors at this level (17,19,44,45). Most studies recommend >500 ng/ml as a therapeutic threshold level to identify adherence and stable SLE (18,19,45). Other researchers reported 620 ng/ml HCQ levels to be a good therapeutic threshold to prevent lupus flares (29). Our study supports levels of ≥750 ng/ml to be clinically meaningful and statistically significant to identify disease flare (change in SLEDAI score ≥3 points) and to predict active disease (SLEDAI score ≥6).

Despite the strengths of being the first study to systematically review and meta-analyze the clinical role of monitoring HCQ levels, we acknowledge the limitations. First, overall the number of studies measuring HCQ levels and nonadherence or SLEDAI score was limited (n = 4). Second, most of the studies that examined the correlation between reported adherence and HCQ blood levels were performed in Europe, and there was only 1 small US study (n = 31) (35). Therefore, generalizability for our findings could be limited because of differences in cultural beliefs, social issues, and insurance/medical coverage in populations from diverse countries. Third, there was heterogeneity among the included studies. Severe heterogeneity (I^2 >80%) was found in included studies examining the correlation between SLEDAI scores and HCQ levels, and sensitivity and individual patient data analyses did not significantly change results. Fourth, variation in reporting flare frequency by patient- versus event-level limited our ability to calculate a pooled odds ratio. Finally, there were few studies assessing the association between cutaneous SLE and HCQ levels.

In summary, our meta-analysis found strong association between reported nonadherence and low HCQ levels. Our systematic review reports improvements in adherence and disease activity with longitudinal monitoring of HCQ levels. Therefore, we recommend using objective measurement of HCQ blood levels to assess nonadherence during routine clinic visits, or using self-report questionnaires when measuring HCQ levels is not feasible. Our meta-analysis suggested a trend of higher lupus activity in groups with low HCQ levels in the a priori analysis. Individual patient data and the pooled analysis using HCQ levels showed a statistically significant decrease in the SLEDAI score by 1.4–1.7 in patients with HCQ levels of ≥500 ng/ml. Levels of ≥750 ng/ml showed both a

clinically meaningful and statistically significant decrease in SLEDAI scores by 3.2 points, suggesting that <750 ng/ml is a clinically relevant threshold for predicting disease flare or activity.

Our systematic review consistently reported a strong correlation with flares in patients with low HCQ levels. Future studies are needed to confirm clinically relevant thresholds to identify nonadherence and risk for active SLE. Finally, we also recommend that future studies assess the clinical role of periodic HCQ-level monitoring versus self-reported nonadherence in routine SLE care, particularly in settings with lower resources.

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. Garg 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. Garg, Unnithan, Hansen, Bartels.
Acquisition of data. Garg, Unnithan, Costedoat-Chalumeau, Bartels.
Analysis and interpretation of data. Garg, Hansen, Costedoat-Chalumeau, Bartels.

REFERENCES

1. Ruiz-Irastorza G, Ramos-Casals M, Brito-Zeron P, Khamashta MA. Clinical efficacy and side effects of antimalarials in systemic lupus erythematosus: a systematic review. Ann Rheum Dis 2010;69:20–8.
2. Akhavan PS, Su J, Lou W, Gladman DD, Urowitz MB, Fortin PR. The early protective effect of hydroxychloroquine on the risk of cumulative damage in patients with systemic lupus erythematosus. J Rheumatol 2013;40:831–41.
3. 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.
4. Tsakonas E, Joseph L, Esdaile JM, Choquette D, Senecal JL, Cividino A, et al. A long-term study of hydroxychloroquine withdrawal on exacerbations in systemic lupus erythematosus: the Canadian Hydroxychloroquine Study Group. Lupus 1998;7:80–5.
5. Fasano S, Pierro L, Pantano I, Iudici M, Valentini G. Longterm hydroxychloroquine therapy and low-dose aspirin may have an additive effectiveness in the primary prevention of cardiovascular events in patients with systemic lupus erythematosus. J Rheumatol 2017;44:1032–8.
6. Alarcon GS, McGwin G, Bertoli AM, Fessler BJ, Calvo-Alen J, Bastian HM, et al. Effect of hydroxychloroquine on the survival of patients with systemic lupus erythematosus: data from LUMINA, a multiethnic US cohort (LUMINA L). Ann Rheum Dis 2007;66:1168–72.
7. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Ann Rheum Dis 2019;78:736–45.
8. Hahn BH, McMahon MA, Wilkinson A, Wallace WD, Daikh DI, Fitzgerald JD, et al. American College of Rheumatology guidelines for screening, treatment, and management of lupus nephritis. Arthritis Care Res (Hoboken) 2012;64:797–808.
9. Mehat P, Atiquzzaman M, Esdaile JM, Aviña-Zubieta A, De Vera MA. Medication nonadherence in systemic lupus erythematosus: a systematic review. Arthritis Care Res (Hoboken) 2017;69:1706–13.
10. Feldman CH, Collins J, Zhang Z, Subramanian SV, Solomon DH, Kawachi I, et al. Dynamic patterns and predictors of hydroxychloroquine nonadherence among Medicaid beneficiaries with systemic lupus erythematosus. Semin Arthritis Rheum 2018;48:205–13.
11. Costedoat-Chalumeau N, Houssiau F, Izmirly P, Le Guern V, Navarra S, Jolly M, et al. A prospective international study on adherence to treatment in 305 patients with flaring SLE: assessment by drug levels and self-administered questionnaires. Clin Pharmacol Ther 2018;103:1074–82.
12. Costedoat-Chalumeau N, Amoura Z, Hulot JS, Aymard G, Leroux G, Marra D, et al. Very low blood hydroxychloroquine concentration as an objective marker of poor adherence to treatment of systemic lupus erythematosus. Ann Rheum Dis 2007;66:821–4.
13. Dobbels F, Berben L, De Geest S, Drent G, Lennerling A, Whittaker C, et al. The psychometric properties and practicability of self-report instruments to identify medication nonadherence in adult transplant patients: a systematic review. Transplantation 2010;90:205–19.
14. Feldman CH, Zhang Z, Desai RJ, Lin T, Collins J, Subramanian SV, et al. Association between hydroxychloroquine nonadherence and adverse outcomes among patients with systemic lupus erythematosus [abstract]. Annual Scientific Meeting of the American College of Rheumatology, San Diego, CA, November 2017.
15. 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.
16. Chasset F, Arnaud L, Costedoat-Chalumeau N, Zahr N, Bessis D, Frances C. The effect of increasing the dose of hydroxychloroquine (HCQ) in patients with refractory cutaneous lupus erythematosus (CLE): an open-label prospective pilot study. J Am Acad Dermatol 2016;74:693–9.
17. Al-Rawi H, Meggitt SJ, Williams FM, Wahie S. Steady-state pharmacokinetics of hydroxychloroquine in patients with cutaneous lupus erythematosus. Lupus 2018;27:847–52.
18. Arnaud L, Zahr N, Costedoat-Chalumeau N, Amoura Z. The importance of assessing medication exposure to the definition of refractory disease in systemic lupus erythematosus. Autoimmun Rev 2011;10:674–8.
19. Durcan L, Clarke WA, Magder LS, Petri M. Hydroxychloroquine blood levels in systemic lupus erythematosus: clarifying dosing controversies and improving adherence. J Rheumatol 2015;42:2092–7.
20. Costedoat-Chalumeau N, Amoura Z, Hulot JS, Hammoud HA, Aymard G, Cacoub P, et al. Low blood concentration of hydroxychloroquine is a marker for and predictor of disease exacerbations in patients with systemic lupus erythematosus. Arthritis Rheum 2006;54:3284–90.
21. Costedoat-Chalumeau N, Galicier L, Aumaitre O, Frances C, Le Guern V, Liote F, et al. Hydroxychloroquine in systemic lupus erythematosus: results of a French multicentre controlled trial (PLUS Study). Ann Rheum Dis 2013;72:1786–92.
22. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
23. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
24. 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.
25. Viswanathan M, Patnode CD, Berkman ND, Bass EB, Chang S, Hartling L, et al. Assessing the risk of bias in systematic reviews of health care interventions: methods guide for effectiveness and comparative effectiveness reviews. Rockville (MD): Agency for Healthcare Research and Quality; 2008.

26. Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol 2002;29:288–91.
27. Geraldino-Pardilla L, Perel-Winkler A, Miceli J, Neville K, Danias G, Nguyen S, et al. Association between hydroxychloroquine levels and disease activity in a predominantly Hispanic systemic lupus erythematosus cohort. Lupus 2019;28:862–7.
28. Jallouli M, Galicier L, Zahr N, Aumaître O, Francès C, Le Guern V, et al. Determinants of hydroxychloroquine blood concentration variations in systemic lupus erythematosus. Arthritis Rheumatol 2015;67:2176–84.
29. Cunha C, Alexander S, Ashby D, Lee J, Chusney G, Cairns TD, et al. Hydroxycloroquine blood concentration in lupus nephritis: a determinant of disease outcome? Nephrol Dial Transplant 2018;33:1604–10.
30. Iudici M, Pantano I, Fasano S, Pierro L, Charlier B, Pingeon M, et al. Health status and concomitant prescription of immunosuppressants are risk factors for hydroxychloroquine non-adherence in systemic lupus patients with prolonged inactive disease. Lupus 2018;27:265–72.
31. Jolly M, Galicier L, Aumaitre O, Frances C, Le Guern V, Liote F, et al. Quality of life in systemic lupus erythematosus: description in a cohort of French patients and association with blood hydroxychloroquine levels. Lupus 2016;25:735–40.
32. Lee JY, Lee J, Kwok SK, Ju JH, Park KS, Park SH. Factors related to blood hydroxychloroquine concentration in patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69:536–42.
33. Mok CC, Penn HJ, Chan KL, Tse SM, Langman LJ, Jannetto PJ. Hydroxychloroquine serum concentrations and flares of systemic lupus erythematosus: a longitudinal cohort analysis. Arthritis Care Res (Hoboken) 2016;68:1295–302.
34. Morita S, Takahashi T, Yoshida Y, Yokota N. Population pharmacokinetics of hydroxychloroquine in Japanese patients with cutaneous or systemic lupus erythematosus. Ther Drug Monit 2016;38:259–67.
35. Ting TV, Kudalkar D, Nelson S, Cortina S, Pendl J, Budhani S, et al. Usefulness of cellular text messaging for improving adherence among adolescents and young adults with systemic lupus erythematosus. J Rheumatol 2012;39:174–9.
36. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.
37. Carmichael SJ, Day RO, Tett SE. A cross-sectional study of hydroxychloroquine concentrations and effects in people with systemic lupus erythematosus. Intern Med J 2013;43:547–53.
38. Francès C, Anne Cosnes MD, Pierre Duhaut MD, Zahr N, Boutros Soutou B, Ingen-Housz-Oro S, et al. Low blood concentration of hydroxychloroquine in patients with refractory cutaneous lupus erythematosus. Arch Dermatol 2012;148:479–84.
39. Sleath B, Chewning B, de Vellis BM, Weinberger M, de Vellis RF, Tudor G, et al. Communication about depression during rheumatoid arthritis patient visits. Arthritis Rheum 2008;59:186–91.
40. Waimann CA, Marengo MF, de Achaval S, Cox VL, Garcia-Gonzalez A, Reveille JD, et al. Electronic monitoring of oral therapies in ethnically diverse and economically disadvantaged patients with rheumatoid arthritis: consequences of low adherence. Arthritis Rheum 2013;65:1421–9.
41. Costedoat-Chalumeau N, Le Guern V, Piette JC. Determining adherence to therapeutic regimens in patients with chronic illness: comment on the article by Waimann et al [letter]. Arthritis Rheumatol 2014;66:770.
42. Mushlin AI, Appel FA. Diagnosing potential noncompliance: physicians' ability in a behavioral dimension of medical care. Arch Intern Med 1977;137:318–21.
43. Yazdany J, Marafino BJ, Dean ML, Bardach NS, Duseja R, Ward MM, et al. Thirty-day hospital readmissions in systemic lupus erythematosus: predictors and hospital- and state-level variation. Arthritis Rheumatol 2014;66:2828–36.
44. Lafyatis R, York M, Marshak-Rothstein A. Antimalarial agents: closing the gate on toll-like receptors? [editorial]. Arthritis Rheum 2006;54:3068–70.
45. Costedoat-Chalumeau N, Le Guern V, Piette JC. Routine hydroxychloroquine blood concentration measurement in systemic lupus erythematosus reaches adulthood. J Rheumatol 2015;42:1997–9.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 717–721
DOI 10.1002/acr.24360

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Soluble Flt-1, Placental Growth Factor, and Vascular Endothelial Growth Factor Serum Levels to Differentiate Between Active Lupus Nephritis During Pregnancy and Preeclampsia

Guilherme R. de Jesús,[1] Marcela I. Lacerda,[1] Bruna C. Rodrigues,[1] Flávia C. dos Santos,[1] Adriana P. do Nascimento,[1] Luís Cristóvão Porto,[1] Nilson R. de Jesús,[1] Roger A. Levy,[2] and Evandro M. Klumb[1]

Objective. To evaluate mean serum levels of vascular endothelial growth factor (VEGF), placental growth factor (PlGF), and soluble Flt-1 (sFlt-1) in pregnant patients with systemic lupus erythematosus (SLE) with inactive disease, active lupus nephritis, and preeclampsia for differential diagnosis between these conditions.

Methods. Pregnant women with SLE, with singleton pregnancies and no other autoimmune diseases, were classified according to disease activity (inactive SLE and active lupus nephritis) and the presence of preeclampsia. Serum samples were collected within 3 weeks of delivery and frozen for subsequent blinded analysis through the enzyme-linked immunosorbent assay method.

Results. A total of 71 women were included, with 41 classified as having inactive SLE (group 1; Systemic Lupus Erythematosus Pregnancy Disease Activity Index [SLEPDAI] score <4), 15 with a diagnosis of active lupus nephritis (group 2, SLEPDAI score ≥4, including renal criteria), and 15 with a diagnosis of preeclampsia (group 3). Patients in group 3 had higher mean levels of sFlt-1 and lower mean levels of PlGF compared to groups 1 and 2, both findings with statistical significance. The sFlt-1:PlGF ratio was also significantly higher in patients with preeclampsia, while mean VEGF levels were higher in pregnant woman with active lupus nephritis compared to patients with preeclampsia or inactive SLE.

Conclusion. Evaluation of serum VEGF, PlGF, and sFlt-1 levels can differentiate between preeclampsia, inactive SLE, and active lupus nephritis during pregnancy.

INTRODUCTION

Pregnancy in patients with systemic lupus erythematosus (SLE) is associated with significant morbidity and mortality compared to the general population, including an increased risk of disease activity, hypertension, pregnancy loss, preterm delivery, intrauterine growth restriction, and preeclampsia (1,2). Active lupus nephritis (LN) during pregnancy makes the differential diagnosis with preeclampsia troublesome in clinical practice, because both conditions can present with hypertension, edema, proteinuria, low platelet count, and worsening of renal function. The classical biomarkers, such as anti–double-stranded DNA (anti-dsDNA) and complement plasma levels, are not always able to differentiate the 2 conditions, which require different treatment approaches (3).

Evaluation of serum angiogenic factors, like vascular endothelial growth factor (VEGF) and placental growth factor (PlGF), and antiangiogenic factors, such as soluble Flt-1 (sFlt-1), has been proposed to help the differentiation between these 2 conditions (4,5). VEGF and PlGF are necessary for physiologic development of pregnancy, because they promote angiogenesis and induce the vasodilatory prostacyclins and nitric oxide in endothelial cells, resulting in reduced vascular tone and blood pressure. They have also been related to glomerular healing and accelerated renal recovery in animal models (6).

Presented in part at the Annual Scientific Meeting of the American College of Rheumatology, Atlanta, Georgia, November 2019, and published in abstract form: De Jesús G, Lacerda M, Rodrigues B, dos Santos F, Nascimento A, Porto L, et al. sFlt-1, PlGF, and VEGF in the differential diagnosis between active SLE nephritis during pregnancy and preeclampsia. Arthritis Rheumatol 2019;71 Suppl 10.

Supported by the Centro de Estudos em Reumatologia Pedro Ernesto, from the Universidade do Estado do Rio de Janeiro.

[1]Guilherme R. de Jesús, MD, PhD, Marcela I. Lacerda, MD, Bruna C. Rodrigues, MD, Flávia C. dos Santos, MD, PhD, Adriana P. do Nascimento, PhD, Luís Cristóvão Porto, MD, PhD, Nilson R. de Jesús, MD, Evandro M. Klumb, MD, PhD: Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil; [2]Roger A. Levy, MD, PhD: GlaxoSmithKline, Upper Providence, Pennsylvania.

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

Address correspondence to Guilherme R. de Jesús, MD, PhD, Departamento de Obstetrícia, Hospital Universitário Pedro Ernesto, Universidade do Estado do Rio de Janeiro, Avenida Professor Manoel de Abreu, 500, 1° andar, 20550-170, Rio de Janeiro, Brazil. Email: guilhermedejesus@gmail.com.

Submitted for publication March 19, 2020; accepted in revised form June 16, 2020.

SIGNIFICANCE & INNOVATIONS

- Pregnant patients with lupus and preeclampsia have increased serum levels of soluble Flt-1 and lower levels of placental growth factor when compared to inactive systemic lupus erythematosus (SLE) and active lupus nephritis.
- Levels of vascular endothelial growth factor were higher in patients with active lupus nephritis compared to inactive SLE and preeclampsia.
- Evaluation of angiogenic and antiangiogenic factors can be a new tool to differentiate preeclampsia from lupus nephritis during pregnancy in clinical practice.

On the other hand, sFlt-1 is a splice variant of VEGF endothelial receptor Flt-1, but it lacks transmembrane and cytoplasmatic domains. It works as a strong antagonist of VEGF and PlGF and induces hypertension, endothelial dysfunction, and nephrotic proteinuria when administered to animal models (6). In humans, previous publications have demonstrated an angiogenic imbalance (increased serum sFlt-1 with low PlGF and VEGF) in patients who develop preeclampsia (7), including patients with lupus (8,9).

Although promising for the differential diagnosis with preeclampsia, few data are available regarding the behavior of these cytokines among pregnant patients with SLE who present quiescent or active LN, and most data come from case reports (4). The objective of this study was to evaluate serum levels of VEGF, PlGF, and sFlt-1 in pregnant women with SLE with inactive disease, active LN, and preeclampsia.

PATIENTS AND METHODS

This was a cross-sectional study of patients with SLE, diagnosed according to American College of Rheumatology (ACR) criteria (10), with singleton pregnancies followed at a high-risk prenatal care clinic in a tertiary health unit (Universidade do Estado do Rio de Janeiro, Brazil). Patients were prospectively included according to regular prenatal follow-up visits and were accompanied by obstetricians and rheumatologists experienced in evaluating pregnant patients with SLE. Women with other autoimmune diseases, including antiphospholipid syndrome, and with end-stage renal disease, were excluded because these conditions could influence the results of the tests.

Disease activity was established according to the Systemic Lupus Erythematosus Pregnancy Disease Activity Index (SLEPDAI) (11), and the occurrence of preeclampsia followed the American College of Obstetricians and Gynecologists proposed criteria (12). Clinical data were obtained by physical examination and medical chart reviews, with an initial classification of activity or preeclampsia performed at the time of blood collection by both rheumatologist and obstetrician in all cases. After delivery, all initial diagnostic results were retrospectively reviewed to ensure that there was no misdiagnosis at first impression. Information about SLE characteristics (clinical and laboratory manifestations before pregnancy, medications, Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index [SDI]) and outcomes of current pregnancy (gestational age at delivery, birth weight, Apgar score) were also recorded.

Blood samples were collected through venous puncture, within 3 weeks of delivery, during regular prenatal visits for patients with inactive SLE, or if disease activity or preeclampsia was suspected during the third trimester. Serum sample aliquots were frozen at –80°C for subsequent blinded analysis by enzyme-linked immunosorbent assay kits (PlGF: DRG Instruments; sFlt-1 and VEGF: R&D Systems) according to manufacturers' recommendations. All samples were run in duplicate in 2 batches in the same laboratory, and average rates were reported. Manufacturers' controls were used, and the assay was repeated if there was a variation >10% between duplicates.

The results were compared between groups, using Pearson's chi-square test, the Mann-Whitney U test, and analysis of variance as appropriate. A receiver operating characteristic curve was created for the sFlt-1:PlGF ratio and VEGF to determine cutoff values and analyze the accuracy of the tests. This study was approved by the Universidade do Estado do Rio de Janeiro Institutional Review Board.

RESULTS

A total of 74 women were prospectively added according to the inclusion criteria. One patient with a diagnosis of secondary antiphospholipid syndrome and 2 patients who presented with nonrenal SLE activity were excluded from this analysis. For the purpose of this study, only patients with active renal manifestations were included in the active SLE group.

A total of 41 patients had inactive or mildly active SLE (group 1: SLEPDAI score <4), 15 had active LN (group 2: SLEPDAI score ≥4, including renal criteria), and 15 had preeclampsia (group 3) at the time of blood collection. Among patients of group 1, 36.5% had a history of LN but had no clinical or laboratory manifestations of active renal disease. Patients with active LN and preeclampsia had higher proteinuria and serum creatinine levels compared to patients with inactive SLE, although there were no cases of severe renal disfunction (all patients had creatinine <1.2 mg/dl). Patients with active LN more frequently had positive anti-dsDNA than the other 2 groups (11 of 15 versus 10 of 56; $P < 0.001$), but there was no significant difference in the number of patients with hypocomplementemia (6 of 15 versus 13 of 56; $P = 0.19$).

Two patients who had been initially classified as having active LN were reclassified as having preeclampsia due to subtle normalization of hypertension and proteinuria a few days after delivery without considerable change of medications. Demographics and clinical characteristics of included patients are described in Table 1.

Table 1. Demographic and clinical characteristics and gestational results of patients with inactive SLE, active LN, and SLE with preeclampsia*

	Inactive SLE (n = 41)	Active LN (n = 15)	SLE with preeclampsia (n = 15)	*P*†
Age at inclusion, years	27.2 ± 6	29.4 ± 4.4	30.1 ± 5.8	0.17
Gestational age at blood collection, weeks	36.8 ± 1.7	33.7 ± 4.2	34.5 ± 2.5	<0.001‡
Gestational age at delivery, weeks	38.7 ± 1.9	35.7 ± 3.8	35.8 ± 2.5	<0.001‡
SDI	0.3 ± 0.6	0.3 ± 0.5	0.1 ± 0.4	0.45
History of LN, no. (%)	15 (35.7)	15 (100)	8 (53.3)	NA
Birth weight	2,976.8 ± 532.4	2,448.4 ± 759.2	2,174.3 ± 834.6	<0.001‡
Serum creatinine at inclusion, mg/dl	0.57 ± 0.14	0.76 ± 0.22	0.67 ± 0.21	<0.001‡
Proteinuria at inclusion, grams/24 hours	0.2 + 0.1	2.2 + 1.5	1.7 + 1.9	<0.001‡
5th-minute Apgar score	9.0 ± 0.6	8.7 ± 2.5	8.7 ± 0.6	0.60
Small for gestational age newborn, no. (%)§	5 (12.1)	4 (26.6)	9 (60)	NA

* Values are the mean ± SD unless indicated otherwise. LN = lupus nephritis; NA = not applicable; SLE = systemic lupus erythematosus; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index.
† *P* values by analysis of variance.
‡ Statistically significant.
§ Defined as birth weight below 10th percentile.

Blood sample collections were performed at a mean gestational age of 36.8, 33.7, and 34.5 weeks, respectively, in groups 1, 2, and 3, and delivery occurred at 38.7, 35.7, and 35.8 weeks, respectively. The mean gestational age of blood collection and delivery was significantly higher in patients with inactive SLE ($P < 0.001$ for both). The mean SDI score was similar in all groups, as were 5th-minute Apgar scores. The mean birth weight was considerably lower in patients with active LN and even more in those with preeclampsia ($P < 0.001$).

Medications used during pregnancy are described in Table 2. Patients in group 2 (active LN) used prednisone and azathioprine more frequently compared to the other groups. Only 2 patients in the study were not using hydroxychloroquine and >80% were using low-dose aspirin.

Mean levels of VEGF, PlGF, and sFlt-1 of each group are reported on Table 3 and Supplementary Figures 1–4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24360/abstract. Patients with SLE and preeclampsia had significantly lower mean serum levels of PlGF, while sFlt-1 was significantly higher in patients with preeclampsia compared to pregnant patients with inactive SLE or active LN. The sFlt-1:PlGF ratio was also significantly higher in patients of group 3 (preeclampsia) compared to other patients with SLE (groups 1 and 2). VEGF was higher in patients with LN compared to inactive SLE and SLE with preeclampsia, while PlGF and sFlt-1 were similar when both groups with SLE without preeclampsia were compared.

The positive predictive value (PPV) for preeclampsia with a sFlt-1:PlGF ratio of 44 was 56.5% (13 of 23), with a negative predictive value (NPV) of 95.8% (46 of 48). The sensitivity was 86.6% (13 of 15) and specificity was 80.3% (45 of 56). For a VEGF cutoff of 10.4 pg/ml, the sensitivity for active LN was 53.3% (8 of 15) and specificity of 87.5% (49 of 56), with a PPV of 53.3% (8 of 15) and an NPV of 87.5% (49 of 56).

DISCUSSION

The differential diagnosis between active LN and preeclampsia in patients with SLE is crucial for better outcomes, because the first condition is treated with immunosuppressive therapy and the latter has considerable improvement of manifestations after delivery (2). This study provides new insights for this conundrum, because serum levels of sFlt-1 and the sFlt-1:PlGF ratio were higher in patients with preeclampsia, while PlGF levels were significantly lower compared to pregnant patients with SLE without this obstetric morbidity.

Table 2. Medications used by included patients with inactive SLE, active LN, and SLE with preeclampsia*

	Inactive SLE (n = 41)	Active LN (n = 15)	SLE with preeclampsia (n = 15)	*P*†
Prednisone	24 (57.1); 9.4 ± 6.9	14 (93.3); 27.9 ± 23.7	8 (53.3); 10.3 ± 6.6	<0.0001‡
Hydroxychloroquine	40 (95.2); 385.0 ± 53.3	15 (100); 400.0 ± 0	15 (100); 386.7 ± 51.6	0.56
Azathioprine	18 (42.8); 111.1 ± 36.6	13 (86.6); 130.8 ± 38.4	6 (40.0); 100.0 ± 31.6	0.06
Antihypertensive, no. (%)	1 (2.3)	4 (26.6)	4 (26.6)	NA
Low-dose aspirin, no. (%)	34 (80.9)	14 (93.3)	12 (80)	NA

* Values are the number (%); mean ± SD unless indicated otherwise. LN = lupus nephritis; SLE = systemic lupus erythematosus; NA = not applicable.
† *P* values by analysis of variance.
‡ Statistically significant.

Table 3. Mean values of VEGF, PlGF, sFlt-1, and sFlt-1:PlGF ratio for patients with inactive SLE, active LN, and SLE with preeclampsia*

	Inactive SLE (n = 41)	Active LN (n = 15)	SLE with preeclampsia (n = 15)	*P*†		
				Inactive SLE × LN	Inactive SLE × PE	LN × PE
VEGF, pg/ml	5.6 ± 7	12.3 ± 10.1	4.1 ± 5	0.006‡	0.45	0.009‡
PlGF, pg/ml	189.8 ± 146.1	198.7 ± 134.8	61.4 ± 127.3	0.83	0.003‡	0.007‡
sFlt-1, pg/ml	1,804.2 ± 668.3	1,832.1 ± 760.9	2,517.0 ± 431.9	0.90	<0.001‡	0.006‡
sFlt-1:PlGF ratio	22.9 ± 25.1	23.3 ± 35.5	781.1 ± 1,211.3	0.96	0.02‡	0.02‡

* Values are the mean ± SD unless indicated otherwise. LN = lupus nephritis; PE = preeclampsia; PlGF = placental growth factor; sFlt-1 = soluble Flt-1; SLE = systemic lupus erythematosus; VEGF = vascular endothelial growth factor.
† *P* values by Mann-Whitney U test.
‡ Statistically significant.

Levine et al described a case–control study of healthy nulliparous women, with increased serum levels of sFlt-1 in patients with preeclampsia compared to controls, while PlGF and VEGF were significantly lower. The authors suggest that the physiologic proangiogenic state of the second trimester (high PlGF and low sFlt-1) is converted to an antiangiogenic state during late pregnancy, with higher sFlt-1 and lower PlGF to control placental vascular growth. Patients with preeclampsia would have this conversion at an earlier stage and more abruptly, with an exaggeration of the normal process of placental growth and function (7).

Two publications have validated, in a prospective fashion, the use of angiogenic and antiangiogenic factors in patients with SLE and preeclampsia, demonstrating the same pattern as healthy women and also the possibility to predict patients who will develop this obstetric condition (8,9). Nonetheless, the researchers did not evaluate the levels of those factors in patients with active renal SLE, precluding the use for differential diagnosis between LN and preeclampsia and did not include VEGF in their analysis.

Angiogenic factor imbalance can also be used as predictor of adverse obstetric outcomes, such as preeclampsia, fetal/neonatal death, fetal growth restriction, and indicated preterm delivery. In the Predictors of Pregnancy Outcome: Biomarkers in APL Syndrome and SLE study, sFlt-1 and PlGF levels between 12 and 15 weeks were significantly altered in patients with SLE with severe adverse obstetric outcomes (13). Another publication identified a higher sFlt-1:PlGF ratio between 24 and 29 weeks in women with SLE who developed adverse obstetric outcomes compared to uncomplicated pregnancies and 5 patients with SLE flare, none with renal activity (14).

We have previously reported that, although nonpregnant patients with a history of LN had increased sFlt-1 levels compared to controls, PlGF was also higher in these patients, which is a different pattern compared to preeclampsia (15). The current study confirms the potential use of these angiogenic and antiangiogenic factors for a differential diagnosis of preeclampsia and LN, also demonstrating that serum VEGF is higher in patients with active LN compared to those with inactive lupus and preeclampsia. This result is in consonance with previous publications that demonstrated increased serum VEGF in nonpregnant patients with SLE with active disease compared to SLE controls (16). Some authors have suggested that a low sFlt-1:PlGF ratio may rule out preeclampsia for a few weeks in patients without SLE, based on the very high reported NPV of this ratio (17), a result that was also found in this study.

Considering the fact that that hydroxychloroquine treatment for pregnant patients with SLE may reduce the incidence of preeclampsia (18), in vitro studies have investigated the effect of this medication on human placental explants from term gestations exposed to hypoxic injury. A protective effect on endothelial function has been described, but there was no influence on sFlt-1 and soluble endoglin release (19). In a similar study, azathioprine significantly increased sFlt-1 and PlGF expressions on term placenta explants after 24 hours of incubation when compared to controls (20). However, there are no studies evaluating whether these drugs affect angiogenic and antiangiogenic levels in pregnant women.

Our results may help physicians with prenatal care of pregnant patients with SLE, considering that differential diagnosis between LN and preeclampsia can be challenging, and sometimes impossible, using currently available methods. Serum complement levels, usually low in patients with proliferative glomerulonephritis, may be normal due to physiologic changes of pregnancy. Dysmorphic hematuria is not always present in nephritis and anti-dsDNA can be persistently positive in some patients with SLE, while serum uric acid is normally elevated in preeclampsia but is not specific for the disease (3).

Two patients had an initial diagnosis of LN at inclusion, but maintenance of abnormalities despite early immunosuppressive treatment, in addition to rapid reversal of hypertension and proteinuria after delivery, switched this diagnosis to preeclampsia. Retrospective blinded analysis of studied factors also indicated a diagnosis of preeclampsia in both patients (high sFlt-1 and low PlGF), suggesting that this information could have changed the initial recommended treatment. Similarly, Hirashima et al published a case report about a woman with an initial diagnosis of preeclampsia who did not reverse proteinuria and hypertension for >30 days after delivery, receiving a final diagnosis of new onset of LN during pregnancy (4). Evaluation of blood samples retrieved before delivery demonstrated normal serum levels of sFlt-1 and a sFlt-1:PlGF ratio, so the authors suggest that if these results had

been available at the time, they could have ruled out preeclampsia and proceeded to appropriate diagnosis and treatment (4).

The small number of patients is a limitation of this study, but the exclusion of patients with SLE activity without LN and women with other autoimmune diseases, especially antiphospholipid syndrome, makes the results more reliable for the intended differential diagnosis. Either way, this is the largest study evaluating these factors for this purpose, led by obstetricians and rheumatologists with considerable experience in prenatal care of lupus patients.

In conclusion, this study demonstrated that pregnant patients with SLE who developed preeclampsia had similar angiogenic and antiangiogenic profile of patients with preeclampsia without SLE, low serum PlGF, and high serum sFlt-1, with a high sFlt-1:PlGF ratio. This pattern differs from patients with inactive SLE or active LN, the latter condition being the main differential diagnosis during gestation of SLE patients. In addition, there is an increase in serum VEGF in patients with active LN, which is not expected in preeclampsia. Evaluation of angiogenic and antiangiogenic factors can be a new tool to differentiate preeclampsia from LN during pregnancy in clinical practice.

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. G. R. de Jesús 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. G. R. de Jesús, N. R. de Jesús, Levy, Klumb.

Acquisition of data. G. R. de Jesús, Lacerda, Rodrigues, dos Santos, do Nascimento, Porto, N. R. de Jesús, Levy.

Analysis and interpretation of data. G. R. de Jesús, Lacerda, N. R. de Jesús, Levy, Klumb.

REFERENCES

1. Lateef A, Petri M. Systemic lupus erythematosus and pregnancy. Rheum Dis Clin North Am 2017;43:215–26.
2. Lightstone L, Hladunewich MA. Lupus nephritis and pregnancy: concerns and management. Semin Nephrol 2017;37:347–53.
3. De Jesus GR, Mendoza-Pinto C, de Jesus NR, Dos Santos FC, Klumb EM, Carrasco MG, et al. Understanding and managing pregnancy in patients with lupus. Autoimmune Dis 2015;2015:943490.
4. Hirashima C, Ogoyama M, Abe M, Shiraishi S, Sugase T, Niki T, et al. Clinical usefulness of serum levels of soluble Fms-like tyrosine kinase 1/placental growth factor ratio to rule out preeclampsia in women with new-onset lupus nephritis during pregnancy. CEN Case Rep 2019;8:95–100.
5. Verdonk K, Visser W, Russcher H, Danser AH, Steegers EA, van den Meiracker AH. Differential diagnosis of preeclampsia: remember the soluble Fms-like tyrosine kinase 1/placental growth factor ratio. Hypertension 2012;60:884–90.
6. Maynard SE, Min JY, Merchan J, Lim KH, Li J, Mondal S, et al. Excess placental soluble Fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. J Clin Invest 2003;111:649–58.
7. Levine RJ, Maynard SE, Qian C, Lim KH, England LJ, Yu KF, et al. Circulating angiogenic factors and the risk of preeclampsia. N Engl J Med 2004;350:672–83.
8. Leaños-Miranda A, Campos-Galicia I, Berumen-Lechuga MG, Molina-Pérez CJ, García-Paleta Y, Isordia-Salas I, et al. Circulating angiogenic factors and the risk of preeclampsia in systemic lupus erythematosus pregnancies. J Rheumatol 2015;42:1141–9.
9. Mayer-Pickel K, Stern C, Eberhard K, Lang U, Obermayer-Pietsch B, Cervar-Zivkovic M. Angiogenic factors in pregnancies of women with antiphospholipid syndrome and systemic lupus erythematosus. J Reprod Immunol 2018;127:19–23.
10. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.
11. Andreoli L, Gerardi MC, Fernandes M, Bortoluzzi A, Bellando-Randone S, Brucato A, et al. Disease activity assessment of rheumatic diseases during pregnancy: a comprehensive review of indices used in clinical studies. Autoimmun Rev 2019;18:164–76.
12. ACOG practice bulletin no. 202: gestational hypertension and preeclampsia. Obstet Gynecol 2019;133:e1–25.
13. Kim MY, Buyon JP, Guerra MM, Rana S, Zhang D, Laskin CA, et al. Angiogenic factor imbalance early in pregnancy predicts adverse outcomes in patients with lupus and antiphospholipid antibodies: results of the PROMISSE study. Am J Obstet Gynecol 2016;214:108.e1–14.
14. Rodríguez-Almaraz ME, Herraiz I, Gómez-Arriaga PI, Vallejo P, Gonzalo-Gil E, Usategui A, et al. The role of angiogenic biomarkers and uterine artery Doppler in pregnant women with systemic lupus erythematosus or antiphospholipid syndrome. Pregnancy Hypertens 2018;11:99–104.
15. De Jesus GR, de Jesus NR, Levy RA, Klumb EM. The use of angiogenic and antiangiogenic factors in the differential diagnosis of pre-eclampsia, antiphospholipid syndrome nephropathy and lupus nephritis. Lupus 2014;23:1299–301.
16. Zhou L, Lu G, Shen L, Wang L, Wang M. Serum levels of three angiogenic factors in systemic lupus erythematosus and their clinical significance. Biomed Res Int 2014;2014:627126.
17. Zeisler H, Llurba E, Chantraine F, Vatish M, Staff AC, Sennström M, et al. Predictive value of the sFlt-1:PlGF ratio in women with suspected preeclampsia. N Engl J Med 2016;374:13–22.
18. Seo MR, Chae J, Kim YM, Cha HS, Choi SJ, Oh S, et al. Hydroxychloroquine treatment during pregnancy in lupus patients is associated with lower risk of preeclampsia. Lupus 2019;28:722–30.
19. Rahman RA, Murthi P, Singh H, Gurungsinghe S, Leaw B, Mockler JC, et al. Hydroxychloroquine mitigates the production of 8-isoprostane and improves vascular dysfunction: implications for treating preeclampsia. Int J Mol Sci 2020;21.
20. Gomes SZ, Araujo F, Bandeira CL, Oliveira LG, Hoshida MS, Zugaib M, et al. The impact of immunosuppressive drugs on human placental explants. Reprod Sci 2019;26:1225–34.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 722–731
DOI 10.1002/acr.24149

American College of Rheumatology
Empowering Rheumatology Professionals

Contribution of Sex and Autoantibodies to Microangiopathy Assessed by Nailfold Videocapillaroscopy in Systemic Sclerosis: A Systematic Review of the Literature

Nina M. van Leeuwen,[1] Jacopo Ciaffi,[2] Jan W. Schoones,[1] Tom W. J. Huizinga,[1] and Jeska K. de Vries-Bouwstra[1]

Objective. Microangiopathy and dysregulation of the immune system play important roles in the pathogenesis of systemic sclerosis (SSc). Factors that trigger vascular injury in SSc have not been elucidated so far. We undertook this study to evaluate whether sex or expression of specific antinuclear autoantibodies might associate with the degree of microangiopathy through performance of a systematic review that summarizes what is known about these associations.

Methods. A standardized search of PubMed, Embase, Web of Science, and the Cochrane Library were performed to identify studies that described autoantibodies in SSc patients and microangiopathy and, for the second search, those that described sex and microangiopathy.

Results. We included 11 studies that described the relationship between SSc-specific autoantibodies and microangiopathy and 6 studies that reported on the association between sex and microangiopathy. Contradictory results were found on the association between SSc-specific autoantibodies and microangiopathy, and no association was found between sex and microangiopathy based on the current literature.

Conclusion. Based on this review of the literature, we can conclude that sex does not seem to influence degree of microangiopathy in SSc, while results on association between SSc-specific autoantibodies and degree of microangiopathy were inconclusive.

INTRODUCTION

Systemic sclerosis (SSc) is characterized by a triad of microvascular damage, dysregulation of innate and adaptive immunity, and generalized fibrosis that can affect skin and internal organs (1). In SSc, the most frequent symptom of microvascular damage is Raynaud's phenomenon (RP), which is present in up to 96% of patients and often represents the earliest manifestation of the disease. Current concepts indicate that microangiopathy is a key factor in early pathogenesis of SSc. In RP that is evolving to definite SSc, presence of microvascular damage and SSc-specific autoantibodies indicate a very high probability of developing SSc (2). The frequency of progression is higher with both the presence of SSc autoantibodies and microvascular damage (79.5%) than with the presence of 1 of these predictors (32.2%) (3). In addition to its diagnostic value, the degree of microangiopathy is also a valuable prognostic marker in SSc patients, as it contributes to the prediction of future organ complications (3–5). The SSc-specific autoantibodies are associated with specific clinical characteristics and therefore are of additional prognostic value. Anticentromere antibodies (ACAs) are associated with a decreased risk of lung (odds ratio [OR] 0.12) and heart (OR 0.39) involvement, while patients who are anti–topoisomerase I antibody (ATA) positive have an increased risk for these complications (OR 6.66 and OR 2.12, respectively) (6,7). Strikingly, the degree of microangiopathy was comparable between ACA+ and ATA+ patients (late SSc pattern; ACA 33%, ATA 25%), which suggests that the presence of a specific antinuclear antibody (ANA) is independent of the development of microangiopathy.

In some studies, however, an association between microvascular damage and autoantibodies has been described (8). ANAs, found in 95% of patients with SSc, have been mentioned as 1 of the possible triggers for vascular injury by causing acceleration of vascular endothelial cell senescence and therefore inducing

Dr. Ciaffi's work was supported by the European Alliance of Associations for Rheumatology (bursary awarded for scientific training).

[1]Nina M. van Leeuwen, MD, Jan W. Schoones, drs., MsC, Tom W. J. Huizinga, MD, PhD, Jeska K. de Vries-Bouwstra, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands; [2]Jacopo Ciaffi, MD: Leiden University Medical Center, Leiden, The Netherlands, Azienda Policlinico of Modena, University of Modena and Reggio Emilia, Modena, Italy, and IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy.

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

Address correspondence to Nina M. van Leeuwen, MD, Department of Rheumatology Leiden University Medical Center, C1-R, PO Box 9600, 2300 RC, Leiden, The Netherlands. Email: n.m.van_leeuwen@lumc.nl.

Submitted for publication August 7, 2019; accepted in revised form January 14, 2020.

SIGNIFICANCE & INNOVATIONS

- Degree of microangiopathy is used as a diagnostic and prognostic tool in systemic sclerosis (SSc).
- Factors that influence microangiopathy are not completely elucidated.
- Based on the current literature in SSc, there is no association between sex and degree of microangiopathy, but for SSc-specific autoantibodies, the results are contradictory, advocating further evaluation.

RP (9,10). Other studies suggest that autoantibody production occurs secondary to vasculopathy, and as such these autoantibodies should be viewed as a bystander in disease pathogenesis (7,11,12).

Vasculopathy in SSc involves all layers of the peripheral blood vessels and is caused by a dysfunction of the endothelium, resulting in an imbalance of vasoactive factors. In particular, endothelin 1 plays a prominent role in the regulation of vascular tone through its receptors. RP induces prolonged ischemia-reperfusion injury, which may cause persistent endothelial activation, resulting in apoptosis, microvascular damage, and other toxic stimuli. Recent insights showed that impaired functioning of endothelial progenitor cells could be involved in angiogenic response and in the pathogenesis of SSc. Microvascular tone alterations and cell apoptosis trigger the opening of intercellular junctions in the endothelial barrier. This loss of integrity favors further migration and homing of inflammatory cells, inducing increased microvascular permeability and progressive vascular leak (13). Infective stimuli, environmental exposures, sex, and endocrine disturbances have all been proposed as contributors to microangiopathy (14,15).

In SSc, there is a marked sex imbalance, with higher prevalence of the disease in women than in men (4:1). Also, distribution of ANA is disbalanced, with women showing more frequently ACA positivity and men showing more frequent ATA positivity. In general, disease course is more severe in men, resulting in lower survival rates (45% versus 23% after 10 years) (16–20). The most frequent disease-related causes of death also differ between men and women, with interstitial lung disease in men and pulmonary hypertension (PH) in women (21). The higher incidence of PH in women and the fact that unopposed estrogen replacement therapy has been associated with increased RP suggest a contribution of hormonal factors to microangiopathic manifestations (22); however, little information is known about the relationship between sex and microangiopathy in SSc.

As microvascular damage is one of the hallmarks of SSc, different imaging techniques have been applied to evaluate structural and functional abnormalities of the finger microcirculation in patients with SSc (23–26) (see Supplementary Appendix A and Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24149/abstract). However, nailfold videocapillaroscopy (NVC) is considered the most reliable tool to distinguish between primary and secondary RP. NVC is widely applied and provides the opportunity to directly visualize the evolving obliterative microangiopathy and nailfold capillary abnormalities characteristic of SSc, that have been classified as scleroderma pattern (27).

Given the role of microangiopathy in the pathogenesis of SSc, insights in the factors responsible for microvascular damage could contribute to our understanding of disease pathophysiology. Therefore, we decided to evaluate and summarize in this comprehensive review what is known about the association between the expression of specific autoantibodies and microangiopathy, and between sex and microangiopathy in SSc.

MATERIALS AND METHODS

Literature search. A systematic literature search was performed (JWS), including studies published before June 17, 2019. The databases used were Medline (via PubMed), Web of Science, the Cochrane Library, and Embase. No restrictions on date were applied, and only manuscripts published in English or Dutch were selected. The search strategy intended to include all relevant reports describing adult patients with SSc, in which microangiopathy of the hand was evaluated and where association with SSc-specific autoantibodies was assessed. A second systematic literature search performed on the same day intended to include all relevant reports describing adult patients with SSc, in which microangiopathy of the hand was evaluated and a comparison between male and female patients was described (for search strategies, see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24149/abstract).

Two reviewers (NMvL and JC) independently screened the titles of retrieved articles and, in the case that 1 or both reviewers identified a publication as possibly relevant, the study proceeded to abstract screening. In case of discrepancies in agreement, abstracts were reviewed by a third investigator (JKdV-B). Full text reading was performed for the selected abstracts (NMvL and JC).

Screening process and study selection criteria. For the review on autoantibodies and microangiopathy, the following criteria were applied: 1) adult participants (ages >18 years) with a clinical diagnosis of SSc; 2) fulfillment of either American College of Rheumatology (ACR) 2013, ACR 1980, or LeRoy and Medsger criteria (28,29); 3) report on prevalence of SSc-related autoantibodies, including at least ATAs or ACAs, and additionally, anti–RNA polymerase III (anti–RNAP III), anti–RNAP I, antifibrillarin, anti-PM/Scl, or anti-Th/To antibodies; and 4) assessment of microangiopathy using ≥1 imaging modality, including NVC, laser dermoscopy, Doppler confocal microscopy, laser speckle contrast analysis (LASCA)/video image analysis, and photomicroscopy.

For the review on sex and microangiopathy, the following criteria were applied: 1) adult participants (ages >18 years) with a

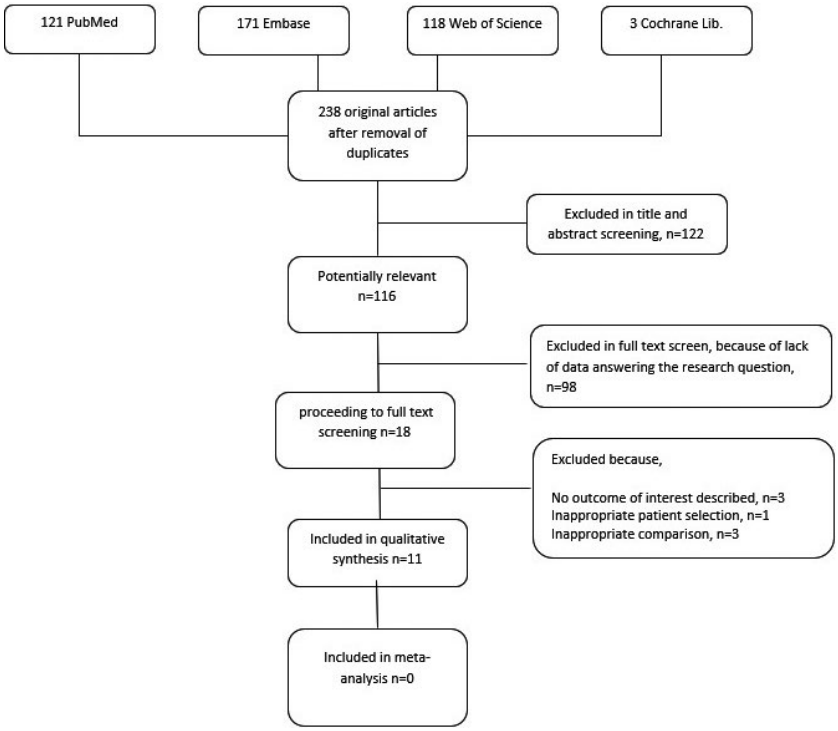


Figure 1. Flow chart of the association of autoantibodies and microangiopathy.

clinical diagnosis of SSc; 2) fulfillment of either ACR 2013, ACR 1980, or LeRoy and Medsger criteria (28,29); 3) report on the comparison between female and male patients, with at least n = 3 and 10% male patients included in the study; and 4) assessment of microangiopathy using ≥1 imaging modality, including NVC, laser dermoscopy, Doppler confocal microscopy, LASCA/video image analysis, and photomicroscopy. Exclusion criteria for both search strategies were animal studies, editorials, reviews, letters to the editor, unpublished material, case-reports, and manuscripts written in languages other than English or Dutch.

Quality assessment. The Newcastle-Ottawa scale was used for assessment of quality of case–control studies, whereas the National Institutes of Health quality assessment tool was used for observational cohort studies (30,31). Discrepancies in scoring and implications for interpretation of the findings were discussed (NMvL and JC).

Evaluation of capillaroscopic descriptions throughout the studies. In the literature a variety of definitions are used to describe NVC. In this review, we will report the NVC findings in a standardized way by evaluating the used terminology to describe NVC characteristics per included article. In line with the European League Against Rheumatism recommendations on capillaroscopy, the NVC characteristics can be evaluated quantitatively, qualitatively, or semi-quantitively (32) (see Supplementary Appendix A and Supplementary Table 1, available online at http://onlinelibrary.wiley.com/doi/10.1002/acr.24149/abstract). When available, all these NVC characteristics were extracted throughout the included articles.

RESULTS

Literature search and study description. Figures 1 and 2 show flow charts of the systematic review processes. Eleven studies that demonstrated the association between autoantibodies and microangiopathy (7,8,11,33–40) and 6 studies that demonstrated the association between sex and microangiopathy (33,37,40–43) were included. Three studies addressed both associations (33,37,40). All of the included articles were cohort or case–control studies, but many were limited by small sample sizes. In the majority of the included articles, except for 4 (8,11,42,43), the association of interest was not the primary outcome of the study. Characteristics of all included studies are provided in Table 1. In all, these studies included 4,704 women (83%) and 971 men (17%), with a mean age of 49 years. Subtypes of SSc were specified in all but 1 article (diffuse cutaneous SSc [n = 1,473 (28%)] and limited cutaneous SSc [n = 3,746 (72%)]). Disease duration was defined either as time since onset of RP, as time since onset of first sign or symptom attributable to SSc different from RP, or as time since diagnosis, and ranged between 6 months and 37 years.

Comprehensiveness of reporting. The comprehensiveness of reporting was variable. Although all selected studies used NVC, the parameters to describe microangiopathy and to classify severity of microvascular changes differed between the studies.

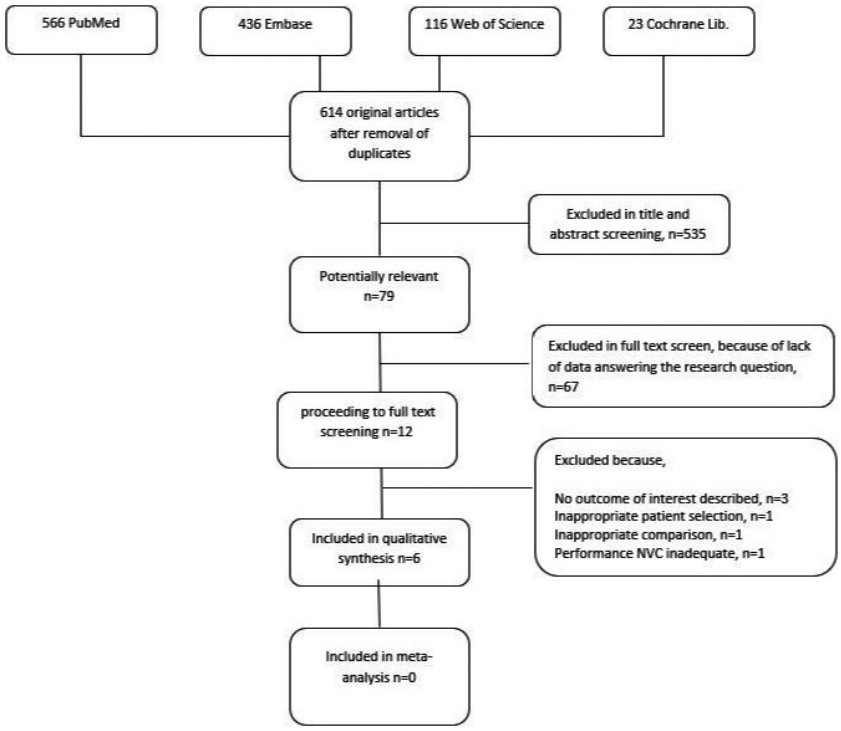


Figure 2. Flow chart of the association of sex and microangiopathy.

Risk of bias. Study quality is summarized in Supplementary Table 1, available online at http://onlinelibrary.wiley.com/doi/10.1002/acr.24149/abstract. Three articles were assessed as high quality (7,8,33), 9 as medium quality (11,33–39,42), and 2 as low quality due to selection bias, performance bias, and incomplete outcome data (40,43). Because of the limited number of studies reporting on the association between autoantibodies, sex, and microangiopathy, we chose to also include medium- and low-quality articles.

Autoantibodies and microangiopathy. A meta-analysis could not be conducted due to heterogeneity of the studies and the use of different outcome measures. In total, 11 studies described the associations between autoantibodies and microangiopathy (Table 2).

Qualitative assessment of NVC. Caramaschi et al performed NVC in 103 SSc patients and the degree of microangiopathy was defined as early, active, or late SSc pattern according to Cutolo et al (quality score good) (32,33). The distribution of ANA, ACA, and ATA positivity did not differ between patients with early, active, or late SSc patterns. De Santis et al investigated 44 SSc patients using NVC to identify early, active, or late SSc patterns (quality score medium) (34). No significant differences in the SSc patterns were found between ACA+ and ATA+ patients. In a study that included 287 SSc patients, ACA, ATA, anti-RNP, anti–RNAP III, anti-fibrillarin, anti-PM/Scl, anti-Th/To, and anti-Ku antibodies were evaluated, and early, active, or late SSc patterns were described on NVC (quality score good) (7). The prevalence of NVC patterns was equally distributed among patients with different specific autoantibodies. On the contrary, Pizzorni et al investigated 33 SSc patients and classified the degree of microangiopathy according to the 3 SSc patterns: early, active, or late (quality score medium) (37). ATA+ patients showed a late SSc pattern ($P = 0.002$) more frequently, while in ACA+ patients early or active SSc patterns were more common ($P = 0.03$). Cutolo et al evaluated NVC patterns and serum autoantibodies in 241 SSc patients (quality score good) (8). NVC was described as early, active, or late SSc pattern. ATA positivity was significantly less frequent in the early SSC patterns (5%) than in the active (25%), or late (24%) SSc patterns.

Presence of ATA was shown to be related to earlier expression of the active and late SSc patterns of microvascular damage. On the other hand, ACA positivity was found more frequently, although not significantly, in the early pattern. The authors concluded that specific autoantibodies do not seem directly linked to the expression of a singular NVC pattern, but that autoantibodies might be related to the rate of progression of microvascular damage. In a study by Ingegnoli et al, data from the European Scleroderma Trials and Research group were used to investigate NVC in 2,754 SSc patients (quality score medium) (38). NVC patterns were described as early, active, or late SSc pattern. Late pattern was present in 47% of ATA+ and in 28% of ACA+ ($P < 0.05$) patients, while early and active patterns were more frequent in ACA+ than in ATA+ patients (44% versus 28%; $P < 0.05$). Significant associations were found between ATA positivity and late SSc pattern, and between ACA positivity and early/active SSc pattern ($P = 0.03$). Sulli et al

Table 1. Baseline characteristics of articles included in the systematic review: association of sex and autoantibodies with the degree of microangiopathy*

Study	Country	Patients, no.	Age, mean years	Sex, female/male	Disease duration, years since diagnosis	SSc type	Methodological framework	Main topic
Markusse et al, 2017 (7)	Netherlands	287	53.9	202/85	3.7 since onset of RP	141 lcSSc/56 dcSSc	Observational cohort, cross-sectional	Evaluate anti-ENA antibodies in SSc and predictive power of combination of autoantibodies and NVC
Cutolo et al, 2004 (8)	Italy	241	57	227/14	5.6/13.7 since onset of RP	148 lcSSc/93 dcSSc	Observational cohort, cross-sectional	Relation NVC pattern autoantibodies and subset cutaneous involvement
Sulli et al, 2013 (11)	Italy, Belgium	42	47	NA	5 since onset of RP	NA	Observational cohort, longitudinal	Correlation between ANA patterns and NVC stage in SSc
Caramaschi et al, 2007 (33)	Italy	103	54.3	91/12	7	68 lcSSc/35 dcSSc	Observation cohort, cross- sectional	NVC pattern and clinical characteristics
De Santis et al, 2016 (34)	Italy	44	66	42/2	9	34 lcSSc/10 dcSSc	Observational cohort, cross-sectional	Correlation NVC and clinical SSc phenotype
Fichel et al, 2014 (35)	France	88	54.9	81/7	16.5 since onset of RP	51 lcSSc/15 dcSSc/ 12 noncutaneous	Observational cohort, cross-sectional	Characteristics SSc patients with normal or abnormal NVC
Ghizzoni et al, 2015 (36)	Italy	275	54.9	253/22	36.9	242 lcSSc/33 dcSSc	Observational cohort, longitudinal	Prevalence, evolution of NVC and analysis of characteristics according to capillaroscopic features
Pizzorni et al, 2017 (37)	Italy	33	59	28/5	6.6	30 lcSSc/3 dcSSc	Observational cohort, cross-sectional	Evaluate use of MES assessment with qualitative analysis of NVC and telangiectasia
Ingegnoli et al, 2013 (38)	Italy	2,754	54.9	2,148/606	7.6	1,622 lcSSc/803 dcSSc	Observatinal cohort, cross-sectional	Frequency of NVC patterns and their disease phenotype
Tieu et al, 2018 (39)	Australia	152	43.7	121/31	10.9 since onset of RP	99 lcSSc/30 dcSSc	Observational cohort, longitudinal	Investigate possible utility of NVC in predicting survival
Chandran et al, 1995 (40)	Australia	148	50	44/8	5 since onset of RP	81 lcSSc/13 dcSSc	Observational cohort, cross-sectional	Role of NVC in identification and prognostication
Caramaschi et al, 2009 (41)	Italy	49	52.4	44/5	8	31 lcSSc/18 dcSSc	Observational cohort, longitudinal	NVC changes after iloprost treatment
Freire et al, 2017 (42)	Spain	1,506	45.6	1,341/165	6.4	1,151 lcSSc/355 dcSSc	Observational cohort, longitudinal	Influence sex on survival
Simeon et al, 1996 (43)	Spain	91	52.5	82/9	6 months and 63 years since onset of RP	70 lcSSc/19 dcSSc	Observational cohort, cross-sectional	Relationship disease pattern and sex

* Disease duration was defined differently in the articles, either as time since onset of Raynaud's phenomenon (RP) or non-RP, or time since diagnosis. ANA = antinuclear autoantibody; dcSSc = diffuse cutaneous systemic sclerosis; ENA = extractable nuclear antigen; lcSSc = limited cutaneous systemic sclerosis; MES = microangiopathy evolution score; NA = not available; NVC = nailfold videocapillaroscopy; SSc = systemic sclerosis.

Table 2. Association between autoantibodies and microangiopathy*

Study, type	Patients, no.	Antibodies	NVC assessment	Significance	Conclusion
Markusse et al, 2017 (7); qualitative†	253	ACA, ATA, RNAPIII, RNP, U3 RNP, Pm/Scl	Early; active; late SSc pattern	$P > 0.10$	No significant difference
Cutolo et al, 2004 (8); qualitative	241	ACA, ATA	Early; active; late SSc pattern	$P < 0.01$	ATA+ more frequent in active and late patterns than in early
Sulli et al, 2013 (11); qualitative	42	ACA, ATA	Early; active; late SSc pattern	$P = 0.03$ (OR 8.0 [1.4-47.0])	ATA more often present in late pattern than in early and active
Sulli et al, 2013 (11); semiquantitative	42	ACA, ATA	MES	ANA vs. ACA, $P = 0.09$, ANA vs. ATA, $P = 0.05$	No significant differences
Caramaschi et al, 2007 (33); qualitative	103	ACA, ATA	Early; active; late SSc pattern	Nonsignificant (not specified)	No significant difference
De Santis et al, 2016 (34); qualitative	44	ACA, ATA	Early; active; late SSc pattern	$P < 0.05$	No significant difference
De Santis et al, 2016 (34); quantitative	44	ACA, ATA	Giants, neoangiogenesis, avascular areas, density	$P > 0.05$	No significant differences
Fichel et al, 2014 (35); qualitative	88	ACA, ATA	Normal; SSc pattern	ACA normal/SSc pattern, $P = 0.90$ (OR 0.90 [0.3-2.6]); ATA normal/SSc pattern, $P = 0.34$ (OR 0.50 [0.1-2.6])	No significant difference
Ghizzoni et al, 2015 (36); qualitative	275	ACA, ATA	Normal; SSc pattern	Nonsignificant (not specified)	No significant difference
Pizzorni et al, 2017 (37); qualitative	33	ACA, ATA	Early; active; late SSc pattern	ACA early and active/late, $P = 0.03$; ATA early and active/late, $P = 0.02$	Early-active pattern more often present in ACA patients; late pattern more often present in ATA patients.
Pizzorni et al, 2017 (37); semiquantitative†	33	ACA, ATA	MES	ACA MES <6/>6, $P = 0.72$; ATA MES <6/>6, $P = 0.43$	No significant differences
Ingegnoli et al, 2013 (38); qualitative	2,754	ACA, ATA	Early; active; late SSc pattern	$P < 0.05$	ATA more often present in late pattern than in early and active
Tieu et al, 2018 (39); semiquantitative	152	ACA, ATA, RNP, RNAPIII	Mean capillary damage score; mean capillary dropout score	RNAPIII > capillary damage compared with ACA and RNP ($P < 0.001$); ATA and RNAPIII > dropout compared with ACA (P = unknown)	Difference found between autoantibodies and capillary damage and capillary dropout
Chandran et al, 1995 (40); semiquantitative	52	ACA, ATA, RNP	Moderate loss and enlargement; extreme capillary dropout; class 1 to 5	Not mentioned	ATA+ patients more severe nailfold changes compared to ACA and RNP+

* ACA = anticentromere antibody; ANA = antinuclear antibody; ATA = anti–topoisomerase 1 antibody; MES = microangiopathy evolution score; NVC = nailfold videocapillaroscopy; RNAP III = RNA polymerase III; SSc = systemic sclerosis.
† Same article used 2 techniques for NVC assessment.

found that the prevalence of ATA was significantly higher in patients with the late SSc pattern (n = 42; quality score medium) (11). Fichel et al described the characteristics of 88 SSc patients with normal, nonspecific, or SSc-specific NVC pattern (quality score medium) (35). The frequencies of ANA, ACA ($P = 0.90$), and ATA ($P = 0.34$) positivity were comparable for normal/nonspecific and SSc-specific NVC patterns. This is in line with the results of Ghizzoni et al who described NVC features, demographic, clinical, and serologic manifestations of 275 SSc patients (quality score medium) (36). No differences in the percentage of ACA or ATA positivity were found between patients with SSc patterns compared to patients with normal/nonspecific NVC patterns (ACA: 15.2% versus 14.6%, ATA: 31.8% versus 23.6%; all nonsignificant).

Quantitative assessment of NVC. Besides the SSc-specific NVC patterns, de Santis et al also described the amount of giants, neoangiogenesis, avascular areas, and the capillary density and compared these characteristics between ACA+ and ATA+ patients (34). No significant differences were found.

Semiquantitative assessment of NVC. A study by Tieu et al included 152 SSc patients and investigated capillary dropout during follow-up (quality score medium) (39). Patients with anti–RNAP III had a significantly higher nailfold capillary total damage index compared with ACA+, ATA+, and anti-RNP+ patients. Patients with ATA or anti–RNAP III had greater capillary dropout than patients with ACA, despite a significantly shorter disease duration. Finally, a study by Chandran et al demonstrated that in 52 SSc patients, the ATA+ cases had more severe nailfold changes (quality score low) (40). However, in this study only 4 ATA+ patients were included and 2 of them had severe NVC changes, whereas of the 22 ACA+ patients, 3 had severe NVC changes. Two studies, by Pizzorni et al and by Sulli et al (quality score medium) used the microangiopathy evolution score (MES) to semiquantitatively evaluate the degree of microvascular damage. No significant differences in the MES were found between ACA+ and ATA+ patients (11,37).

In conclusion, weighing the results shown in Table 2, the total number of patients in the studies that found an association between autoantibodies and microangiopathy was 2,364, compared to 742 patients in the studies that did not find an association. This would implicate that specific autoantibodies are associated with the degree of microangiopathy; however, when only high-quality studies were evaluated (7,8,33), an association was found only in 241 patients, while in 390 patients no association between autoantibodies and microangiopathy was noted.

Sex and microangiopathy. In total, 6 studies reported on sex and microangiopathy in patients with SSc (Table 3). A meta-analysis could not be conducted due to the heterogeneity of the studies.

Qualitative assessment. A study by Caramaschi et al included 103 SSc patients (12 men, 91 women) and the microvascular alterations were classified as early, active, and late SSc patterns (quality score good) (33). In this study, no significant differences in NVC patterns were found between male and female patients. Freire et al studied 1,506 SSc patients (165 men, 1,341 women) and assessed microangiopathy with the use of NVC and described the degree of microangiopathy as slow or active pattern (quality score medium) (42). No significant difference in the distribution of patterns was observed between men and women (46% versus 53% for slow pattern and 37% versus 33% for active pattern). Pizzorni et al evaluated 33 patients, including 5 men, and found no difference in the prevalence of SSc patterns in men or women (37). One of 6 studies suggested a possible sex difference regarding microangiopathy (41). In 49 SSc patients who were treated with iloprost and underwent 2 NVC examinations with a 3-year interval, improvement of SSc pattern was found to be associated with male sex ($r = 9.07$, $P = 0.019$).

Quantitative and semiquantitative assessment. None of the included studies evaluated the association between sex and quantitative assessment of microangiopathy. Chandran et al performed a study on prevalence, subset characteristics, and NVC patterns of SSc patients in South Australia (quality score low) (40). The study included 44 men and 8 women, and an equal proportion of men and women had severe capillary changes of class IV (moderate loss of capillaries) and V (extreme capillary dropout). Simeon et al evaluated 91 SSc patients, of which 9 were men (quality score low) (43). The NVC patterns were described using capillary loss and megacapillaries as parameters. No significant NVC differences were found between male and female patients. In line with these results, Pizzorni et al compared

Table 3. Association between sex and microangiopathy*

Study, type	Patients, no.	Sex, female/male	NVC assessment	Significance	Conclusion
Caramaschi et al, 2007 (33); qualitative	103	91/12	Early; active; SSc pattern	Nonsignificant (not specified)	No significant difference
Pizzorni et al, 2017 (37); qualitative†	33	28/5	Early; active; late SSc pattern	$P = 0.623$	No significant difference
Pizzorni et al, 2017 (37); semiquantitative†	33	28/5	MES 0–9, <6 or >6 dichotomized	$P = 0.625$	No significant difference
Chandran et al, 1995 (40); semiquantitative	52	44/8	Moderate loss and enlargement; extreme capillary dropout; class 1 to 5	Not mentioned	No significant difference
Caramaschi et al, 2009 (41); qualitative	49	44/5	Early; active; late SSc pattern	$P < 0.05$	Improvement of NVC associated with male sex
Freire et al, 2017 (42); qualitative	1,506	1,341/165	Slow (giants and minimal loss) or active pattern (capillary loss and nonvascularization)	$P = 0.126$ (slow pattern male/female); $P = 0.420$ (active pattern male/female)	No significant difference
Simeon et al, 1996 (43); semiquantitative	91	82/9	Capillary loss and megacapillaries	$P = 0.71$ (capillary loss); $P = 1.00$ (megacapillaries)	No significant difference

* MES = microangiopathy evolution score; NVC = nailfold videocapillaroscopy; SSc = systemic sclerosis.
† Same article used 2 techniques for NVC assessment.

MES between men and women, and no significant difference was found (37).

In conclusion, of the 6 included articles, 5 studies including 1,614 women and 204 men did not show an association between sex and microangiopathy. The only study showing a significant difference included 44 women and 5 men and, importantly, male patients were more often treated with cyclophosphamide, but a multivariate analysis to identify the contribution of sex corrected for the prescribed treatment was not performed (41).

DISCUSSION

Microangiopathy can be secondary to different causes. Research in different fields shows that many factors can affect microangiopathy, including biological, environmental, and socioeconomic factors (44,45). In addition, sex-specific factors have been postulated as men and women develop different types of ischemic heart disease with different pathophysiologic background (3,4). Atherosclerosis is more common in men, while in women vasoreactivity prevails, characterized by spasm and endothelial alterations. Microvascular dysfunction with perfusion problems seems to be present more often in women with cardiovascular disease, and takotsubo cardiomyopathy, heart failure, and stroke are more common in women (46,47).

Similarly, it has been recognized that there are clinical differences between female and male patients with systemic autoimmune rheumatic diseases in which microangiopathy plays a role, such as systemic lupus erythematosus (SLE) and SSc (48). SLE is rare in men, and men with SLE are more likely to experience cardiovascular complications and myocardial infarction and are less likely to have dermatologic manifestations (48). Nevertheless, it remains unknown why SLE in men differs substantially from SLE in women.

Although there is a growing interest, the exact interplay between autoantibodies and microangiopathy in autoimmune diseases remains to be elucidated. In SLE, a difference in autoantibody prevalence has been suggested between men and women. Anticardiolipin antibodies, anti–double-stranded DNA antibodies, and lupus anticoagulant were found to be more prevalent in men in a few studies (49). Some studies showed that in lupus nephritis, antiphospholipid antibodies and lupus anticoagulant were more frequently observed in patients with thrombotic microangiopathy of the kidney. Additionally, among the autoantibodies mainly implicated in neuropsychiatric (NP) SLE, anti-β_2-glycoprotein I antibodies are preferentially involved in focal NP events that are a consequence of noninflammatory microangiopathy; otherwise, anti–ribosomal P protein antibodies and anti–*N*-methyl-D-aspartate receptor antibodies might cause diffuse NP events (49). In dermatomyositis, anti-MDa5 autoantibodies have a strong correlation with vasculopathy (50). Irrespective of these specific cases, little information is available on the association between sex or autoantibodies and microangiopathy in connective tissue diseases, both for SSc and for other systemic autoimmune diseases.

As the assessment of microangiopathy has an established diagnostic and prognostic role in SSc patients (51), we value possible factors that could influence microangiopathy as relevant. In this review of the literature, we focused on the influence of sex and autoantibodies on microangiopathy in SSc patients. We can conclude that sex does not associate with degree of microangiopathy in SSc, while the results on association between specific autoantibodies and degree of microangiopathy were inconclusive. When summarizing the findings of the positive studies for autoantibodies and microangiopathy, presence of ATA might be associated with more severe microangiopathy as reflected by a late pattern. Indeed, both more severe damage and presence of ATA associate with more severe disease in SSc. However, the degree of microangiopathy can change over time and possible confounders such as age, disease duration, comorbidities, or medications were not taken into account in any of the included studies. When evaluating the high-quality studies only, no clear association between ATA and more severe microangiopathy was shown. However, even in these studies the results were not adjusted for confounders. Therefore, we believe that further prospective controlled studies are needed to better explore the association between presence of specific antibodies and the degree of microangiopathy.

Regarding sex and microangiopathy, no clear association was found in the included articles. However, only 6 studies were retrieved and 2 evaluated sex differences as primary outcome (42,43). Also, a relatively limited number of men was included in the studies. Although several studies focused on sex differences in SSc, a possible difference between males and females in the degree of microangiopathy was disregarded in most studies. To account for the sex gap and disease dissimilarities in SSc, a role of sex hormones has been proposed. Estrogens act as enhancers of the immune system and of cell proliferation, as also demonstrated in cultures of cells harvested from skin biopsies of SSc patients (52–54). A recent study demonstrated a protective effect of estrogens in dermal fibrosis, as estrogens reduce transforming growth factor β–dependent activation of dermal fibroblasts, and estrogen inhibition leads to a more severe experimental dermal fibrosis, but their effects on vasculature are largely unknown (55). At macrovascular level, hormone replacement therapy might be protective against the risk of pulmonary arterial hypertension, and short- or long-term administration of conjugated estrogens induced flow-mediated dilatation in the brachial artery of SSc patients (56–58). Regarding microvasculature, little is known about the effects of estrogen in patients with SSc (22). A recent study investigated the influence of cumulative endogenous estrogen exposure (CEEE) in patients with SSc on the degree of microvascular damage observed through NVC, and no association between length of CEEE and degree of microvascular impairment was found (59).

We aimed to summarize the available evidence about the association between sex, or specific autoantibodies,

and microangiopathy in SSc, but our review is not without limitations. We could include only a limited number of articles, with variable quality and, due to the heterogeneity of patients and outcomes, a meta-analysis could not be conducted.

Contradictory results were found about the association between autoantibodies and microangiopathy and no firm conclusions can be drawn. As NVC has prognostic relevance in the global assessment of each single SSc patient, we believe that the identification of factors possibly affecting microangiopathy is of relevance to elucidate the pathophysiology of microangiopathy and for clinical risk stratification. Therefore, in consideration of the paucity of available data, and especially the lack of data derived from high-quality research, we advocate further prognostic cohort studies to evaluate factors contributing to the degree of microangiopathy in SSc.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Van Leeuwen, Ciaffi, Schoones, de Vries-Bouwstra.

Acquisition of data. Van Leeuwen, Ciaffi, de Vries-Bouwstra.

Analysis and interpretation of data. Van Leewen, Ciaffi, Huizinga, de Vries-Bouwstra.

REFERENCES

1. Geyer M, Muller-Ladner U. The pathogenesis of systemic sclerosis revisited. Clin Rev Allergy Immunol 2011;40:92–103.
2. Walker UA, Tyndall A, Czirjak L, Denton C, Farge-Bancel D, Kowal-Bielecka O, et al. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials and Research group database. Ann Rheum Dis 2007;66:754–63.
3. Koenig M, Joyal F, Fritzler MJ, Roussin A, Abrahamowicz M, Boire G, et al. Autoantibodies and microvascular damage are independent predictive factors for the progression of Raynaud's Phenomenon to systemic sclerosis: a twenty-year prospective study of 586 patients, with validation of proposed criteria for early systemic sclerosis. Arthritis Rheum 2008;58:3902–12.
4. Sulli A, Paolino S, Pizzorni C, Ferrari G, Pacini G, Pesce G, et al. Progression of nailfold capillaroscopic patterns and correlation with organ involvement in systemic sclerosis: a 12 year study. Rheumatology (Oxford) 2020;59:1051–8.
5. Paxton D, Pauling JD. Does nailfold capillaroscopy help predict future outcomes in systemic sclerosis? A systematic literature review. Semin Arthritis Rheum 2018;48:482–94.
6. Kayser C, Fritzler MJ. Autoantibodies in systemic sclerosis: unanswered questions. Front Immunol 2015;6:167.
7. Markusse IM, Meijs J, de Boer B, Bakker JA, Schippers HPC, Schouffoer AA, et al. Predicting cardiopulmonary involvement in patients with systemic sclerosis: complementary value of nailfold videocapillaroscopy patterns and disease-specific autoantibodies. Rheumatology (Oxford) 2017;56:1081–8.
8. Cutolo M, Pizzorni C, Tuccio M, Burroni A, Craviotto C, Basso M, et al. Nailfold videocapillaroscopic patterns and serum autoantibodies in systemic sclerosis. Rheumatology (Oxford) 2004;43:719–26.
9. Steen VD. Autoantibodies in systemic sclerosis. Semin Arthritis Rheum 2005;35:35–42.
10. Shen CY, Li KJ, Lai PH, Yu CL, Hsieh SC. Anti-CENP-B and anti-TOPO-1-containing sera from systemic sclerosis-related diseases with Raynaud's phenomenon induce vascular endothelial cell senescence not via classical p53–p21 pathway. Clin Rheumatol 2018;37:749–56.
11. Sulli A, Ruaro B, Smith V, Pizzorni C, Zampogna G, Gallo M, et al. Progression of nailfold microvascular damage and antinuclear antibody pattern in systemic sclerosis. J Rheumatol 2013;40:634–9.
12. Stern EP, Denton CP. The pathogenesis of systemic sclerosis. Rheum Dis Clin North Am 2015;41:367–82.
13. Cutolo M, Soldano S, Smith V. Pathophysiology of systemic sclerosis: current understanding and new insights. Expert Rev Clinical Immunol 2019;15:753–64.
14. Kahaleh B. Vascular disease in scleroderma: mechanisms of vascular injury. Rheum Dis Clin North Am 2008;34:57–71;vi.
15. Nietert PJ, Silver RM. Systemic sclerosis: environmental and occupational risk factors. Curr Opin Rheumatol 2000;12:520–6.
16. Hussein H, Lee P, Chau C, Johnson SR. The effect of male sex on survival in systemic sclerosis. J Rheumatol 2014;41:2193–200.
17. Ferri C, Valentini G, Cozzi F, Sebastiani M, Michelassi C, La Montagna G, et al. Systemic sclerosis: demographic, clinical, and serologic features and survival in 1,012 Italian patients. Medicine (Baltimore) 2002;81:139–53.
18. Medsger Jr TA, Masi AT. Survival with scleroderma. II. A life-table analysis of clinical and demographic factors in 358 male U.S. veteran patients. J Chronic Dis 1973;26:647–60.
19. Reveille JD, Fischbach M, McNearney T, Friedman AW, Aguilar MB, Lisse J, et al. Systemic sclerosis in 3 US ethnic groups: a comparison of clinical, sociodemographic, serologic, and immunogenetic determinants. Semin Arthritis Rheum 2001;30:332–46.
20. Walker UA, Tyndall A, Czirjak L, Denton CP, Farge-Bancel D, Kowal-Bielecka O, et al. Geographical variation of disease manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials and Research (EUSTAR) group database. Ann Rheum Dis 2009;68:856–62.
21. Peoples C, Medsger TA Jr, Lucas M, Rosario BL, Feghali-Bostwick CA. Gender differences in systemic sclerosis: relationship to clinical features, serologic status and outcomes. J Scleroderma Relat Disord 2016;1:177–240.
22. Fraenkel L, Zhang Y, Chaisson CE, Evans SR, Wilson PW, Felson DT. The association of estrogen replacement therapy and the Raynaud phenomenon in postmenopausal women. Ann Intern Med 1998;129:208–11.
23. Campbell PM, LeRoy EC. Pathogenesis of systemic sclerosis: a vascular hypothesis. Semin Arthritis Rheum 1975;4:351–68.
24. Rosato E, Gigante A, Barbano B, Cianci R, Molinaro I, Pisarri S, et al. In systemic sclerosis macrovascular damage of hands digital arteries correlates with microvascular damage. Microvasc Res 2011;82:410–5.
25. Bregenzer N, Distler O, Meyringer R, Scholmerich J, Muller-Ladner U, Lock G. Doppler ultrasound identifies increased resistive indices in SSc. Ann Rheum Dis 2004;63:109–10.
26. Schmidt WA, Krause A, Schicke B, Wernicke D. Color Doppler ultrasonography of hand and finger arteries to differentiate primary from secondary forms of Raynaud's phenomenon. J Rheumatol 2008;35:1591–8.
27. Maricq HR, Harper FE, Khan MM, Tan EM, LeRoy EC. Microvascular abnormalities as possible predictors of disease subsets in Raynaud phenomenon and early connective tissue disease. Clin Exp Rheumatol 1983;1:195–205.
28. Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.

29. LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. J Rheumatol 2001;28:1573–6.
30. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010;25:603–5.
31. NIH National Heart, Lung and Blood Institute. Study quality assessment tools. URL: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools. 2013.
32. Cutolo M, Melsens K, Wijnant S, Ingegnoli F, Thevissen K, De Keyser F, et al. Nailfold capillaroscopy in systemic lupus erythematosus: a systematic review and critical appraisal. Autoimmun Rev 2018;17:344–52.
33. Caramaschi P, Canestrini S, Martinelli N, Volpe A, Pieropan S, Ferrari M, et al. Scleroderma patients nailfold videocapillaroscopic patterns are associated with disease subset and disease severity. Rheumatology (Oxford) 2007;46:1566–9.
34. De Santis M, Ceribelli A, Cavaciocchi F, Crotti C, Massarotti M, Belloli L, et al. Nailfold videocapillaroscopy and serum VEGF levels in scleroderma are associated with internal organ involvement. Auto Immun Highlights 2016;7:5.
35. Fichel F, Baudot N, Gaitz JP, Trad S, Barbe C, Frances C, et al. Systemic sclerosis with normal or nonspecific nailfold capillaroscopy. Dermatology 2014;228:360–7.
36. Ghizzoni C, Sebastiani M, Manfredi A, Campomori F, Colaci M, Giuggioli D, et al. Prevalence and evolution of scleroderma pattern at nailfold videocapillaroscopy in systemic sclerosis patients: clinical and prognostic implications. Microvasc Res 2015;99:92–5.
37. Pizzorni C, Giampetruzzi AR, Mondino C, Facchiano A, Abeni D, Paolino S, et al. Nailfold capillaroscopic parameters and skin telangiectasia patterns in patients with systemic sclerosis. Microvasc Res 2017;111:20–4.
38. Ingegnoli F, Ardoino I, Boracchi P, Cutolo M, Airo P, Ananieva LP, et al. Nailfold capillaroscopy in systemic sclerosis: data from the EULAR scleroderma trials and research (EUSTAR) database. Microvasc Res 2013;89:122–8.
39. Tieu J, Hakendorf P, Woodman RJ, Patterson K, Walker J, Roberts-Thomson P. The role of nailfold capillary dropout on mortality in systemic sclerosis. Intern Med J 2018;48:517–23.
40. Chandran G, Smith M, Ahern MJ, Roberts-Thomson PJ. A study of scleroderma in South Australia: prevalence, subset characteristics and nailfold capillaroscopy. Aust N Z J Med 1995;25:688–94.
41. Caramaschi P, Volpe A, Pieropan S, Tinazzi I, Mahamid H, Bambara LM, et al. Cyclophosphamide treatment improves microvessel damage in systemic sclerosis. Clin Rheumatol 2009;28:391–5.
42. Freire M, Rivera A, Sopena B, Tolosa VC, Guillen-Del CA, Colunga AD, et al. Clinical and epidemiological differences between men and women with systemic sclerosis: a study in a Spanish systemic sclerosis cohort and literature review. Clin Exp Rheumatol 2017;35 Suppl S:89S–97S.
43. Simeon CP, Castro-Guardiola A, Fonollosa V, Armadans L, Clemente C, Solans R, et al. Systemic sclerosis in men: clinical and immunological differences. Br J Rheumatol 1996;35:910–1.
44. Okroglic S, Widmann CN, Urbach H, Scheltens P, Heneka MT. Clinical symptoms and risk factors in cerebral microangiopathy patients. PloS One 2013;8:e53455.
45. Sörensen BM, Houben A, Berendschot T, Schouten J, Kroon AA, van der Kallen CJ, et al. Cardiovascular risk factors as determinants of retinal and skin microvascular function: The Maastricht Study. PloS One 2017;12:e0187324.
46. Regitz-Zagrosek V, Kararigas G. Mechanistic pathways of sex differences in cardiovascular disease. Physiol Rev 2017;97:1–37.
47. Volgman AS, Bairey Merz CN, Aggarwal NT, Bittner V, Bunch TJ, Gorelick PB, et al. Sex differences in cardiovascular disease and cognitive impairment: another health disparity for women? J Am Heart Assoc 2019;8:e013154.
48. Tan TC, Fang H, Magder LS, Petri MA. Differences between male and female systemic lupus erythematosus in a multiethnic population. J Rheumatol 2012;39:759–69.
49. Rekvig OP, Putterman C, Casu C, Gao HX, Ghirardello A, Mortensen ES, et al. Autoantibodies in lupus: culprits or passive bystanders? Autoimmun Rev 2012;11:596–603.
50. Fujimoto M, Watanabe R, Ishitsuka Y, Okiyama N. Recent advances in dermatomyositis-specific autoantibodies. Curr Opin Rheumatol 2016;28:636–44.
51. Smith V, Decuman S, Sulli A, Bonroy C, Piettte Y, Deschepper E, et al. Do worsening scleroderma capillaroscopic patterns predict future severe organ involvement? A pilot study. Ann Rheum Dis 2012;71:1636–9.
52. Cutolo M, Brizzolara R, Atzeni F, Capellino S, Straub RH, Puttini PC. The immunomodulatory effects of estrogens: clinical relevance in immune-mediated rheumatic diseases. Ann N Y Acad Sci 2010;1193:36–42.
53. Aida-Yasuoka K, Peoples C, Yasuoka H, Hershberger P, Thiel K, Cauley JA, et al. Estradiol promotes the development of a fibrotic phenotype and is increased in the serum of patients with systemic sclerosis. Arthritis Res Ther 2013;15:R10.
54. Soldano S, Montagna P, Brizzolara R, Sulli A, Parodi A, Seriolo B, et al. Effects of estrogens on extracellular matrix synthesis in cultures of human normal and scleroderma skin fibroblasts. Ann N Y Acad Sci 2010;1193:25–9.
55. Avouac J, Pezet S, Gonzalez V, Baudoin L, Cauvet A, Ruiz B, et al. Estrogens counteract the profibrotic effects of TGF-beta and their inhibition exacerbates experimental dermal fibrosis. J Invest Dermatol 2020;140:593–601.
56. Beretta L, Caronni M, Origgi L, Ponti A, Santaniello A, Scorza R. Hormone replacement therapy may prevent the development of isolated pulmonary hypertension in patients with systemic sclerosis and limited cutaneous involvement. Scand J Rheumatol 2006;35:468–71.
57. Lekakis J, Papamichael C, Mavrikakis M, Voutsas A, Stamatelopoulos S. Effect of long-term estrogen therapy on brachial arterial endothelium-dependent vasodilation in women with Raynaud's phenomenon secondary to systemic sclerosis. Am J Cardiol 1998;82:1555–7, A8.
58. Lekakis J, Mavrikakis M, Papamichael C, Papazoglou S, Economou O, Scotiniotis I, et al. Short-term estrogen administration improves abnormal endothelial function in women with systemic sclerosis and Raynaud's phenomenon. Am Heart J 1998;136:905–12.
59. Ciaffi J, van Leeuwen NM, Huizinga TW, de Vries-Bouwstra JK. Cumulative endogenous estrogen exposure is not associated with severity of peripheral microangiopathy in patients with systemic sclerosis. Clin Exp Rheumatol 2019;37 Suppl S:82S–7S.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 732–741
DOI 10.1002/acr.24167

Clinical Features of Systemic Sclerosis–Mixed Connective Tissue Disease and Systemic Sclerosis Overlap Syndromes

Jessica L. Fairley,[1] Dylan Hansen,[2] Susanna Proudman,[3] Joanne Sahhar,[4] Gene-Siew Ngian,[4] Jenny Walker,[5] Gemma Strickland,[2] Michelle Wilson,[2] Kathleen Morrisroe,[2] Nava Ferdowsi,[2] Gabor Major,[6] Janet Roddy,[7] Wendy Stevens,[2] and Mandana Nikpour,[8] for the Australian Scleroderma Interest Group

Objective. To describe the clinical characteristics and outcomes of systemic sclerosis–mixed connective tissue disease (SSc–MCTD) and SSc overlap syndrome.

Methods. We included patients from the Australian Scleroderma Cohort Study who met American College of Rheumatology/European Alliance of Associations for Rheumatology criteria for SSc. Three mutually exclusive groups were created: SSc–MCTD, SSc overlap, and SSc only. Univariate comparison of clinical features was performed by analysis of variance or chi-square test. Survival analysis was performed using Kaplan-Meier (KM) curves and Cox proportional hazards regression models.

Results. Of 1,728 patients, 97 (5.6%) had SSc–MCTD, and 126 (7.3%) had SSc overlap. Those with MCTD–SSc were more commonly Asian (18.3% versus 10.1% in SSc overlap, and 3.6% in SSc only; $P < 0.0001$) and younger at disease onset (38.4 years versus 46.5 or 46.8 years, $P < 0.0001$). Those with SSc–MCTD or SSc overlap were more likely to have limited cutaneous SSc. All 3 groups had similar frequency of interstitial lung disease (ILD), although pulmonary arterial hypertension (PAH) was less common in SSc overlap. Synovitis and myositis were more common in SSc overlap and SSc–MCTD than in SSc only. KM curves showed better survival in SSc–MCTD than SSc overlap or SSc only ($P = 0.011$), but this was not significant after adjustment for sex and age at disease onset. SSc-specific antibodies were survival prognostic markers, with antinuclear antibody centromere or anti-RNP conferring better survival than anti–Scl-70 or anti–RNA polymerase III ($P = 0.005$). Patients with SSc–MCTD and SSc overlap had lower mortality following diagnosis of ILD and PAH than patients with SSc only.

Conclusion. This study provides insights into the clinical characteristics of patients with SSc–MCTD, SSc overlap, and SSc only and shows that anti-RNP antibodies are associated with better survival than anti–Scl-70 and anti-RNA polymerase III antibodies.

INTRODUCTION

Mixed connective tissue disease (MCTD) is a heterogeneous clinical syndrome first described in 1972 (1). It is characterized by overlapping features of multiple connective tissue diseases, including systemic sclerosis (SSc), systemic lupus erythematosus (SLE), polymyositis (PM), and rheumatoid arthritis (RA) (1,2). Features may include synovitis, myositis, finger swelling, Raynaud's phenomenon, and acrosclerosis (1,3). Antibody to the extractable nuclear antigen (ENA) U1-RNP complex is the serologic hallmark of MCTD (2). Multiple diagnostic criteria exist, with the most sensitive and specific being those described by Alargon-Sergovia and Kahn (3). These criteria require positive anti–U1-RNP antibodies (titer ≥1:1,600) in combination with ≥3 of swollen hands, synovitis, myositis, Raynaud's phenomenon, and acrosclerosis (3).

By definition, MCTD shares many features with other connective tissue diseases (2,4). Accordingly, some patients with MCTD will fulfill diagnostic criteria for both MCTD and another connective

Dr. Morrisroe's work was supported by an Arthritis Australia Fellowship. Dr. Nikpour is recipient of a National Health and Medical Research Council Career Development Fellowship (grant APP1126370).

The Australian Scleroderma Cohort Study is supported by Actelion, Scleroderma Australia, Scleroderma Victoria, Arthritis Australia, the Australian Rheumatology Association, the St. Vincent's Hospital Melbourne Research Endowment Fund, GlaxoSmithKline, Pfizer, Bristol Myers Squibb, Roche, and Bayer. The Australian Scleroderma Interest Group is supported by Actelion.

[1]Jessica L. Fairley, MBBS (Hons): The Alfred Hospital and Monash University, Melbourne, Victoria, Australia; [2]Dylan Hansen, BSc (Pharm Sc, Mbiostats), Gemma Strickland, MBBS (Hons), FRACP, Michelle Wilson, BSc (Hons), PhD, Kathleen Morrisroe, MBBS, FRACP, PhD, Nava Ferdowsi, MBBS, FRACP, Wendy Stevens, MBBS, FRACP: St. Vincent's Hospital Melbourne, Melbourne, Victoria, Australia; [3]Susanna Proudman, MBBS (Hons), FRACP: Royal Adelaide Hospital and University of Adelaide, Adelaide, South Australia, Australia; [4]Joanne Sahhar, MBBS (Hons), FRACP, Gene-Siew Ngian, MBBS (Hons), FRACP, PhD: Monash Health and Monash University, Melbourne, Victoria, Australia; [5]Jenny Walker, MBBS, FRACP, PhD: Royal Adelaide Hospital, Adelaide, South Australia, Australia; [6]Gabor Major, MBBS, FRACP: Royal Newcastle Centre John Hunter Hospital and University of Newcastle, Newcastle, New South Wales, Australia; [7]Janet Roddy, MD, FRACP: Fiona

SIGNIFICANCE & INNOVATIONS

- There are some significant differences in clinical features of systemic sclerosis–mixed connective tissue disease (SSc–MCTD) and SSc overlap compared with SSc only.
- Antibodies may be more accurate at predicting prognosis than classification according to these disease groups.
- While development of interstitial lung disease (ILD) or pulmonary arterial hypertension (PAH) was a poor prognostic factor, patients with SSc–MCTD and SSc overlap had lower mortality following diagnosis of ILD and PAH than patients with SSc only.

tissue disease. Of interest is the overlap between patients who meet diagnostic criteria for MCTD and SSc simultaneously.

Patients who have both SSc and clinical features of another connective tissue disease, such as SLE, RA, PM, or Sjögren's syndrome, are often classified as having SSc overlap syndromes (2,4). The wider literature suggests that these patients have a different disease course and organ involvement from those who purely meet criteria for limited or diffuse cutaneous SSc (2); thus, arguably, they should be considered a distinct subgroup.

Our study aims to describe the clinical phenotype of patients in the Australian Scleroderma Cohort Study (ASCS) who are recorded as having MCTD or SSc overlap syndromes. We aim to compare these groups with the remainder of the cohort in terms of clinical manifestations and outcome.

PATIENTS AND METHODS

Patients. Patients were recruited from the ASCS. The ASCS is a multicenter study across 13 participating Australian centers to investigate risk and prognostic factors in SSc. The ASCS has been approved by all human research ethics committees of participating sites, with St. Vincent's Hospital Melbourne Human Research Ethics Committee acting as the coordinating site. Written informed consent was obtained from all patients at recruitment. See Appendix A for members of the ASCS.

Selection of patient groups. We included only those patients who met the American College of Rheumatology/European Alliance of Associations for Rheumatology criteria for SSc (5). We divided patients into 3 mutually exclusive groups for analysis: those with SSc and MCTD (SSc–MCTD), those with SSc overlap syndromes (SSc overlap), and those with SSc only. SSc–MCTD was defined as positive anti-RNP antibodies, and at least 3 of the following clinical features: synovitis, myositis, finger swelling, Raynaud's phenomenon, and acrosclerosis in accordance with accepted diagnostic criteria (1,3). Patients who were positive for anti-RNP but did not have at least 3 clinical features of MCTD as listed above were not included in the SSc–MCTD group but in the SSc only or SSc overlap groups depending on other clinical features recorded in the database and classification according to the treating physician. SSc overlap was designated by the treating physician if there were clinical features of another connective tissue diseases present (e.g., SLE, RA, PM, or Sjögren's syndrome), although it was not mandated that patients independently fulfilled diagnostic criteria for these conditions. In those who were classified as having SSc overlap, physicians were offered the option of nominating the connective tissue disease that patients shared features with, but this was not compulsory. The physician could also nominate >1 overlap condition. Patients who were listed in the database as having both MCTD and an overlap syndrome were included in the SSc–MCTD group. Patients who did not meet criteria for SSc–MCTD or SSc overlap were included in the SSc only group.

Autoantibody testing. Indirect immunofluorescence was used to detect antinuclear antibody (ANA). Antibodies to ENAs and antibodies to RNA–polymerase III were detected by enzyme-linked immunosorbent assay (ELISA), immunoblot, or a combination of these using local laboratory commercial test kits. ELISA was used to determine anti–double-stranded DNA (anti-dsDNA) antibody in most laboratories, with the Farr radioimmunoassay used by 2 laboratories. Autoantibody positivity was defined by a positive result according to the local laboratory protocol.

Data collection. Demographic and disease data were prospectively collected at baseline and annual reviews thereafter in a standardized fashion as part of the ASCS. All disease data or antibody results were defined as present if they had ever been reported from the time of diagnosis. Disease onset and duration was defined as time from the first non–Raynaud's phenomenon manifestation. The LeRoy criteria were used to determine disease subtype (diffuse or limited) (6). Pulmonary arterial hypertension (PAH) was diagnosed by right heart catheterization, using a mean pulmonary artery pressure ≥25 mm Hg in association with a pulmonary arterial wedge pressure ≤15 mm Hg. High-resolution

Stanley Hospital, Perth, Western Australia, Australia; [8]Mandana Nikpour, MBBS, FRACP, FRCPA, PhD: St. Vincent's Hospital Melbourne, Melbourne, and The University of Melbourne at St Vincent's Hospital (Melbourne), Fitzroy, Victoria, Australia.

Dr. Proudman has received consulting fees and/or honoraria from Actelion and Boehringer Ingelheim (less than $10,000 each). Dr. Major has received honoraria from Roche, Bristol Myers Squibb, Boehringer Ingelheim, and Novartis (less than $10,000 each). Dr. Stevens has received speaking fees from Janssen (less than $10,000). No other disclosures relevant to this article were reported.

Address correspondence to Mandana Nikpour, MBBS, FRACP, FRCPA, PhD, The University of Melbourne at St. Vincent's Hospital, 41 Victoria Parade Fitzroy, Melbourne, Victoria 3065, Australia. Email: m.nikpour@unimelb.edu.au.

Submitted for publication September 19, 2019; accepted in revised form February 4, 2020.

computed tomography (HRCT) of the chest was used to diagnose interstitial lung disease (ILD), usually performed in response to clinical examination findings (chest crepitations) or abnormal respiratory function testing. Severity of ILD was defined by the extent of involvement on HRCT (mild <20%, moderate 20–30%, severe >30%). In the absence of symptoms suggestive of PAH or ILD, patients within the ASCS were screened annually with both transthoracic echocardiography and pulmonary function testing. Definitive diagnostic testing as discussed above was arranged if these were abnormal. Scleroderma renal crisis was diagnosed in the presence of 2 of 3 criteria: new-onset hypertension with no alternate cause, unexplained rise in serum creatinine, or microangiopathic hemolytic anemia. Small intestinal bacterial overgrowth was diagnosed by concurrent diarrhea and use of cyclical antibiotics. Endoscopy was used to diagnose gastric antral vascular ectasia, reflux esophagitis, and esophageal strictures. Hospitalization data were collected annually based on patient-reported admissions of >24 hours. Malignancy was defined by presence of melanoma, solid organ, or hematologic malignancy.

Statistical analysis. Characteristics of patients in the study are presented as the mean ± SD for continuous variables or number (%) for categorical variables. We compared continuous variables among the 3 groups using 1-way analysis of variance. Discrete variables were compared using chi-square test. All-cause mortality was used for survival analysis. Kaplan-Meier (KM) curves and the Wilcoxon test were used to estimate survival from SSc onset, PAH diagnosis, and ILD diagnosis according to disease group and antibody status. Cox proportional hazards regression analysis was used to determine multivariable predictors of mortality. Patient characteristics that were clinically significant were considered for multivariate analysis of survival, including sex, age at disease onset, and one of either disease group or autoantibody status. In the antibody model, any patient who had multiple antibodies (among ANA centromere, anti-RNP, anti–Scl-70, and anti–RNA polymerase III) was excluded from the analysis. Any patients who did not have any of these antibodies were included in the "no antibody" category. Characteristics with a P value ≤0.05 who did not violate the proportional hazards assumption were included in the multivariate model. Due to collinearity between disease groups and SSc-specific antibodies, these variables were each included in separate multivariable models. The results were reported as hazard ratios (HRs) with accompanying 95% confidence intervals (95% CIs). All statistical analyses were performed using Stata, version 15.1.

RESULTS

Description of whole cohort. A total of 1,728 patients fulfilled the inclusion criteria for this study. The characteristics of this cohort are summarized in Table 1. Of the 1,728 patients included, 1,489 (86%) were female, 1,285 (74%) had limited disease, and 442 (26%) diffuse disease. In total, 97 patients (5.6%) were identified as having both SSc and MCTD (SSc–MCTD), while 126 (7.3%) were identified as having SSc overlap syndrome. Patients with SSc–MCTD or SSc overlap were more likely to have limited cutaneous SSc than patients with SSc only (84.5% and 83.3% versus 73%). Most patients were White (92%), followed by Asian ethnicity (4.9%). Ethnicity was similar between SSc only and SSc overlap, although in the SSc–MCTD group, Asian background was significantly more common (18.3% versus 10.1% [SSc overlap] and 3.6% [SSc only]; $P < 0.0001$). Patients with SSc–MCTD were younger at disease onset (38.4 years versus 46.5 or 46.8 years; $P < 0.0001$). Mean duration of follow-up was similar between groups at ~4.5 years. Among those who had SSc overlap, 49 (38.9%) were listed as overlap with RA, 17 (13.5%) with SLE, 22 (17.5%) with polymyositis, 3 (2.4%) with dermatomyositis, 43 (34.1%) with Sjögren's syndrome, and in 2 (1.6%), the overlap condition was not specified.

Autoantibody profile of the 3 disease groups. The autoantibody profile of the cohort is summarized in Table 1. In accordance with our definition of MCTD, antibodies to RNP were positive in all patients with SSc–MCTD. Anti-RNP was positive in 2.4% of patients with SSc overlap and 0.3% with SSc only. These RNP-positive patients did not otherwise meet diagnostic criteria for MCTD. Anti–Scl-70 was more commonly positive in patients with SSc only (15.0%) or SSc overlap (20.2%) than those with SSc–MCTD (7.5%). There was a higher frequency of anti–RNA polymerase III positivity in SSc only (14.6%) compared with the other 2 groups (SSc overlap 8.0%, SSc–MCTD 2.1%; $P = 0.0135$). Anti–Jo-1 positivity was more common in patients with SSc overlap (1.7%) or SSc–MCTD (2.1%) than SSc only (0.3%; $P = 0.0073$). Anti-Ro and anti-La were both more common in SSc–MCTD (26.3% and 5.3%, respectively) and SSc overlap (28.8% and 4.2%, respectively) than SSc only (6.6% and 1.3%, respectively) ($P < 0.0001$ and $P = 0.0023$, respectively). Anti-Sm positivity was significantly more common in SSc–MCTD (25.8%) than SSc overlap (3.4%) or SSc only (0.3%; $P < 0.0001$).

Antineutrophil cytoplasmic antibodies (ANCAs) were more common in patients with SSc–MCTD (25.0%) or SSc overlap (25.2%) than SSc only (13.2%; $P = 0.0001$), although with no significant differences in frequency of anti-MPO or anti-PR3 positivity. Anti-CCP antibody was most common in SSc overlap (9.4%), followed by SSc–MCTD (7.1%) and SSc only (2.8%; $P = 0.0451$), without significant difference in frequency of rheumatoid factor positivity. Anti-dsDNA antibody was significantly more common in SSc–MCTD (19.8%) and SSc overlap (15.0%) than SSc only (5.9%; $P < 0.0001$).

Clinical characteristics and organ involvement. The clinical characteristics of patients is presented in Table 2. Regarding cardiopulmonary involvement, PAH was more common in those with SSc–MCTD (12.4%) and SSc only (11.1%) than in patients with SSc overlap (4.8%), although this was not statistically significant ($P = 0.0751$). There was no significant difference

Table 1. Demographic and autoantibody profiles of study participants (n = 1,728)*

Variable	SSc only group	SSc–MCTD overlap group	SSc overlap group	*P*
Sex				
Female	1,293 (85.9)	83 (85.6)	113 (89.7)	0.4924
Male	212 (14.1)	14 (14.4)	13 (10.3)	
Disease subtype				
Diffuse	406 (27.0)	15 (15.5)	21 (16.7)	0.0024
Limited	1,098 (73.0)	82 (84.5)	105 (83.3)	
Race				
Aboriginal Islander	16 (1.1)	1 (1.1)	1 (0.8)	<0.0001
Asian	52 (3.6)	17 (18.3)	12 (10.1)	
White	1,341 (93.3)	71 (76.3)	105 (88.2)	
Hispanic	12 (0.8)	1 (1.1)	0 (0.0)	
Other	16 (1.1)	3 (3.2)	1 (0.8)	
Age at recruitment	57.6 (12.50)	49.9 (13.67)	57.5 (12.4)	<0.0001
Age at onset of SSc	46.8 (14.1)	38.4 (14.4)	46.5 (15.2)	<0.0001
Follow-up in ASCS, years	4.5 (3.33)	4.6 (2.98)	4.4 (3.0)	0.8699
Ever smoked	742 (49.3)	50 (51.5)	64 (50.8)	0.9885
Autoantibody profile				
ANA positive	1,395 (95.4)	94 (96.9)	114 (92.7)	0.2884
ANA centromere	717 (49.6)	11 (11.6)	51 (42.1)	<0.0001
ANA homogeneous	278 (19.5)	17 (18.1)	28 (23.5)	0.5293
ANA nucleolar	329 (22.9)	8 (8.5)	25 (21.0)	0.0045
ANA speckled	352 (24.6)	76 (79.2)	34 (28.1)	<0.0001
Anti–RNA polymerase III	141 (14.6)	1 (2.1)	7 (8.0)	0.0135
Anti-RNP	5 (0.3)	97 (100.0)	3 (2.4)	<0.0001
Anti–Jo-1	4 (0.3)	2 (2.1)	2 (1.7)	0.0073
Anti-La	19 (1.3)	5 (5.3)	5 (4.2)	0.0023
Anti-Ro†	93 (6.6)	25 (26.3)	34 (28.8)	<0.0001
Anti–Scl-70	214 (15.0)	7 (7.5)	24 (20.2)	0.0374
Anti-Sm	4 (0.3)	24 (25.8)	4 (3.4)	<0.0001
ANCAs	175 (13.2)	21 (25.0)	27 (25.2)	0.0001
Anti-MPO	19 (1.4)	2 (2.4)	3 (2.8)	0.4558
Anti-PR3	24 (1.8)	3 (3.6)	4 (3.7)	0.2340
Anti-dsDNA	70 (5.9)	17 (19.8)	17 (15.0)	<0.0001
Anti-CCP	11 (2.8)	2 (7.1)	5 (9.3)	0.0451
Rheumatoid factor	391 (28.9)	31 (34.1)	38 (33.3)	0.3769
Anti–PM-Scl	21 (1.5)	1 (1.1)	2 (1.7)	0.9312
No SSc-specific antibody‡	376 (27.9)	71 (81.6)	42 (36.2)	<0.0001

* Values are the number (%) unless indicated otherwise. ANA = antinuclear antibody; ANCAs = antineutrophil cytoplasmic antibodies; ASCS = Australian Scleroderma Cohort Study; CCP = cyclic citrullinated peptide; dsDNA = double-stranded DNA; MCTD = mixed connective tissue disease; PR3 = proteinase 3; SSc = systemic sclerosis.
† Anti-Ro60 antibody.
‡ This refers to the absence of scleroderma-specific antibodies, i.e., none of antinuclear antibody centromere, anti-RNP, anti–Scl-70, or anti–RNA polymerase III.

in mean pulmonary artery pressure at PAH diagnosis among groups. No significant differences existed for frequency of ILD between groups. No difference existed in frequency of pericardial or myocardial disease.

In terms of gastrointestinal involvement, patients with SSc overlap were significantly more likely to have experienced dysphagia (60.3% versus SSc–MCTD 45.4%, SSc only 45.5%; $P = 0.0006$) than those with SSc only or SSc–MCTD. There was a higher frequency of esophageal strictures in this group (SSc overlap 24.6% versus SSc–MCTD 23.7% and SSc only 16.7%; $P = 0.0221$). Lowest recorded body mass index was within normal range in all groups, although lower in those with SSc–MCTD (23.6 versus SSc only 25.0 and SSc overlap 24.4; $P = 0.0260$).

In terms of musculoskeletal and mucocutaneous manifestations, patients with SSc–MCTD were less likely to experience non-hand skin ulcers (SSc only 8.8% versus SSc overlap 7.9%, SSc–MCTD 4.1%; $P = 0.0018$), calcinosis (SSc–MCTD 21.6% versus SSc only 41.3%, SSc overlap 37.3%; $P = 0.0011$), and joint contractures than other groups (SSc–MCTD 25.8% versus SSc only 39.7%, SSc overlap 42.9%; $P = 0.0344$). There was no difference in frequency of sclerodactyly. Highest recorded modified Rodnan skin thickness scores were greater in those with SSc only (mean ± SD 11.9 ± 9.6) rather than SSc–MCTD (mean ± SD 8.8 ± 7.8) or SSc overlap (mean ± SD 9.4 ± 7.9; $P = 0.0003$). Sicca symptoms were more common in those with SSc overlap than

Table 2. Organ involvement and immunosuppressive treatment in study participants (n = 1,728)*

Variable	SSc only group	SSc–MCTD group	SSc overlap group	*P*
Cardiopulmonary				
PAH†	167 (11.1)	12 (12.4)	6 (4.8)	0.0751
PAP at PAH diagnosis, mean ± SD mm Hg	35.4 ± 10.1	38.8 ± 13.0	29.3 ± 5.2	0.1742
ILD (on HRCT)†	393 (66.4)	25 (61.0)	34 (70.8)	0.6178
Severity of ILD‡				
Mild (<20%)	212 (56.5)	17 (70.8)	17 (53.1)	0.6317
Moderate (20–30%)	111 (29.6)	4 (16.7)	11 (34.4)	
Severe (>30%)	52 (13.9)	3 (12.5)	4 (12.5)	
Lowest FVC, mean ± SD†	89.3 ± 22.1	81.5 ± 19.5	87.7 ± 21.6	0.0029
Pericardial effusion†	126 (8.5)	7 (7.3)	7 (5.7)	0.5396
Myocardial disease†	114 (7.6)	6 (6.2)	10 (7.9)	0.8667
Gastrointestinal				
Malabsorption†	56 (3.7)	7 (7.2)	9 (7.1)	0.0549
Rectal prolapse†	30 (2.0)	0 (0.0)	5 (4.0)	0.1103
GAVE†	169 (11.2)	5 (5.2)	10 (7.9)	0.1010
Esophageal stricture†	251 (16.7)	23 (23.7)	31 (24.6)	0.0221
Fecal incontinence†	439 (29.2)	16 (16.5)	34 (27.0)	0.0577
Dysphagia†	685 (45.5)	44 (45.4)	76 (60.3)	0.0006
Reflux esophagitis†	1,248 (82.9)	80 (82.5)	107 (84.9)	0.8382
Vomiting†	320 (21.3)	18 (18.6)	45 (35.7)	0.0044
Lowest BMI score†	25.0 (5.4)	23.6 (4.8)	24.4 (5.0)	0.0260
Musculoskeletal and mucocutaneous				
SSc skin changes present†	1,365 (93.0)	81 (85.3)	111 (89.5)	0.0115
Skin ulcers (non-hand)†	132 (8.8)	4 (4.1)	10 (7.9)	0.0018
Highest MRSS score†	11.9 (9.6)	8.8 (7.8)	9.4 (7.9)	0.0003
Synovitis†	541 (35.9)	56 (57.7)	74 (58.7)	<0.0001
Myositis†	64 (4.3)	18 (18.6)	28 (22.2)	<0.0001
Calcinosis†	622 (41.3)	21 (21.6)	47 (37.3)	0.0011
Joint contractures	597 (39.7)	25 (25.8)	54 (42.9)	0.0344
Large joint contractures	54 (3.6)	1 (1.0)	7 (5.6)	0.0074
Small joint contractures	268 (17.8)	7 (7.2)	31 (24.6)	0.0159
Puffy digits/scleredema†	1,041 (69.2)	80 (82.5)	84 (66.7)	0.0522
Sclerodactyly†	1,337 (88.8)	84 (86.6)	114 (90.5)	0.7123
Dry eyes†	941 (62.5)	64 (66.0)	100 (79.4)	0.0056
Dry mouth†	1,085 (72.1)	67 (69.1)	110 (87.3)	0.0048
Tendon friction rubs†	130 (8.6)	8 (8.2)	12 (9.5)	0.9810
Renal				
Renal crisis†	55 (3.7)	0 (0.0)	5 (4.0)	0.1548
Glomerular filtration rate (lowest)†				
<30	40 (2.8)	1 (1.1)	5 (4.1)	0.1545
30–60	347 (24.4)	14 (14.9)	29 (24.0)	
>60	1,036 (72.8)	79 (84.0)	87 (71.9)	
Vascular				
Raynaud's phenomenon†	1,494 (99.3)	97 (100.0)	126 (100.0)	0.4404
Digital gangrene/amputation†	199 (13.2)	11 (11.3)	10 (7.9)	0.4191
Digital ulcers†	777 (51.6)	42 (43.3)	56 (44.4)	0.2876
Telangiectasia†	1,300 (86.4)	82 (84.5)	96 (76.2)	0.0094
Malignancy				
All malignancies§	314 (20.9)	15 (15.5)	26 (20.6)	0.4430
Biochemistry/laboratory parameters				
Low C3†	238 (17.6)	21 (22.8)	31 (27.0)	0.0260
Low C4†	234 (17.3)	31 (33.7)	30 (26.1)	0.0001
Highest ESR†	27.5 (24.7)	34.6 (25.7)	30.5 (23.6)	0.0145
Highest CK†	132.6 (134.9)	209.1 (403.7)	199.7 (378.1)	<0.0001
Lowest hemoglobin†	123.6 (17.2)	120.8 (14.2)	123.1 (17.3)	0.2650
Lowest albumin†	37.5 (4.5)	37.1 (5.6)	37.0 (4.2)	0.4580
Lowest platelet count†	244.3 (71.0)	225.7 (69.9)	246.7 (64.4)	0.0869

(Continued)

Table 2. (*Cont'd*)

Variable	SSc only group	SSc–MCTD group	SSc overlap group	*P*
Immunomodulatory/immunosuppressive treatments				
Abatacept	2 (0.1)	0 (0.0)	2 (1.6)	0.0044
Rituximab/anti-CD20	10 (0.7)	5 (5.2)	9 (7.2)	<0.0001
Azathioprine	106 (7.1)	14 (14.4)	23 (18.4)	<0.0001
Calcineurin inhibitor	21 (1.4)	4 (4.1)	2 (1.6)	0.1116
Cyclophosphamide	135 (9.0)	9 (9.3)	12 (9.6)	0.9715
Hydroxychloroquine	262 (17.5)	48 (49.5)	54 (43.2)	<0.0001
Leflunomide	10 (0.7)	4 (4.1)	5 (4.0)	<0.0001
Methotrexate	292 (19.4)	43 (44.3)	61 (48.8)	<0.0001
Mycophenolate	152 (10.1)	11 (11.3)	16 (12.8)	0.6109
Penicillamine	119 (7.9)	2 (2.1)	8 (6.4)	0.0927
Prednisolone	646 (43.0)	63 (64.9)	79 (63.2)	<0.0001
TNF inhibitors	6 (0.4)	3 (3.1)	6 (4.8)	<0.0001
Tocilizumab	5 (0.3)	1 (1.0)	1 (0.8)	0.4488
Intravenous immunoglobulin	4 (0.3)	0 (0.0)	6 (4.8)	<0.0001

* Values are the number (%) unless indicated otherwise. BMI = body mass index; CK = creatine kinase; ESR = erythrocyte sedimentation rate; FVC = forced vital capacity; GAVE = gastric antral vascular ectasia; HRCT = high-resolution computed tomography; ILD = interstitial lung disease; MCTD = mixed connective tissue disease; MRSS = modified Rodnan skin thickness score; PAH = pulmonary arterial hypertension; PAP = pulmonary artery pressure; SSc = systemic sclerosis; TNF = tumor necrosis factor.
† Ever during follow-up or from SSc diagnosis.
‡ ILD severity based on extent (%) of lung involvement on HRCT of the lung.
§ Recorded malignancies included bowel, breast, hematologic, lung, melanoma, and nonmelanoma skin cancers.

SSc only or SSc–MCTD (dry eyes 79.4% versus SSc–MCTD 66.0% and SSc only 62.5%; $P = 0.0056$; dry mouth 87.3% versus SSc only 72.1% and SSc–MCTD 69.1%; $P = 0.0048$). Synovitis was equally common in SSc–MCTD and SSc overlap groups (57.7% and 58.7%, respectively) compared with SSc only (35.9%; $P < 0.0001$), as was myositis (SSc–MCTD 18.6% and SSc overlap 22.2% versus SSc only 4.3%; $P < 0.0001$). Puffy digits were more common in patients with SSc–MCTD (82.5%) than in those with SSc only (69.2%) or SSc overlap (66.7%; $P = 0.0522$).

No significant differences in the frequency of renal crisis or vascular manifestations existed between groups, with the exception of telangiectasia, more commonly seen in SSc only (86.4%) or SSc–MCTD (84.5%) than SSc overlap (76.2%; $P = 0.0094$). In terms of biochemical and laboratory parameters, patients with SSc overlap were more likely to have had a low C3 reading (27.0% versus SSc–MCTD 22.8% and SSc only 17.6%; $P = 0.0260$), while patients with SSc–MCTD were more likely to have had a low C4 reading (33.7% versus SSc overlap 26.1% and SSc only 17.3%; $P = 0.0001$). Highest recorded erythrocyte sedimentation rate was greater in those with SSc–MCTD (34.6) than those with SSc overlap (30.5) or SSc only (24.7; $P = 0.0145$).

Patients with SSc–MCTD and SSc overlap recorded similar mean peak creatine kinase levels (209.1 and 199.7, respectively), significantly higher than those with SSc only (132.6; $P < 0.0001$). Other parameters were similar between groups.

Treatment data. Exposure to immunosuppressive or immunomodulatory treatment was generally much more common in those with SSc–MCTD or SSc overlap than SSc only (Table 2). Patients with SSc overlap were more likely than those with SSc–MCTD to have been exposed to biologic medications, including abatacept, rituximab, and anti–tumor necrosis factor agents, as well as azathioprine and intravenous immunoglobulin ($P < 0.01$ for all). Both groups were equally likely to have been exposed to synthetic disease-modifying antirheumatic drugs (DMARDs), including hydroxychloroquine, leflunomide, and methotrexate, as well as prednisolone, than those with SSc only ($P < 0.0001$ for all). Hydroxychloroquine and methotrexate were the most commonly used DMARDs in our population, both significantly more common in those with SSc–MCTD and SSc overlap than SSc only. All 3 groups were equally exposed to cyclophosphamide, mycophenolate, calcineurin inhibitors, and tocilizumab.

Patients with ILD were more likely to be treated with rituximab, azathioprine, cyclophosphamide, mycophenolate, and prednisolone than those without ILD (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24167/abstract). Patients with PAH were more likely to be treated with cyclophosphamide (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24167/abstract). Frequency of immunosuppressive therapies in SSc overlap patients by overlap condition is presented in Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24167/abstract.

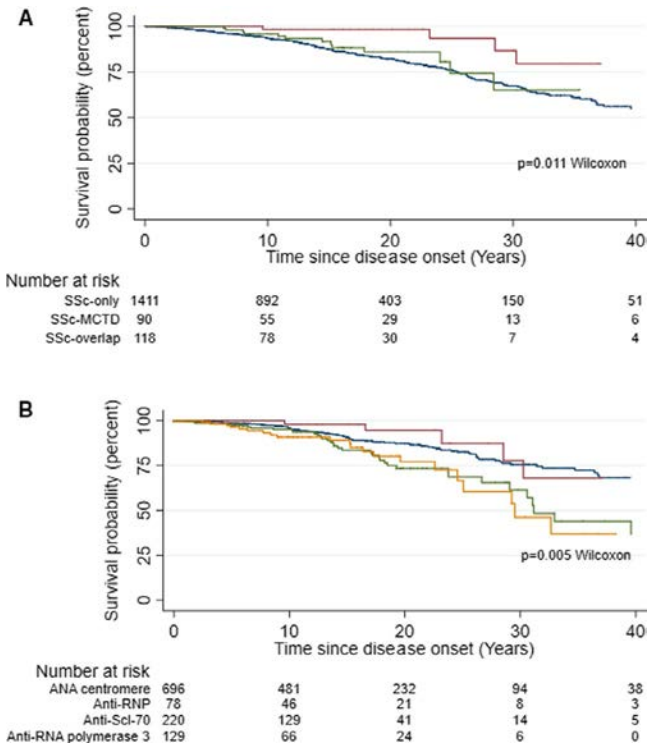


Figure 1. **A**, Survival by disease group, according to Kaplan-Meier survival estimates, for the systemic sclerosis (SSc) only group (blue), the SSc–mixed connective tissue disease (MCTD) group (red), and the SSc overlap group (green). **B**, Survival by autoantibody status for the antinuclear antibody (ANA) centromere group (blue), the anti-RNP group (red), the anti–Scl-70 group (green), and the anti–RNA polymerase III group (yellow).

There was no difference between groups in exposure to antihypertensives, vasodilators, or anticoagulants, or therapies targeting gastrointestinal manifestations (see Supplementary Table 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24167/abstract). No major differences existed in the frequency of PAH treatments.

Survival and risk factors for mortality. Comparing survival by diagnosis, in KM analysis, patients with SSc–MCTD had a better prognosis than those with SSc only ($P = 0.011$) (Figure 1A). Those with SSc–MCTD also had better survival than patients with SSc overlap ($P = 0.037$). However, in multivariable Cox proportional hazards models, after adjusting for sex and age at disease onset, there was no significant difference in survival among the 3 groups (Table 3).

Compared to KM analysis of survival according to disease group, in KM analysis according to antibody positivity, differences were more pronounced (Figure 1B). Those who were positive for ANA centromere had a similar survival to patients positive for anti-RNP (Figure 1B). Patients with anti–Scl-70 or anti–RNA polymerase III positivity had worse survival than patients positive for ANA centromere or anti-RNP. In a multivariate Cox proportional hazards model of survival according to antibody status (using ANA centromere as a reference category), anti–Scl-70 positivity conferred a significantly worse prognosis (HR 2.75 [95% CI 1.88–4.04], $P < 0.001$), as did anti–RNA polymerase III positivity (HR 1.79 [95% CI 1.11–2.89], $P = 0.018$) and absence of any positive SSc-specific antibody (HR 1.85 [95% CI 1.38–2.47], $P < 0.001$). There was no significant difference in survival between patients positive for ANA centromere and anti-RNP (Table 3). Male sex (HR 2.20 [95% CI 1.62–2.97], $P < 0.001$) and older age at disease onset (HR 1.10 [95% CI 1.09–1.11], $P < 0.001$) were predictors of mortality, independently of disease groups (Table 3).

Patients with SSc–MCTD and SSc overlap had lower all-cause mortality following diagnosis of ILD than those with SSc only ($P = 0.024$) (Figure 2A). A similar pattern was seen in all-cause mortality following diagnosis of PAH, but this was not statistically significant ($P = 0.058$) (Figure 2B). However, when SSc–MCTD and SSc overlap patients were combined, they had significantly better survival than those with SSc only ($P = 0.019$) (Figure 2C). SSc myocardial disease or history of renal crisis did not predict increased mortality in this cohort.

DISCUSSION

In this large cohort of patients with SSc, 5.6% of patients were identified as having SSc–MCTD, and 7.3% had SSc overlap. Compared with SSc–MCTD, patients with SSc overlap or SSc only were more likely to have positive SSc-specific antibodies, including ANA centromere, anti–Scl-70, and anti–RNA polymerase III. SSc overlap patients were more likely to have positive anti-CCP antibodies than those with SSc–MCTD. SSc–MCTD patients were more likely to be positive for anti-Sm and anti-dsDNA. Patients with SSc–MCTD or SSc overlap were more likely than those with SSc only to have a number of other positive autoantibodies (anti-Ro, anti-La, anti–Jo-1, and ANCA).

Clinically, both groups had similar frequency of ILD and PAH. Patients with SSc overlap had higher frequency of multiple gastrointestinal manifestations and cutaneous disease than those with SSc only or SSc–MCTD. Synovitis was equally common in SSc overlap and SSc–MCTD groups, although puffy digits were more common in those with SSc–MCTD. Myositis was equally common in those with SSc–MCTD and SSc overlap. Patients with SSc overlap or SSc–MCTD were significantly more likely to be exposed to a range of immunosuppressive medications, including prednisolone, than those with SSc only, with the most commonly used DMARDs being hydroxychloroquine and methotrexate. This increased frequency of immunomodulatory and immunosuppressive therapies likely reflects a greater frequency of inflammatory manifestations (e.g., synovitis and myositis) in the SSc–MCTD and SSc overlap groups than in the SSc only group.

In terms of survival, scleroderma-specific antibodies were a more reliable indicator of survival than disease groups. ANA centromere or anti-RNP conferred consistently better survival

Table 3. Multivariable hazards models for mortality*

Variable	Model with disease group (n = 1,555)			Model with autoantibodies (n = 1,094)		
	HR	95% CI	*P*	HR	95% CI	*P*
Sex						
Female	1	–		1	–	
Male	1.85	1.33–2.56	<0.001	1.72	1.10–2.68	0.018
Age at disease onset	1.10	1.09–1.11	<0.001	1.12	1.10–1.14	<0.001
Disease group						
SSc only	1	–				
SSc–MCTD	0.42	0.15–1.14	0.090			
SSc overlap	0.89	0.47–1.69	0.718			
Antibody						
ANA centromere				1	–	
Anti-RNP				0.85	0.33–2.19	0.740
Anti–Scl-70				2.16	1.28–3.65	0.004
Anti–RNA polymerase III				1.22	0.66–2.25	0.529
Interstitial lung disease						
No	1	–		1	–	
Mild/moderate	1.66	1.25–2.22	0.001	1.39	0.91–2.14	0.132
Severe	5.23	3.47–7.89	<0.001	4.14	2.16–7.95	<0.001
Pulmonary arterial hypertension						
No	1	–		1	–	
Yes	2.99	2.26–3.94	<0.001	3.26	2.30–4.63	<0.001
Renal crisis						
No	1	–		1	–	
Yes	1.42	0.79–2.54	0.237	1.14	0.45–2.93	0.775
Myocardial involvement						
No	1	–		1	–	
Yes	1.08	0.72–1.63	0.696	1.36	0.82–2.27	0.234

* 95% CI = 95% confidence interval; ANA = antinuclear antibody; HR = hazard ratio; MCTD = mixed connective tissue disease; SSc = systemic sclerosis.

than anti–Scl-70 or anti–RNA polymerase III, while disease groups were not associated with consistent differences in survival. Despite no difference in prognosis between groups in a multivariable model accounting for younger age of patients with SSc–MCTD, anti-RNP positivity continued to confer a survival benefit. Furthermore, absence of any SSc-specific antibody was associated with worse prognosis than ANA centromere or RNP positivity. Despite similar severity of ILD, patients with SSc–MCTD and SSc overlap had consistently lower all-cause mortality following diagnosis of PAH or ILD than patients with SSc only. It may be that this difference is related to a lower frequency of diffuse disease in those with SSc overlap and SSc–MCTD or a protective effect of particular antibodies, e.g., anti-RNP or ANA centromere. Other potential explanations include different pathogenic mechanisms or greater exposure to immunosuppressive medications.

In the wider literature, infrequent and inconsistent data describe features of patients with SSc overlap, and there is a paucity of data about patients with SSc–MCTD. While one study supports our finding that patients with SSc overlap, in addition to SSc–MCTD, were more likely to have limited skin involvement (2), another showed that those with SSc and myositis overlap were more likely to have diffuse skin disease (7). We identified a similar frequency of PAH and ILD in those with SSc–MCTD and SSc overlap. Other data support a similar frequency of PAH in those with SSc overlap and limited cutaneous SSc (2), albeit with a higher risk of ILD in SSc overlap than in limited cutaneous SSc (2) or SSc in general (7). Data in the wider literature consistently report higher rates of myositis (2) and arthritis/synovitis (7,8) in those with SSc overlap than SSc only, as was the case in our study. Furthermore, multiple studies in patients with SSc overlap (2,7) support our finding that patients with SSc overlap and SSc–MCTD are more likely to have >1 detectable autoantibody compared to those with SSc only.

Multiple studies have shown lower mortality in SSc patients with ANA centromere antibody positivity compared to those with anti–Scl-70 (9,10). In another cohort of patients with SSc, improved survival was demonstrated in those with anti-RNP or anticentromere antibody than in those with anti–Scl-70 (11). Interestingly, unlike in our study, this study did not demonstrate significantly worse survival in those with anti–RNA polymerase III positivity (11). Once ILD or PAH was diagnosed, patients with SSc–MCTD or SSc overlap had a better prognosis than those with SSc only, despite a similar severity of ILD. In the wider literature, disease subtype has not been shown to impact survival in those with PAH, suggesting that PAH is the most important factor (12). However, in our data, patients with SSc–MCTD had a lower all-cause mortality following ILD diagnosis.

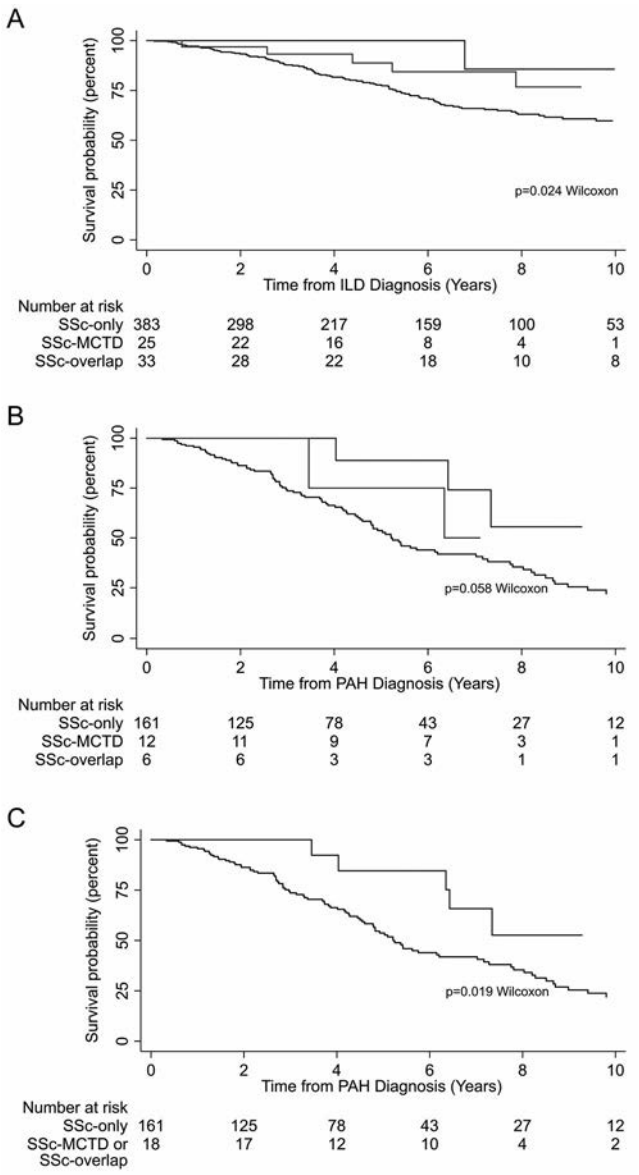


Figure 2. **A**, Survival by diagnosis, according to Kaplan-Meier survival estimates, of interstitial lung disease (ILD) for the systemic sclerosis (SSc) only group (bottom line), the SSc–mixed connective tissue disease (MCTD) group (top line), and the SSc overlap group (middle line). **B**, Survival from diagnosis of pulmonary arterial hypertension (PAH) for the SSc only group (bottom line), the SSc–MCTD group (top line), and the SSc overlap group (middle line). **C**, Survival from diagnosis of PAH for the SSc only group (bottom line) and the SSc–MCTD or SSc overlap group (top line).

MCTD is a controversial entity that some argue is a disease defined by an antibody (13). Among those who fulfill classification criteria for SSc, the diagnostic label of MCTD has limited usefulness beyond prognostic significance of anti-RNP relative to other SSc-specific antibodies. We would suggest that the most important step is identifying anti-RNP positivity rather than making an additional diagnosis of MCTD. Not all patients with anti-RNP positivity meet criteria for MCTD; we have identified a small number of patients in our cohort with SSc who are anti-RNP positive but did not fulfill criteria for MCTD.

To our knowledge, this is the largest study to investigate clinical features of patients with both SSc–MCTD and SSc overlap. Our study includes a comprehensive analysis of disease features, serologic profile, and survival using prospectively collected data. However, our study does have limitations. We did not have scope within our study to investigate patients with MCTD who do not fulfill criteria for SSc, as our database includes only those with SSc. Furthermore, on average, patients in our study were recruited >10 years after diagnosis of their disease, which may mean that there is a degree of survivor bias, as those with more aggressive disease and early mortality are less likely to have survived to be recruited into our study. This is likely to underestimate differences in survival between patients with SSc–MCTD and SSc only. While data were collected prospectively, analysis was performed retrospectively. Furthermore, while the study overall included a large number of patients, SSc–MCTD was relatively uncommon, leading to small numbers in some subgroup analyses.

In conclusion, this study reveals significant differences between patients with SSc–MCTD, SSc overlap, and SSc only. We have identified a number of similarities between patients with SSc overlap and SSc–MCTD, including prognosis and frequency of PAH, ILD, myositis, synovitis, and autoantibody positivity. Furthermore, this study highlights the critical importance of antibody profile in determining prognosis, which has greater accuracy than profile by disease group. Patients with anti-RNP positivity display better long-term survival than those with anti–Scl-70 or anti–RNA polymerase III positivity. Furthermore, patients with SSc–MCTD and SSc overlap had better survival following ILD or PAH diagnosis, despite similar severity. These data suggest that testing for antibody to RNP is a valuable prognostic tool in patients with SSc. Whether a diagnostic label of SSc–MCTD or SSc–RNP is more appropriate in this setting is a point of contention. Regardless, this group of patients has a distinct phenotypic profile. While we did not have a consensus or a priori definition of SSc overlap, it was clear that treating physicians were indeed able to identify a group of patients with SSc with overlap features who also had distinct clinical features and outcomes that differed significantly from those with SSc only. Furthermore, the study may assist with determining risk of specific organ manifestations in patients with SSc, particularly if an overlap syndrome is suspected. Further data are required to better understand these patients' conditions.

ACKNOWLEDGMENTS

We thank the following members of the Australian Scleroderma Interest Group for being collaborators in the ASCS: Helen Cooley (Hobart Private Hospital, Hobart, Tasmania), Lucy Croyle (Monash Health, Melbourne, Victoria), Catherine Hill (University of Adelaide, South Australia), Lauren Host (Fiona Stanley Hospital, Perth, Western Australia), Sue Lester (University of Adelaide, South Australia), Peter Nash (University of Queensland, Sunshine Coast, Queensland), Maureen Rischmueller (University of Adelaide, South Australia),

Tien Tay (Westmead Hospital, Sydney, New South Wales), Kathleen Tymms (Australian National University, Canberra, Australian Capital Territory), and Peter Youssef (University of Sydney, New South Wales).

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. Nikpour 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. Fairley, Proudman, Wilson, Morrisroe, Ferdowsi, Stevens, Nikpour.

Acquisition of data. Stevens, Nikpour.

Analysis and interpretation of data. Fairley, Hansen, Proudman, Sahhar, Ngian, Walker, Strickland, Wilson, Morrisroe, Ferdowsi, Major, Roddy, Nikpour.

REFERENCES

1. Sharp GC, Irvin WS, Tan EM, Gould RG, Holman HR. Mixed connective tissue disease: an apparently distinct rheumatic disease syndrome associated with a specific antibody to an extractable nuclear antigen (ENA). Am J Med 1972;52:148–59.
2. Moinzadeh P, Aberer E, Ahmadi-Simab K, Blank N, Distler JH, Fierlbeck G, et al. Disease progression in systemic sclerosis-overlap syndrome is significantly different from limited and diffuse cutaneous systemic sclerosis. Ann Rheum Dis 2015;74:730–7.
3. Alarcon-Segovia D, Cardiel MH. Comparison between 3 diagnostic criteria for mixed connective tissue disease: study of 593 patients. J Rheumatol 1989;16:328–34.
4. Jablonska S, Blaszczyk M. Scleroderma overlap syndromes. Adv Exp Med Biol 1999;455:85–92.
5. Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
6. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15:202–5.
7. Balbir-Gurman A, Braun-Moscovici Y. Scleroderma overlap syndrome. Isr Med Assoc J 2011;13:14–20.
8. Avouac J, Clements PJ, Khanna D, Furst DE, Allanore Y. Articular involvement in systemic sclerosis. Rheumatology (Oxford) 2012;51:1347–56.
9. Scussel-Lonzetti L, Joyal F, Raynauld JP, Roussin A, Rich E, Goulet JR, et al. Predicting mortality in systemic sclerosis: analysis of a cohort of 309 French Canadian patients with emphasis on features at diagnosis as predictive factors for survival. Medicine (Baltimore) 2002;81:154–67.
10. Ferri C, Valentini G, Cozzi F, Sebastiani M, Michelassi C, La Montagna G, et al. Systemic sclerosis: demographic, clinical, and serologic features and survival in 1,012 Italian patients. Medicine (Baltimore) 2002;81:139–53.
11. Hesselstrand R, Scheja A, Shen GQ, Wiik A, Akesson A. The association of antinuclear antibodies with organ involvement and survival in systemic sclerosis. Rheumatology (Oxford) 2003;42:534–40.
12. Morrisroe K, Stevens W, Huq M, Prior D, Sahhar J, Ngian GS, et al. Survival and quality of life in incident systemic sclerosis-related pulmonary arterial hypertension. Arthritis Res Ther 2017;19:122.
13. Jury EC, D'Cruz D, Morrow WJ. Autoantibodies and overlap syndromes in autoimmune rheumatic disease. J Clin Pathol 2001;54:340–7.

APPENDIX A: THE AUSTRALIAN SCLERODERMA INTEREST GROUP

Members of the Australian Scleroderma Interest Group are as follows: Mandana Nikpour (University of Melbourne, Victoria), Susanna Proudman (University of Adelaide, South Australia), Wendy Stevens (St Vincent's Hospital Melbourne, Victoria), Joanne Sahhar (Monash Health, Melbourne, Victoria), Helen Cooley (Hobart Private Hospital, Hobart, Tasmania), Lucy Croyle (Monash Health, Melbourne, Victoria), Nava Ferdowsi (University of Melbourne, Victoria), Catherine Hill (University of Adelaide, South Australia), Lauren Host (Fiona Stanley Hospital, Perth, Western Australia), Sue Lester (University of Adelaide, South Australia), Gabor Major (Royal Newcastle Centre, New South Wales), Kathleen Morrisroe (University of Melbourne, Victoria), Peter Nash (University of Queensland, Sunshine Coast, Queensland), Gene-Siew Ngian (Monash Health, Melbourne, Victoria), Maureen Rischmueller (University of Adelaide, South Australia), Janet Roddy (Fiona Stanley Hospital, Perth, Western Australia), Gemma Strickland (Barwon Rheumatology Service, Geelong, Victoria), Tien Tay (Westmead Hospital, Sydney, New South Wales), Kathleen Tymms (Australian National University, Canberra, Australian Capital Territory), Jennifer Walker (Flinders University, Adelaide, South Australia), and Peter Youssef (University of Sydney, New South Wales).

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 742–754
DOI 10.1002/acr.24473

Anatomic Distribution of Sacroiliac Joint Lesions on Magnetic Resonance Imaging in Patients With Axial Spondyloarthritis and Control Subjects: A Prospective Cross-Sectional Study, Including Postpartum Women, Patients With Disc Herniation, Cleaning Staff, Runners, and Healthy Individuals

Sengül Seven,[1] Mikkel Østergaard,[1] Lone Morsel-Carlsen,[2] Inge J. Sørensen,[3] Birthe Bonde,[4] Gorm Thamsborg,[3] Jens J. Lykkegaard,[3] and Susanne J. Pedersen[1]

Objective. To investigate the anatomic location and distribution of lesions on magnetic resonance imaging (MRI) in the sacroiliac (SI) joints in patients with axial spondyloarthritis (SpA), women with and without postpartum pain (childbirth within 4–16 months), patients with disc herniation, cleaning staff, runners, and healthy persons.

Methods. In a prospective cross-sectional study of 204 participants, MRI of the entire cartilaginous compartment of the SI joint was scored blindly by 2 independent, experienced readers, according to Spondyloarthritis Research Consortium of Canada definitions of SI joint inflammation and structural lesions in each SI joint quadrant or half and in each of 9 slices. The locations of the lesions (unilateral/bilateral, upper/lower, sacral/iliac, and anterior/central/posterior slices) were analyzed based on concordant reads.

Results. Bone marrow edema (BME) occurred in all quadrants in nearly all participant groups, but rarely bilaterally, except in patients with axial SpA and women with postpartum pain. Fat lesions were mainly found in axial SpA and occurred in all quadrants, but mostly bilaterally in sacral quadrants. Erosion was rare, except in axial SpA, where it was mainly iliac and often bilateral. Sclerosis was exclusively iliac and most frequent in women with postpartum pain.

Conclusion. The location and distribution of common SI joint lesions in axial SpA and non–axial SpA were reported, and group-specific patterns were revealed. BME distributed bilaterally or unilaterally, both locally and more widespread in the SI joint, is common in both postpartum women with pain and axial SpA patients, which limits the use of BME to differentiate these groups. This study indicates that the presence of fat lesions, especially when widespread, and/or erosion, particularly when located centrally or posteriorly, are diagnostically important and should be investigated further.

INTRODUCTION

The presence of bone marrow edema (BME) on magnetic resonance imaging (MRI) of the sacroiliac (SI) joints is pivotal in the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axial spondyloarthritis (SpA) (1–3). However, differentiating axial SpA from other conditions remains a challenge because the presence of BME has been reported in a high frequency of individuals with nonspecific back pain (6–23%) (4–6), in pregnant and postpartum women

Supported by Rigshospitalet and the Danish Rheumatism Association (study grants R102-A2132-B98 and R150-A4467-B98). Dr. Seven is recipient of a PhD grant from the Danish Rheumatism Association (grant R159-A5061).

[1]Sengül Seven, MD, Mikkel Østergaard, MD, PhD, DMSc, Susanne J. Pedersen, MD, PhD: Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Center of Head and Orthopedics, Rigshospitalet, Glostrup, and University of Copenhagen, Copenhagen, Denmark; [2]Lone Morsel-Carlsen, MD: Rigshospitalet, Glostrup, Denmark; [3]Inge J. Sørensen, MD, PhD, Gorm Thamsborg, MD, DMSc, Jens J. Lykkegaard, MD: Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Center of Head and Orthopedics, Rigshospitalet, Glostrup, Denmark; [4]Birthe Bonde, PT: The Birthe Bonde Clinic of Physiotherapy, Copenhagen, Denmark.

Dr. Seven has received consulting fees, speaking fees, and/or honoraria from Novartis, UCB, Sanofi, and AbbVie (less than $10,000 each). Dr. Østergaard has received consulting fees and/or speaking fees from AbbVie, Bristol Myers Squibb, Boehringer Ingelheim, Celgene, Eli Lilly and Company, Hospira, Janssen, Merck, Novartis, Novo, Orion, Pfizer, Regeneron, Roche, Sandoz, Sanofi, and UCB (less than $10,000 each) and research support from those companies. Dr. Pedersen has received consulting fees and/or speaking fees from MSD, Pfizer, AbbVie, Novartis, UCB (less than $10,000 each) and Novartis (more than $10,000) and research support from AbbVie, MSD, and Novartis. No other disclosures relevant to this article were reported.

Address correspondence to Sengül Seven, MD, Rigshospitalet, Glostrup, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Center of Head and Orthopedics, Valdemar Hansens Vej 17, 2600 Glostrup, Denmark. Email: Sengul.seven@dadlnet.dk.

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

SIGNIFICANCE & INNOVATIONS

- This is the first study to investigate in detail the anatomic location and distribution of typical sacroiliac (SI) joint lesions on magnetic resonance imaging (MRI) in both patients with axial spondyloarthritis (SpA) and various control groups.
- Bilateral and unilateral bone marrow edema on MRI was frequently seen in the axial SpA group and in women with postpartum pain, while unilateral lesions occasionally were seen in all other groups.
- Widespread fat lesions distributed unilaterally or bilaterally were only seen in patients with axial SpA. In non-axial SpA groups, erosion was rare, and backfill and ankylosis were absent.
- The location and distribution of SI joint inflammatory and structural lesions may be helpful in diagnosing axial SpA.

(21–63%) (5–9), and healthy persons (0–23%) (4,6,9), including athletes (4–41%) (6,9,10), military recruits (23%) (11), and the general population (17%) (12). Although not required in ASAS definitions of SpA on MRI, the presence of concomitant structural lesions may contribute to diagnosing axial SpA (2). Several studies have examined the presence of inflammatory (4–8,10–13) and structural (4,5,7–14) SI joint features on MRI, such as erosion, fat lesions, sclerosis, backfill, and ankylosis, either individually or in various combinations (14,15), in axial SpA and/or in different conditions. In a recent publication (9), we investigated the diagnostic utility of BME and different structural lesions (i.e., fat lesions, erosion, backfill, and ankylosis) in SI joints identified by MRI according to the Spondyloarthritis Research Consortium of Canada (SPARCC) method (16,17) for differentiating patients with axial SpA from control subjects with or without buttock or pelvic pain. We found that BME and fat lesions were most frequent in patients with axial SpA but also present in other groups of control subjects, especially women with postpartum pain. On the other hand, erosion scores above a certain threshold and presence of backfill and ankylosis were highly specific for axial SpA. Although previous studies of certain non–axial SpA groups have suggested that BME is more frequent in certain locations (e.g., posterior lower ilium and the anterior upper sacrum in athletes [10], posterior lower ilium of healthy participants [6], and the lower ilium and sacrum of peripartum women [8]), the importance of the location of the different lesions on MRI has not been investigated in a single, larger study including patients with axial SpA, postpartum women, and several groups of controls.

In this study, we aimed to provide a detailed description of the anatomic location and distribution of BME and structural lesions identified by MRI based on lesions being present unilaterally or bilaterally in the upper, lower sacral, and iliac quadrants, and in the anterior, central, and posterior sections of the SI joints of patients with axial SpA, in women with and without postpartum pain, and in patients with disc herniation, cleaning staff, runners, and healthy individuals.

SUBJECTS AND METHODS

Subjects. This prospective cross-sectional study (a scientific investigation of MRI and biochemical markers in patients with axial SpA, back pain for other reasons, subjects with strain on the SI joints, and healthy subjects) was conducted during 2013–2016 at Rigshospitalet, Glostrup, in the Capital Region of Copenhagen, Denmark. The study protocol was approved by the local ethics committee (approval number H-17034960) and conducted according to Danish legislation and the Declaration of Helsinki. Written informed consent was provided by all participants prior to study inclusion.

A total of 204 participants were included, of which 41 were patients with axial SpA (1) with active disease (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24473/abstract). Sixty participants were women who had given birth 4–16 months prior to study inclusion. Of these, 46 had persistent (i.e., ongoing) postpartum buttock or pelvic pain following pregnancy and/or vaginal birth, while 14 were without buttock or pelvic pain related to pregnancy and/or delivery and in the postpartum period. Additionally, the control subjects comprised 25 patients with lumbar disc herniation, 26 individuals who engaged in hard physical work (hospital cleaning staff), 23 long-distance runners, and 29 healthy men. Post hoc, we defined a subgroup of 38 women with previous childbirths within the groups of disc herniation, cleaning staff, and long-distance runners. Participants were included according to common and group-specific inclusion and exclusion criteria as previously described in detail (9) (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24473/abstract). Information on demographic and clinical characteristics was collected together with information on past medical history. All participants were assessed by a physician for fulfilling the ASAS classification criteria for axial SpA, MRI of the SI joints was performed, and C-reactive protein level and HLA–B27 status were determined.

MRI methodology. MRI of the SI joints was performed at Rigshospitalet, Glostrup, on a 1.5T MRI Avanto scanner (Siemens), version Syngo MR B17, with Numaris/4 software. The acquired images included a semicoronal short tau inversion recovery (STIR) sequence with repetition time (TR) 4,000 msec, time to inversion 160 msec, echo time (TE) 37 msec, slice thickness 4 mm, gap 0.4 mm, field of view (FOV) 26 × 26 cm, and matrix size 205 × 256, and a semicoronal T1-weighted sequence with TR 660 msec, TE 11 msec, slice thickness 4 mm, gap 0.3 mm, FOV 23 × 23 cm, and matrix size 320 × 256. T1-weighted and STIR images were evaluated simultaneously in an anonymized manner and in random order by 2 experienced, independent readers (an experienced rheumatologist and a radiologist), who were blinded to all clinical and biochemical data.

MRI reads. The complete cartilaginous compartment was covered in 9 MRI slices. The SI joints were systematically assessed for the presence of inflammatory and structural lesions on each SI joint quadrant at each slice. The 1st slice (slice 1) was the most anterior slice where the joint space and at least the iliac bone was visible, and the 9th slice (slice 9) was the most posterior slice where the cartilaginous joint was visible (see Figure 1 for the location of slices 1–9 of the SI joint and Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24473/abstract, for the corresponding MRI slices). The number of subjects in whom the cartilaginous compartment required all 9 slices to be covered was 143 (70%), whereas 2 (1%), 50 (25%), and 9 (4%) participants had the cartilaginous compartment covered by 7, 8, and 10 slices, respectively, resulting in a mean ± SD (range) number of slices of 9.0 ± 0.5 (8–10) for all male participants and 8.7 ± 0.5 (7–9) for all female participants. No participants with a 10th slice had lesions in this slice. The definition of lesions on MRI according to the SPARCC SI Joint Inflammation Index (17) and SPARCC Sacroiliac Joint Structural scores (16), as well as Morpho definitions (4), were applied, while sclerosis by MRI was defined as a hypointense signal extending at least 5 mm perpendicular to the joint space on both MRI sequences (18). Inflammation, fat lesions, sclerosis, and erosion were scored per quadrant per slice (0 if absent, and 1 if present), and backfill and ankylosis were scored per joint half per slice. Scores were based on concordant reads, i.e., when the 2 readers agreed that lesions were distributed in certain predefined patterns. The investigated patterns were as follows: 1) unilaterally on the same side versus bilaterally; 2) overall in the entire joint; 3) in each quadrant (upper and lower sacrum or ilium) or joint half (upper and lower); and 4) in the same joint section (anterior, central, or posterior) in the same joint quadrant or joint half.

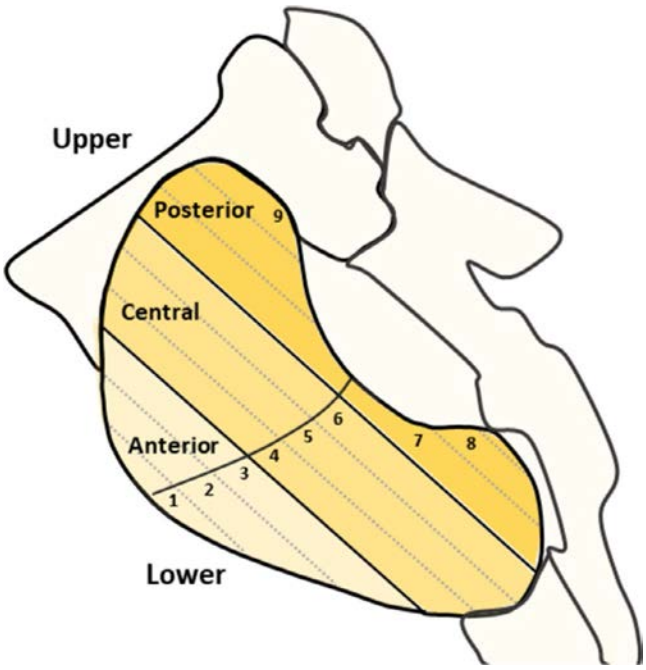


Figure 1. Schematic of the location of slices 1–9 of the sacroiliac joint.

Statistical analysis. SPSS, version 22.0, was used to perform statistical analyses. Anatomic location of the lesions was analyzed on SI joint level, quadrant level, and per slice level. The analyses included the anatomic location of the lesion in 1 or both joint halves (unilateral versus bilateral), per quadrant (upper ilium and sacrum, lower ilium and sacrum, respectively), and in the joint at the anterior section (slices 1–3), the central section (slices 4–6), and the posterior section (slices 7–9). Patient characteristics and clinical, biochemical, and MRI data were characterized by descriptive statistics. A Mann-Whitney U test was applied to compare patients with axial SpA with the other participant groups. The interreader reliability for evaluation of lesions on MRI was assessed by intraclass correlation coefficient (ICC) based on a 2-way, random effects, single-measure model, and the absolute agreements are presented. Values from 0.0–0.2 were considered poor agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 good agreement, and 0.81–1.00 very good agreement (19). The primary analysis was based on reader agreement on the presence of individual lesions (concordant reads). When the calculation of sample size was performed during the planning of the study, it was estimated that a sample size of 200 participants would suffice to reveal differences between patients with axial SpA and other groups by using a nomogram for power calculation in diagnostic studies (20). *P* values less than 0.05 were considered significant.

RESULTS

Subjects and reliability of lesion score. Participant demographics and clinical and biochemical characteristics are presented in Table 1 and have previously been described (9). Table 2 shows lesion scores presented as total scores, joint quadrant scores, and joint half scores for the different participant groups as the mean for the 2 readers, as well as the ICC for the different lesion scores. Overall, there was good or very good agreement between the 2 readers, apart from sclerosis and backfill, which had very low scores.

Unilateral versus bilateral lesions. Table 3 provides the proportion of participants in each group with unilateral versus bilateral SI joint lesions on MRI detected concordantly by the 2 readers. Bilateral SI joint lesions on MRI were primarily found in the axial SpA group (bilateral versus unilateral BME 44% versus 22%, fat lesions 59% versus 17%, sclerosis 2% versus 7%, erosion 34% versus 27%) and in women with postpartum pain (bilateral

Table 1. Demographic, clinical, and biochemical characteristics of the different groups of study participants*

Characteristic	Axial SpA	Women with postpartum pain	Women without postpartum pain	Disc herniation	Cleaning staff	Long-distance runners	Healthy men	Women with ≥1 childbirth†
Total no.	41	46	14	25	26	23	29	38
Age, mean ± SD years	30.9 ± 6.4	32.6 ± 3.3	33.1 ± 4.1	35.2 ± 5.7	39.1 ± 4.6	32.7 ± 6.2	30.9 ± 6.4	38.7 ± 4.4
Median (range)	30.0 (19–44)	32.5 (26–41)‡	32.5 (27–41)	37.0 (21–43)§	39.0 (28–45)¶	32.0 (22–43)	30.0 (20–45)	39.0 (27–45)¶
Male sex	26 (63.4)	0¶	0¶	11 (44.0)	0¶	18 (78.3)	29 (100.0)¶	0¶
No. of childbirths if woman	1.7 (0.8)	1.5 (0.8)	1.9 (0.8)	1.6 (0.9)	2.5 (1.1)	0.5 (1.0)	NA	2.4 (0.9)
Median (range)	2 (0–2)	1 (1–4)	2 (1–3)	2 (0–3)	3 (0–5)	0 (0–2)	NA	2 (1–5)
Time since last childbirth if woman, years	4.9 (4.6)	0.7 (0.3)	0.8 (0.3)	9.1 (7.0)	10.5 (6.3)	5.7	NA	9.7 (6.4)
Median (range)	4.7 (0.8–9.3)	0.7 (0.3–1.3)§	0.9 (0.3–1.1)‡	6.6 (2.0–21.5)	10.3 (1.7–22.3)	5.7 (5.7–5.7)	NA	10.0 (1.7–22.3)
Symptom duration, years	8.4 (5.6)	1.0 (0.8)	NA	1.0 (0.9)	NA	NA	NA	NA
Median (range)	8.4 (1.2–23.8)	(0.3–6.0)¶	NA	0.7 (0.2–3.6)¶	NA	NA	NA	NA
Low back pain VAS score (range 0–10)	3.8 (2.8)	5.5 (2.4)	0.4 (0.7)	5.5 (2.4)	0.8 (1.8)	0.2 (0.5)	0.1 (0.3)	2.4 (3.2)
Median (range)	3.7 (0–10.0)	6.0 (0–9.8)§	0 (0–1.9)¶	6.2 (0.3–9.6)‡	0 (0–6.8)¶	0 (0–1.5)¶	0 (0–1.2)¶	0 (0–9.6)§
HLA–B27 positive	33 (80.5)	5 (10.9)¶	1 (7.1)¶	0¶	0¶	1 (4.3)¶	4 (13.8)¶	0¶
CRP >3 mg/liter	24 (58.5)	8 (17.4)¶	3 (21.4)§	5 (20.0)§	4 (15.4)§	4 (17.4)§	1 (3.4)¶	7 (18.4)¶

* Values are the number (%) unless indicated otherwise. A Mann-Whitney test was applied, and all tests were for patients with axial SpA compared with the other groups. CRP = C-reactive protein; NA = not applicable; SpA = spondyloarthritis; VAS = visual analog scale.
† From disc herniation, cleaning staff, and runner groups. The mean time since last delivery was 9.7 years (range 1.7–22.3). A total of 10 patients (26.3%) had their last delivery <5 years prior, 10 (26.3%) 5–10 years prior, and 18 (47.4%) >10 years prior.
‡ $P < 0.05$.
§ $P < 0.01$.
¶ $P < 0.001$.

Table 2. Sacroiliac joint total on magnetic resonance imaging (MRI), joint quadrant and joint half scores, and MRI interreader reliability*

	Axial SpA	Women with postpartum pain	Women without postpartum pain	Disc herniation	Cleaning staff	Long-distance runners	Healthy men	Women with ≥1 childbirth†	Interreader ICC (95% CI)
Total no.	41	46	14	25	26	23	29	38	204
Bone marrow edema									
Total score (range 0–72)	12.7 ± 13.4	5.4 ± 8.5	2.1 ± 2.9	0.5 ± 1.2	0.6 ± 1.8	0.3 ± 0.7	0.1 ± 0.4	0.5 ± 1.5	0.91 (0.88, 0.93)
Median (range)	10.0 (0–48.5)	1.0 (0–35.0)‡	0.3 (0–10.0)‡	0 (0–6.0)§	0 (0–9.0)§	0 (0–3.0)§	0 (0–2.0)§	0 (0–9.0)§	
Upper ilium (range 0–18)	2.8 ± 4.3	1.2 ± 2.3	0.2 ± 0.3	0 ± 0	0.1 ± 0.5	0 ± 0	0 ± 0	0.1 ± 0.4	0.89 (0.85, 0.91)
Median (range)	0.5 (0–18.0)	0 (0–10.0)‡	0 (0–1.0)¶	0 (0–0)§	0 (0–2.5)§	0 (0–0)§	0 (0–0)§	0 (0–2.5)§	
Lower ilium (range 0–18)	3.5 ± 4.3	2.2 ± 3.5	0.9 ± 1.2	0.1 ± 0.3	0.3 ± 1.1	0.1 ± 0.4	0.1 ± 0.3	0.2 ± 0.9	0.89 (0.85, 0.91)
Median (range)	1.5 (0–17.5)	0 (0–14.0)	0 (0–3.5)	0 (0–1.0)§	0 (0–5.5)§	0 (0–1.5)§	0 (0–1.5)§	0 (0–5.5)§	
Upper sacrum (range 0–18)	3.4 ± 4.0	1.2 ± 1.8	0.6 ± 1.0	0.2 ± 0.4	0.04 ± 0.2	0.2 ± 0.4	0.02 ± 0.1	0.1 ± 0.2	0.89 (0.85, 0.91)
Median (range)	2.0 (0–13.0)	0 (0–6.0)¶	0 (0–2.5)‡	0 (0–1.5)§	0 (0–1.0)§	0 (0–1.5)§	0 (0–0.5)§	0 (0–1.0)§	
Lower sacrum (range 0–18)	3.0 ± 3.2	0.8 ± 1.4	0.4 ± 0.9	0.2 ± 0.6	0.1 ± 0.3	0.04 ± 0.2	0 ± 0	0.1 ± 0.3	0.82 (0.77, 0.86)
Median (range)	2.0 (0–11.5)	0 (0–6.5)§	0 (0–3.5)‡	0 (0–3.0)§	0 (0–1.0)§	0 (0–1.0)§	0 (0–0)§	0 (0–1.0)§	
Fat lesion									
Total score (range 0–72)	17.5 ± 15.9	0.7 ± 2.3	1.0 ± 2.9	0.5 ± 1.2	0 ± 0	0.7 ± 2.2	1.6 ± 6.0	0.1 ± 0.6	0.95 (0.93, 0.96)
Median (range)	15.0 (0–64.0)	0 (0–12.5)§	0 (0–11.0)§	0 (0–5.0)§	0 (0–0)§	0 (0–9.0)§	0 (0–31.0)§	0 (0–4.0)§	
Upper ilium (range 0–18)	3.0 ± 4.2	0.1 ± 0.3	0.1 ± 0.3	0.1 ± 0.4	0 ± 0	0 ± 0	0.7 ± 2.5	0.1 ± 0.3	0.94 (0.92, 0.95)
Median (range)	1.0 (0–16.0)	0 (0–1.5)§	0 (0–1.0)§	0 (0–2.0)§	0 (0–0)§	0 (0–0)§	0 (0–11.0)§	0 (0–2.0)§	
Lower ilium (range 0–18)	4.2 ± 4.7	0.1 ± 0.3	0.1 ± 0.3	0.1 ± 0.3	0 ± 0	0 ± 0	0.5 ± 2.0	0.04 ± 0.2	0.87 (0.83, 0.90)
Median (range)	2.5 (0–16.0)	0 (0–1.5)§	0 (0–1.0)§	0 (0–1.5)§	0 (0–0)§	0 (0–0)§	0 (0–10.0)§	0 (0–1.5)§	
Upper sacrum (range 0–18)	5.5 ± 4.8	0.4 ± 1.5	0.5 ± 1.5	0.3 ± 1.0	0 ± 0	0.5 ± 1.5	0.3 ± 1.2	0.01 ± 0.1	0.91 (0.88, 0.93)
Median (range)	5.5 (0–16.0)	0 (0–8.5)§	0 (0–5.5)§	0 (0–4.0)§	0 (0–0)§	0 (0–6.0)§	0 (0–6.5)§	0 (0–0.5)§	
Lower sacrum (range 0–18)	4.8 ± 4.5	0.2 ± 0.7	0.3 ± 0.9	0.02 ± 0.1	0 ± 0	0.2 ± 0.7	0.1 ± 0.6	0 ± 0	0.92 (0.90, 0.94)
Median (range)	3.0 (0–16.0)	0 (0–4.0)§	0 (0–3.0)§	0 (0–0.5)§	0 (0–0)§	0 (0–2.5)§	0 (0–3.0)§	0 (0–0)§	
Sclerosis									
Total score (range 0–72)	2.6 ± 3.3	3.3 ± 5.7	2.5 ± 3.6	0.8 ± 1.8	1.9 ± 4.0	0.9 ± 1.7	1.7 ± 2.9	1.5 ± 3.4	0.36 (0.14, 0.52)
Median (range)	0.5 (0–10.5)	1.0 (0–24.0)	0.8 (0–12.0)	0 (0–8.0)¶	0.3 (0–20.0)	0 (0–6.0)¶	0 (0–12.0)	0 (0–20.0)	
Upper ilium (range 0–18)	1.8 ± 2.3	1.2 ± 2.5	1.5 ± 2.8	0.5 ± 1.1	1.3 ± 2.5	0.8 ± 1.5	1.5 ± 2.5	1.0 ± 2.2	0.29 (0.1, 0.45)
Median (range)	0 (0–8.5)	0 (0–13.5)	0.3 (0–8.0)	0 (0–3.5)¶	0 (0–11.5)	0 (0–4)	0 (0–10.0)	0 (0–11.5)	
Lower ilium (range 0–18)	0.8 ± 1.4	1.9 ± 3.1	1.0 ± 1.5	0.3 ± 1.0	0.6 ± 1.7	0.1 ± 0.4	0.1 ± 0.4	0.5 ± 1.4	0.50 (0.34, 0.62)
Median (range)	0 (0–5.0)	0 (0–13.0)	0 (0–4.5)	0 (0–4.5)	0 (0–8.0)	0 (0–2.0)¶	0 (0–2.0)‡	0 (0–8.0)	
Upper sacrum (range 0–18)	0 ± 0	0.1 ± 0.4	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.1 ± 0.4	0 ± 0	0.00 (–0.14, 0.14)
Median (range)	0 (0–0)	0 (0–2.5)	0 (0–0)	0 (0–0)	0 (0–0)	0 (0–0)	0 (0–2.0)‡	0 (0–0)	
Lower sacrum (range 0–18)	0.04 ± 0.2	0.2 ± 1.0	0 ± 0	0.02 ± 0.1	0 ± 0	0 ± 0	0 ± 0	0.01 ± 0.1	0.00 (–0.14, 0.14)
Median (range)	0 (0–1.5)	0 (0–6.5)	0 (0–0)	0 (0–0.5)	0 (0–0)	0 (0–0)	0 (0–0)	0 (0–0.5)	
Erosion									
Total score (range 0–72)	5.8 ± 5.7	0.7 ± 2.6	0.1 ± 0.2	0.1 ± 0.2	0 ± 0	0.1 ± 0.2	0.1 ± 0.3	0.03 ± 0.1	0.85 (0.81, 0.89)
Median (range)	3.5 (0–19.0)	0 (0–17.0)§	0 (0–1.0)§	0 (0–1.0)§	0 (0–0)§	0 (0–1.0)§	0 (0–1.0)§	0 (0–1.0)§	
Upper ilium (range 0–18)	2.3 ± 2.8	0.1 ± 0.4	0.04 ± 0.1	0.1 ± 0.2	0 ± 0	0 ± 0	0.1 ± 0.2	0.03 ± 0.1	0.79 (0.73, 0.83)
Median (range)	1.5 (0–12.0)	0 (0–2.0)§	0 (0–0.5)§	0 (0–0.5)§	0 (0–0)§	0 (0–0)§	0 (0–0.5)§	0 (0–0.5)§	
Lower ilium (range 0–18)	2.2 ± 2.3	0.4 ± 1.7	0.04 ± 0	0 ± 0	0 ± 0	0 ± 0	0.02 ± 0.1	0 ± 0	0.86 (0.82, 0.89)
Median (range)	1.5 (0–8.0)	0 (0–11.5)§	0 (0–0.5)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0.5)§	0 (0–0)§	
Upper sacrum (range 0–18)	0.6 ± 1.2	0.0 ± 0.2	0.04 ± 0.1	0 ± 0	0 ± 0	0.04 ± 0.2	0.02 ± 0.1	0 ± 0	0.74 (0.66, 0.79)
Median (range)	0 (0–5.0)	0 (0–1.0)§	0 (0–0.5)¶	0 (0–0)§	0 (0–0)§	0 (0–1.0)‡	0 (0–0.5)§	0 (0–0)§	
Lower sacrum (range 0–18)	0.7 ± 1.0	0.1 ± 0.6	0 ± 0	0 ± 0	0 ± 0	0.02 ± 0.1	0 ± 0	0 ± 0	0.61 (0.52, 0.69)
Median (range)	0 (0–4.0)	0 (0–3.5)§	0 (0–0)‡	0 (0–0)§	0 (0–0)§	0 (0–0.5)§	0 (0–0)§	0 (0–0)§	

(Continued)

Table 2. *(Cont'd)*

	Axial SpA	Women with postpartum pain	Women without postpartum pain	Disc herniation	Cleaning staff	Long-distance runners	Healthy men	Women with ≥1 childbirth†	Interreader ICC (95% CI)
Backfill									
Total score (range 0–36)	6.0 ± 7.0	0.01 ± 0.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.65 (0.56, 0.72)
Median (range)	3.0 (0–24.5)	0 (0–0.5)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	
Upper half (range 0–18)	3.1 ± 3.6	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.60 (0.50, 0.69)
Median (range)	2.5 (0–13.5)	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	
Lower half (range 0–18)	2.9 ± 3.6	0.01 ± 0.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.67 (0.58, 0.74)
Median (range)	1.0 (0–11.5)	0 (0–0.5)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	
Ankylosis									
Total score (range 0–36)	9.6 ± 13.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.72 (0.63, 0.78)
Median (range)	4.0 (0–43.5)	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	
Upper half (range 0–18)	4.4 ± 6.6	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.69 (0.60, 0.77)
Median (range)	1.0 (0–23.0)	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	
Lower half (range 0–18)	5.2 ± 6.7	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.71 (0.63, 0.78)
Median (range)	2.0 (0–20.5)	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	0 (0–0)§	

* Values are the mean ± SD unless indicated otherwise. MRI scores are mean scores of the 2 readers based on 9 MRI slices. Bone marrow edema, fat lesion, sclerosis, and erosion scores correspond to the number of sacroiliac joint quadrants on MRI with a lesion present, providing a total score range of 0–72. Backfill and ankylosis scores correspond to the number of sacroiliac joint halves on MRI with a lesion present, as a lesion is recorded per upper and lower joint halves, providing a total score range of 0–36. A Mann-Whitney test was applied, and all tests were for patients with axial spondyloarthritis (SpA) compared with the other groups. Intraclass correlation coefficients (ICCs) are provided as single measures with 95% confidence interval (95% CI).
† From disc herniation, cleaning staff, and runner groups.
‡ $P < 0.01$.
§ $P < 0.001$.
¶ $P < 0.05$.

Table 3. Number and proportion of participants with ≥1 lesion on magnetic resonance imaging (MRI), stratified according to unilateral (UNI) versus bilateral (BI) location, per entire joint, joint quadrant, and joint half*

	Axial SpA (n = 41)		Women with postpartum pain (n = 46)		Women without postpartum pain (n = 14)		Disc herniation (n = 25)		Cleaning staff (n = 26)		Long-distance runners (n = 23)		Healthy men (n = 29)		Women with ≥1 childbirth (n = 38)†	
	UNI	BI	UNI	BI	UNI	BI	UNI	BI	UNI	BI	UNI	BI	UNI	BI	UNI	BI
Bone marrow edema																
Entire SI joint	9 (22.0)	18 (43.9)	10 (21.7)	9 (19.6)	3 (21.4)	2 (14.3)	2 (8.0)	0	1 (3.8)	1 (3.8)	2 (8.7)	0	0	0	2 (5.3)	1 (2.6)
Upper ilium	12 (29.3)	4 (9.8)	5 (10.9)	5 (10.9)	0	0	0	1 (4.0)	0	1 (3.8)	0	0	0	0	0	1 (2.6)
Lower ilium	10 (24.4)	9 (22.0)	7 (15.2)	7 (15.2)	3 (21.4)	0	0	0	0	1 (3.8)	1 (4.3)	0	0	0	0	1 (2.6)
Upper sacrum	12 (29.3)	11 (26.8)	8 (17.4)	5 (10.9)	4 (28.6)	0	2 (4.3)	0	1 (3.8)	0	0	0	0	0	2 (5.3)	0
Lower sacrum	13 (31.7)	8 (19.5)	4 (8.7)	5 (10.9)	2 (4.3)	0	1 (4.0)	0	0	0	0	0	0	0	0	0
Fat lesion																
Entire SI joint	7 (17.1)	24 (58.5)	1 (2.2)	2 (4.3)	1 (7.1)	1 (7.1)	1 (4.0)	1 (4.0)	0	0	1 (4.3)	0	1 (3.4)	2 (6.9)	1 (2.6)	0
Upper ilium	9 (22.0)	8 (19.5)	1 (2.2)	0	1 (7.1)	0	1 (4.0)	0	0	0	0	0	0	2 (6.9)	1 (2.6)	0
Lower ilium	8 (19.5)	13 (31.7)	0	0	1 (7.1)	0	1 (4.0)	0	0	0	0	0	0	1 (3.4)	1 (2.6)	0
Upper sacrum	10 (24.4)	19 (46.3)	1 (2.2)	2 (4.3)	0	1 (7.1)	0	1 (4.0)	0	0	1 (4.3)	0	1 (3.4)	1 (3.4)	0	0
Lower sacrum	8 (19.5)	18 (43.9)	1 (2.2)	1 (2.2)	2 (4.3)	0	0	0	0	0	1 (4.3)	0	0	1 (3.4)	0	0
Sclerosis																
Entire SI joint	3 (7.3)	1 (2.4)	4 (8.7)	7 (15.2)	0	1 (7.1)	1 (4.0)	1 (4.0)	3 (11.5)	1 (3.8)	1 (4.3)	0	1 (3.4)	2 (6.9)	3 (7.9)	1 (2.6)
Upper ilium	3 (7.3)	1 (2.4)	3 (7.3)	2 (4.3)	0	1 (7.1)	1 (4.0)	1 (4.0)	1 (3.8)	1 (3.8)	0	0	1 (3.4)	2 (6.9)	1 (2.6)	1 (2.6)
Lower ilium	1 (2.4)	0	4 (8.7)	7 (15.2)	1 (7.1)	0	0	1 (4.0)	4 (15.4)	0	1 (4.3)	0	0	0	4 (10.5)	0
Upper sacrum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lower sacrum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Erosion																
Entire SI joint	11 (26.8)	14 (34.1)	3 (6.5)	1 (2.2)	0	0	0	0	0	0	0	0	0	0	0	0
Upper ilium	9 (22.0)	12 (29.3)	2 (4.3)	0	0	0	0	0	0	0	0	0	0	0	0	0
Lower ilium	8 (19.5)	12 (29.3)	1 (2.2)	1 (2.2)	0	0	0	0	0	0	0	0	0	0	0	0
Upper sacrum	5 (12.2)	3 (7.3)	1 (2.2)	0	0	0	0	0	0	0	0	0	0	0	0	0
Lower sacrum	9 (22.0)	1 (2.4)	1 (2.2)	1 (2.2)	0	0	0	0	0	0	0	0	0	0	0	0
Backfill																
Entire SI joint	4 (9.8)	10 (24.4)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Upper half	4 (9.8)	8 (19.5)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lower half	2 (4.9)	9 (22.0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ankylosis																
Entire SI joint	4 (9.8)	11 (26.1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Upper half	3 (7.3)	9 (22.0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lower half	5 (12.2)	11 (26.1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0

* Values are the number (%) of participants with a UNI versus BI (≥1) lesion on MRI based on concordant reads, i.e., both readers agree on lesion presence unilaterally (on the same side) or bilaterally overall in the entire joint, in each quadrant (upper or lower sacrum or ilium), or in half of the joint (upper or lower). Bone marrow edema, fat lesion, sclerosis, and erosion scores correspond to the number of sacroiliac (SI) joint quadrants on MRI with a lesion present, providing a total score range of 0–72. Backfill and ankylosis scores correspond to the number of SI joint halves on MRI with a lesion present, as a lesion is recorded per upper and lower joint halves, providing a total score range of 0–36. SpA = spondyloarthritis.
† From disc herniation, cleaning staff, and runner groups.

versus unilateral BME 20% versus 22%, fat lesions 4% versus 2%, sclerosis 15% versus 9%, erosion 2% versus 7%), while the other groups almost only had unilateral lesions. In the individual quadrants, concordant reads showed that bilateral BME was overall rarer than unilateral BME. Fat lesions were frequently present in the axial SpA group, often bilaterally, and could be located in all 4 quadrants, although most often in the sacral quadrants. In the non–axial SpA groups, fat lesions were few, and they appeared in all quadrants and also both unilaterally and bilaterally. Sclerosis was seen in all groups unilaterally or bilaterally and exclusively in the iliac quadrants, and it was most frequently seen in postpartum women and patients with axial SpA. Erosion was frequently seen in patients with axial SpA where unilateral erosion was frequently present in all 4 quadrants, whereas bilateral erosion was primarily seen in the upper and lower ilium. Erosion was rare in women with postpartum pain, but when present, it was primarily observed unilaterally, whereas erosion was not seen in all the other control groups. Backfill and ankylosis were exclusively found in patients with axial SpA and mostly bilaterally.

Lesion location in anterior, central, and posterior slices. The proportion of participants with a lesion on MRI present in different MRI slices (slice 1 [most anterior] to slice 9 [most posterior]) in the individual groups, based on concordant reads, is shown in Figure 2. BME was seen in all slices in the axial SpA group. In women with postpartum pain, BME was mainly seen in the anterior and central slices (with >20% having BME in slices 2–4). In the other participant groups, BME was recorded mostly in slices 1–4, and only in a few exceptions was BME recorded from slices 5–9 (i.e., in the posterior part of the joint). No BME was recorded in healthy men. Fat lesions were frequently present in the axial SpA group and located in all slices, while in the other groups, fat lesions were located without any distinct slice pattern. Sclerosis was most frequent in women with postpartum pain and found throughout all slices, however, most frequently in the anterior and central slices (up to 17% in slice 3). In the axial SpA group, sclerosis was rare. Erosion was only recorded in the axial SpA group in all slices and in women with postpartum pain, primarily in the anterior slices.

Lesion location in quadrants at different joint sections. Table 4 shows the proportion of participants with specific lesions at different quadrants according to the thirds of the joint based on concordant reads. In patients with axial SpA and women with postpartum pain, BME was present in all quadrants and in all 3 sections of the joint. In women without postpartum pain, BME was present mostly in the anterior section of the lower iliac and the sacral quadrants. In the other groups, BME was rarely present, and if so, in the anterior and central sections. Fat lesions were frequently present in the axial SpA group in all sections and quadrants of the joint. In all other groups, fat lesions were an infrequent finding, and if any, then often more prevalent in sacral quadrants. Sclerosis was present in all groups; however, it was numerically more frequent in the women with postpartum pain and mainly in the iliac quadrants in all sections. Erosion was by far most frequent in patients with axial SpA in all quadrants and sections, but primarily in the ilium. Erosion was, albeit rarely, also seen in women with postpartum pain, and then mainly in the anterior section.

MRI findings in postpartum women with disc herniation, cleaning staff, and long-distance runners. Patient characteristics of the subgroup of women with ≥1 childbirth from the disc herniation, cleaning staff, and runner groups (mean time since last delivery 9.7 years [range 1.7–22.3]) are presented in Table 1. BME, fat lesions, and sclerosis were rare in this subgroup, whereas erosion, backfill, and ankylosis were absent (Tables 3 and 4). Mostly, lesions were unilateral and involved the anterior and/or central sections of the ilium.

DISCUSSION

In this prospective study of 204 participants consisting of patients with axial SpA and different control subjects with or without buttock or pelvic pain, we have in detail mapped out the anatomic location and distribution of inflammatory and structural MRI features in the SI joints (per joint quadrant and half and/or per slice and joint section).

Several MRI studies have over time sought to identify certain patterns of SI joint lesions to differentiate axial SpA from various other conditions (4,13,21). Although evaluating the SI joint by MRI is of great value, discriminating axial SpA from other conditions still poses a major challenge because BME, which plays a central role in the ASAS classification criteria for axial SpA, has also been described in several other non–axial SpA groups of subjects (i.e., patients with nonspecific back pain [4,5], pregnant and postpartum women [5–9], and healthy individuals [4,9], including athletes [6,9,10], military recruits [11], and the general population [12]). In previous publications, it has been shown that not only BME, but also fat lesions (4,5,9–11,14), erosion (4,5,7,9–11,13), and sclerosis (5,8,11,13,15) are present to some extent in non–axial SpA groups, particularly in postpartum women (6–9,15). While there has been much focus on scores for SI joint lesions on MRI (most often, SPARCC Sacroiliac Joint Structural scores) and score thresholds for differential diagnostic purposes (4,13,21), the differences in anatomic location and distribution of different lesions on MRI have only been investigated in a few studies (6,8,10) but have not been systematically addressed in studies assessing both patients with axial SpA and various control groups.

In a retrospective study of 93 peripartum (pregnant and ≤6 months postpartum) women with clinical indications for MRI examination due to a broad range of symptoms such as low back, pelvic, or abdominal pain, neurologic deficit, and postoperative evaluation, Eshed et al (8) found that subchondral sclerosis, and

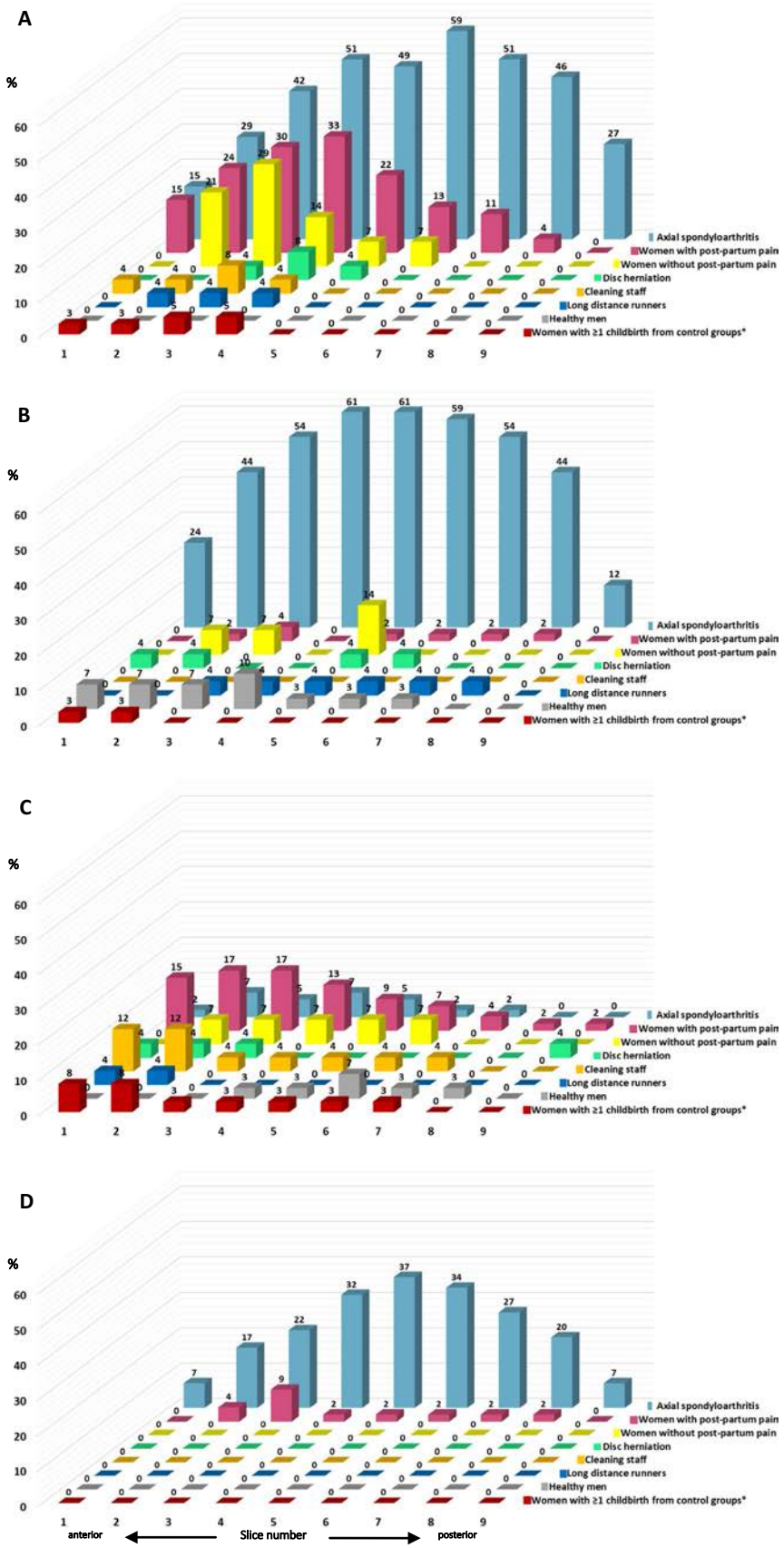


Figure 2. Proportion (%) of participants with ≥1 lesion of bone marrow edema (**A**), fat lesion (**B**), sclerosis (**C**), and erosion (**D**) per slice on magnetic resonance imaging (slices 1–9) stratified according to slice number. * = disc herniation, cleaning staff, and runners.

Table 4. Number and proportion of participants with ≥1 lesion on magnetic resonance imaging (MRI) per joint quadrant or joint half located in either the anterior, central, or posterior section of the joint*

	Axial SpA (n = 41)	Women with postpartum pain (n = 46)	Women without postpartum pain (n = 14)	Disc herniation (n = 25)	Cleaning staff (n = 26)	Long-distance runners (n = 23)	Healthy men (n = 29)	Women with ≥1 childbirth (n = 38)†
Bone marrow edema								
Upper ilium								
ANT	10 (24.4)	10 (21.7)	0	0	1 (3.8)	0	0	1 (2.6)
CENT	16 (39.0)	6 (13.0)	0	0	1 (3.8)	0	0	1 (2.6)
POST	7 (17.1)	1 (2.2)	0	0	0	0	0	0
Lower ilium								
ANT	13 (31.7)	12 (26.1)	3 (21.4)	0	1 (3.8)	1 (4.3)	0	1 (2.6)
CENT	18 (43.9)	11 (23.9)	2 (14.3)	0	1 (3.8)	0	0	1 (2.6)
POST	13 (31.7)	4 (8.7)	0	0	0	0	0	0
Upper sacrum								
ANT	10 (24.4)	5 (10.9)	3 (21.4)	1 (4.0)	1 (3.8)	0	0	1 (2.6)
CENT	20 (48.8)	13 (28.3)	1 (7.1)	1 (4.0)	0	0	0	1 (2.6)
POST	14 (34.1)	4 (8.7)	0	0	0	0	0	0
Lower sacrum								
ANT	7 (17.1)	4 (8.7)	2 (14.3)	1 (4.0)	0	0	0	0
CENT	15 (36.6)	9 (19.6)	1 (7.1)	1 (4.0)	0	0	0	0
POST	14 (34.1)	1 (2.2)	0	0	0	0	0	0
Fat lesion								
Upper ilium								
ANT	16 (39.0)	1 (2.2)	1 (7.1)	1 (4.0)	0	0	2 (6.9)	1 (2.6)
CENT	10 (24.4)	0	0	0	0	0	2 (6.9)	0
POST	5 (12.2)	0	0	0	0	0	0	0
Lower ilium								
ANT	17 (41.5)	0	1 (7.1)	1 (4.0)	0	0	1 (3.4)	1 (2.6)
CENT	12 (29.3)	0	0	0	0	0	1 (3.4)	0
POST	9 (22.0)	1 (2.2)	0	0	0	0	0	0
Upper sacrum								
ANT	15 (36.6)	2 (4.3)	1 (7.1)	0	0	1 (4.3)	0	0
CENT	28 (68.3)	1 (2.2)	1 (7.1)	1 (4.0)	0	1 (4.3)	2 (6.9)	0
POST	20 (48.8)	1 (2.2)	0	0	0	1 (4.3)	1 (3.4)	0
Lower sacrum								
ANT	12 (29.3)	1 (2.2)	0	0	0	1 (4.3)	1 (3.4)	0
CENT	25 (61.0)	0	2 (14.3)	0	0	1 (4.3)	1 (3.4)	0
POST	19 (46.3)	1 (2.2)	0	0	0	0	0	0
Sclerosis								
Upper ilium								
ANT	3 (7.3)	3 (6.5)	1 (7.1)	1 (4.0)	2 (7.7)	0	0	2 (5.3)
CENT	2 (4.9)	5 (10.9)	1 (7.1)	0	1 (3.8)	0	2 (6.9)	1 (2.6)
POST	1 (2.4)	1 (2.2)	0	1 (4.0)	1 (3.8)	0	2 (6.9)	1 (2.6)
Lower ilium								
ANT	1 (2.4)	11 (23.9)	1 (7.1)	1 (4.0)	3 (11.5)	1 (4.3)	0	3 (7.9)
CENT	1 (2.4)	6 (13.0)	1 (7.1)	0	1 (3.8)	0	0	1 (2.6)
POST	0	1 (2.2)	0	0	1 (3.8)	0	0	1 (2.6)
Upper sacrum								
ANT	0	0	0	0	0	0	0	0
CENT	0	0	0	0	0	0	0	0
POST	0	0	0	0	0	0	0	0
Lower sacrum								
ANT	0	0	0	0	0	0	0	0
CENT	0	0	0	0	0	0	0	0
POST	0	0	0	0	0	0	0	0
Erosion								
Upper ilium								
ANT	12 (29.3)	2 (4.3)	0	0	0	0	0	0
CENT	14 (34.1)	0	0	0	0	0	0	0
POST	5 (12.2)	1 (2.2)	0	0	0	0	0	0

(Continued)

Table 4. *(Cont'd)*

	Axial SpA (n = 41)	Women with postpartum pain (n = 46)	Women without postpartum pain (n = 14)	Disc herniation (n = 25)	Cleaning staff (n = 26)	Long-distance runners (n = 23)	Healthy men (n = 29)	Women with ≥1 childbirth (n = 38)†
Lower ilium								
ANT	12 (29.3)	2 (4.3)	0	0	0	0	0	0
CENT	13 (31.7)	1 (2.2)	0	0	0	0	0	0
POST	6 (14.6)	1 (2.2)	0	0	0	0	0	0
Upper sacrum								
ANT	1 (2.4)	2 (4.3)	0	0	0	0	0	0
CENT	4 (9.8)	1 (2.2)	0	0	0	0	0	0
POST	4 (9.8)	0	0	0	0	0	0	0
Lower sacrum								
ANT	2 (4.9)	1 (2.2)	0	0	0	0	0	0
CENT	6 (14.6)	1 (2.2)	0	0	0	0	0	0
POST	2 (4.9)	0	0	0	0	0	0	0
Backfill								
Upper half								
ANT	4 (9.8)	0	0	0	0	0	0	0
CENT	12 (29.3)	0	0	0	0	0	0	0
POST	7 (17.1)	0	0	0	0	0	0	0
Lower half								
ANT	6 (14.6)	0	0	0	0	0	0	0
CENT	10 (24.4)	0	0	0	0	0	0	0
POST	6 (14.6)	0	0	0	0	0	0	0
Ankylosis								
Upper half								
ANT	8 (19.5)	0	0	0	0	0	0	0
CENT	10 (24.4)	0	0	0	0	0	0	0
POST	9 (22.0)	0	0	0	0	0	0	0
Lower half								
ANT	10 (24.4)	0	0	0	0	0	0	0
CENT	13 (31.7)	0	0	0	0	0	0	0
POST	11 (26.8)	0	0	0	0	0	0	0

* Values are the number (%) of participants with ≥1 lesion per joint quadrant or joint half located in either the anterior, central, or posterior section of the joint in 9 slices. Results are based on concordant reads, i.e., the presence of a lesion in the same joint section (anterior, central, or posterior) in the same joint quadrant or joint half. Bone marrow edema, fat lesion, sclerosis, and erosion scores correspond to the number of sacroiliac joint quadrants on MRI with a lesion present, providing a total score range of 0–72. Backfill and ankylosis scores correspond to the number of sacroiliac joint halves on MRI with a lesion present, as a lesion is recorded per upper and lower joint halves, providing a total score range of 0–36. ANT = anterior slices, i.e., slices 1–3; CENT = central slices, i.e., slices 4–6; POST = posterior slices, i.e., slices 7–9; SpA = spondyloarthritis.
† From disc herniation, cleaning staff, and runner groups.

especially BME, were frequent MRI findings (26% and 46% of participants, respectively), and both lesion types appeared mainly in the lower iliac and sacral quadrants. In the 46 women with postpartum pain in our study, we found comparable frequencies of sclerosis and BME (24% and 41%, respectively). In contrast to the findings of Eshed et al (8), we found BME equally distributed in the upper and lower parts of the SI joint (50% and 50%, respectively) but predominantly in the lower ilium (30%) and upper sacrum (28%). However, we also found sclerosis mainly in the lower part of the SI joint (24% versus 11% in the upper part of the SI joint) but exclusively in the ilium. Differences in results may be due to different study populations and different definitions of lesion. The distribution of SI joint BME on MRI has also been examined by de Winter et al (6), who found BME in healthy participants (n = 47) predominantly in the lower iliac quadrants, mainly posteriorly in the joints, while in patients with axial SpA (n = 47) and in women with postpartum back pain (n = 7), it was distributed equally anteriorly and posteriorly, and in frequent runners (n = 24) in the upper ilium, mainly in the anterior part.

For patients with axial SpA and postpartum women with pain, our findings are relatively consistent with those shown in a study by de Winter et al (6), although a comprehensive comparison is not possible because different subdivisions of the SI joint were applied in the 2 studies. However, in contrast to the study by de Winter et al (6), we hardly found any BME in our runners or healthy participants. In a study of the semicoronal SI joint on MRI of 42 athletes (recreational runners and ice hockey players, but no patients with axial SpA), Weber et al (10) found that 30–41% of 42 athletes fulfilled the ASAS definition of sacroiliitis, and that BME was most frequently located in the posterior lower ilium, followed by the anterior upper sacrum. Furthermore, they found that fat metaplasia was rarer than BME and did not show any particular pattern of distribution in the SI joint, while erosion was practically absent. Moreover, when Weber et al (22) added a semiaxial scan

to the evaluation of the semicoronal scan, they reported that BME on 2 perpendicular planes was only present in 25–27% of the athletes, and that the proportion of athletes fulfilling the ASAS definition of sacroiliitis was reduced by 33–56% compared to the first study. These results were ascribed to reduced misinterpretation of the hyperintense signal on the STIR sequence, when the signal was located at 4 constitutional SI joint features (partial volume of vascular signal, deep ligament insertion containing vascular signals, fluid-filled bone cysts, and lumbosacral transitional anomaly), which were better ruled out using 2 scan planes. In line with the study by Weber et al (10), we did not find erosion, nor did we find backfill, in the long-distance runners in our study. However, we found markedly fewer fat lesions (4%) and BME (9%) than Weber et al (10), and we did not find BME in the most posterior part of the SI joint in the non–axial SpA groups, including the runners.

This divergence in the anatomic location of BME may be attributable to the several differences between the studies. First, to avoid misinterpretation at the 4 constitutional features, we recorded these features separately (data not shown) and not as BME. Second, in our study, the difference in the subdivision of the SI joint in the semicoronal scan plane, i.e., in 3 SI joint sections (anterior, central, and posterior), versus the 2 sections (anterior and posterior) in the study by Weber et al (10) may have reduced the number of patients with registered lesions in anterior and posterior sections in our study because these were registered as located in the central section. Other possible contributors to the differences between studies could be differences in study population, i.e., the intensity of training and/or the type and amount of pelvic strain in the groups of athletes. Finally, the MRI acquisition parameters differed between studies because we applied a TE for STIR images of 37 msec, versus 68 msec in the study by Weber et al (10), which may have slightly reduced the detectability of BME lesions (23) in our study.

Only 1 previous study has investigated the symmetry of lesions on MRI in the SI joints (24). A retrospective study of 68 patients with axial SpA (nonradiographic axial SpA [n = 48] versus ankylosing spondylitis [AS] [n = 20]) characterized the distribution and symmetry of lesions on MRI in nonradiographic axial SpA versus AS (24). Although both unilateral and bilateral lesions were observed, BME, fat metaplasia, sclerosis, erosion, and ankylosis were all more frequently bilateral than unilateral, both in nonradiographic axial SpA and AS. These results are in agreement with our results except for sclerosis.

Our study suggests that the presence of BME per se and its location cannot be used to differentiate between patients with axial SpA and women with postpartum pain. However, the findings of more widespread distribution of fat lesions (unilaterally and bilaterally) and erosion in patients with axial SpA may improve differentiation because women with postpartum pain and all other control groups had limited amounts of fat lesions, and erosion was practically absent.

The strengths of the current study include the prospective design, the high number of participants, who were divided into different categories based on predefined, strict inclusion criteria, a standardized MRI protocol covering the entire cartilaginous compartment of the SI joints, and finally, the blinded reads by 2 independent, experienced readers. The limitations include the lack of follow-up for non–axial SpA patients to examine the natural course of the changes. However, in a post hoc analysis of women who had previously given birth in the groups of disc herniation, cleaning staff, and long-distance runners, we found very limited BME, fat lesions, and sclerosis and no erosion, backfill, and ankylosis, which suggests that SI joint findings on MRI in postpartum women may diminish over time. This is in agreement with the observation of Arnbak et al (25), who observed that small and intermediate SI joint BME lesions on MRI in patients with low back pain were mostly transient and rarely developed into extensive BME or structural lesions such as fat lesions and erosion.

In conclusion, we have described typical locations of common SI joint lesions in axial SpA and non–axial SpA. While BME is present in patients with axial SpA and in the vast majority of control groups, especially postpartum women, erosion and fat lesions exhibit more promising results when discriminating patients with axial SpA from control groups, particularly with more widespread distribution in the joints and/or in central and/or posterior slices. Backfill and ankylosis may be considered pathognomonic for axial SpA but are a late finding and not relevant for improving early diagnosis.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of study participants, radiographers, study nurses, and laboratory staff at Rigshospitalet Glostrup who participated in the 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. Seven 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. Østergaard, Morsel-Carlsen, Sørensen, Pedersen.

Acquisition of data. Morsel-Carlsen, Sørensen, Bonde, Thamsborg, Lykkegaard, Pedersen.

Analysis and interpretation of data. Seven, Østergaard, Morsel-Carlsen, Sørensen, Pedersen.

REFERENCES

1. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777–83.
2. 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.

3. Rudwaleit M, Jurik AG, Hermann KG, Landewe R, van der Heijde D, Baraliakos X, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009;68:1520–7.
4. Weber U, Lambert RG, Østergaard M, Hodler J, Pedersen SJ, Maksymowych WP. The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. Arthritis Rheum 2010;62:3048–58.
5. Arnbak B, Jensen TS, Egund N, Zejden A, Horslev-Petersen K, Manniche C, et al. Prevalence of degenerative and spondyloarthritis-related magnetic resonance imaging findings in the spine and sacroiliac joints in patients with persistent low back pain. Eur Radiol 2016;26:1191–203.
6. De Winter J, de Hooge M, van de Sande M, de Jong H, van Hoeven L, de Koning A, et al. Magnetic resonance imaging of the sacroiliac joints indicating sacroiliitis according to the Assessment of SpondyloArthritis international Society definition in healthy individuals, runners, and women with postpartum back pain. Arthritis Rheumatol 2018;70:1042–8.
7. Agten CA, Zubler V, Zanetti M, Binkert CA, Kolokythas O, Prentl E, et al. Postpartum bone marrow edema at the sacroiliac joints may mimic sacroiliitis of axial spondyloarthritis on MRI. AJR Am J Roentgenol 2018;211:1306–12.
8. Eshed I, Miloh-Raz H, Dulitzki M, Lidar Z, Aharoni D, Liberman B, et al. Peripartum changes of the sacroiliac joints on MRI: increasing mechanical load correlating with signs of edema and inflammation kindling spondyloarthropathy in the genetically prone. Clin Rheumatol 2015;34:1419–26.
9. Seven S, Østergaard M, Morsel-Carlsen L, Sorensen IJ, Bonde B, Thamsborg G, et al. Magnetic resonance imaging of lesions in the sacroiliac joints for differentiation of patients with axial spondyloarthritis from control subjects with or without pelvic or buttock pain: a prospective, cross-sectional study of 204 participants. Arthritis Rheumatol 2019;71:2034–46.
10. Weber U, Jurik AG, Zejden A, Larsen E, Jørgensen SH, Rufibach K, et al. Frequency and anatomic distribution of magnetic resonance imaging features in the sacroiliac joints of young athletes: exploring "background noise" toward a data-driven definition of sacroiliitis in early spondyloarthritis. Arthritis Rheumatol 2018;70:736–45.
11. Varkas G, de Hooge M, Renson T, De Mits S, Carron P, Jacques P, et al. Effect of mechanical stress on magnetic resonance imaging of the sacroiliac joints: assessment of military recruits by magnetic resonance imaging study. Rheumatology (Oxford) 2018;57:508–13.
12. Baraliakos X, Richter A, Feldmann D, Ott A, Buelow R, Schmidt CO, et al. Frequency of MRI changes suggestive of axial spondyloarthritis in the axial skeleton in a large population-based cohort of individuals aged <45 years. Ann Rheum Dis 2020;79:186–92.
13. De Hooge M, van den Berg R, Navarro-Compan V, Reijnierse M, van Gaalen F, Fagerli K, et al. Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis? Ann Rheum Dis 2016;75:1308–14.
14. Ziegeler K, Eshkal H, Schorr C, Sieper J, Diekhoff T, Makowski MR, et al. Age- and sex-dependent frequency of fat metaplasia and other structural changes of the sacroiliac joints in patients without axial spondyloarthritis: a retrospective, cross-sectional MRI study. J Rheumatol 2018;45:915–21.
15. Seven S, Østergaard M, Morsel-Carlsen L, Sorensen IJ, Bonde B, Thamsborg G, et al. The utility of magnetic resonance imaging lesion combinations in the sacroiliac joints for diagnosing patients with axial spondyloarthritis: a prospective study of 204 participants including postpartum women, patients with disc herniation, cleaning staff, runners and healthy persons. Rheumatology (Oxford) 2020;59:3237–49.
16. Maksymowych WP, Wichuk S, Chiowchanwisawakit P, Lambert RG, Pedersen SJ. Development and preliminary validation of the Spondyloarthritis Research Consortium of Canada Magnetic Resonance Imaging Sacroiliac Joint Structural Score. J Rheumatol 2015;42:79–86.
17. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Williams M, Stone M, et al. Spondyloarthritis Research Consortium of Canada Magnetic Resonance Imaging Index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005; 53:703–9.
18. 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.
19. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8:23–34.
20. Carley S, Dosman S, Jones SR, Harrison M. Simple nomograms to calculate sample size in diagnostic studies. Emerg Med J 2005; 22:180–1.
21. Hu Z, Wang X, Qi J, Kong Q, Zhao M, Gu J. Backfill is a specific sign of axial spondyloarthritis seen on MRI. Joint Bone Spine 2016; 83:179–83.
22. Weber U, Jurik AG, Zejden A, Larsen E, Jorgensen SH, Rufibach K, et al. MRI of the sacroiliac joints in athletes: recognition of non-specific bone marrow oedema by semi-axial added to standard semi-coronal scans. Rheumatology (Oxford) 2020;59:1381–90.
23. Greese J, Diekhoff T, Sieper J, Schwenke C, Makowski MR, Poddubnyy D, et al. Detection of sacroiliitis by short-tau inversion recovery and T2-weighted turbo spin echo sequences: results from the SIMACT study. J Rheumatol 2019;46:376–83.
24. Motamedi D, Patel R, Devulapalli KK, Gensler LS, Steinbach L. MR distribution of active inflammatory and chronic structural sacroiliac joint changes in axial spondyloarthropathy: challenging conventional wisdom. Clin Imaging 2019;58:70–3.
25. Arnbak B, Jensen TS, Schiottz-Christensen B, Pedersen SJ, Østergaard M, Weber U, et al. What level of inflammation leads to structural damage in the sacroiliac joints? A four-year magnetic resonance imaging follow-up study of low back pain patients. Arthritis Rheumatol 2019;71:2027–33.

Arthritis Care & Research
Vol. 73, No. 5, May 2021, pp 755–764
DOI 10.1002/acr.24402

American College of Rheumatology
Empowering Rheumatology Professionals

Efficacy and Safety of Pharmacologic Interventions in Patients Experiencing a Gout Flare: A Systematic Review and Network Meta-Analysis

Linan Zeng,[1] Anila Qasim,[2] Tuhina Neogi,[3] John D. Fitzgerald,[4] Nicola Dalbeth,[5] Ted R. Mikuls,[6] Gordon H. Guyatt,[2] and Romina Brignardello-Petersen[2]

Objective. To compare the relative efficacy and safety of pharmacologic antiinflammatory interventions for gout flares.

Methods. We searched Ovid Medline, Embase, and Cochrane library for randomized controlled trials (RCTs) that compared pharmacologic antiinflammatory treatment of gout flares. We conducted a network meta-analysis (NMA) using a frequentist framework and assessed the certainty of evidence and made conclusions using the Grading of Recommendations Assessment, Development, and Evaluation for NMA.

Results. In the 30 eligible RCTs, canakinumab provided the highest pain reduction at day 2 and at longest follow-up (mean difference relative to acetic acid derivative nonsteroidal antiinflammatory drugs [NSAIDs] –41.12 [95% confidence interval (95% CI) –53.36, –29.11] on a 0–100 scale at day 2, and mean difference –12.84 [95% CI –20.76, –4.91] at longest follow-up; both moderate certainty; minimum important difference –19). Intravenous or intramuscular corticosteroids were inferior to canakinumab but may be better than the other commonly used interventions (low to very low certainty). For joint tenderness, canakinumab may be the most effective intervention at day 2. Acetic acid derivative NSAIDs improved joint swelling better than ibuprofen NSAIDs at day 2 (mean difference –0.29 [95% CI –0.56, –0.02] on a 0–4 scale; moderate certainty) and improved patient global assessment (PtGA) greater than ibuprofen NSAIDs at the longest follow-up (mean difference –0.44 [95% CI –0.86, –0.02]; moderate).

Conclusion. Canakinumab may be superior to other alternatives and intravenous or intramuscular corticosteroids may be the second best in pain reduction. Acetic acid derivative NSAIDs may be superior to ibuprofen NSAIDs in improving joint swelling and PtGA.

INTRODUCTION

Gout is the most common inflammatory arthritis worldwide, caused by deposition of monosodium urate crystals in joint structures and other sites (1). Despite advances in understanding of the pathophysiology and therapy, gout continues to impair individual's health-related quality of life and consume health care resources (2). For management of gout flares, pharmacologic therapies focus on rapid and effective control of the inflammatory response to monosodium urate crystals, thereby reducing joint pain and inflammation (3). Despite the consistent recommendations of first-line options for gout flare from the American College of Rheumatology (ACR), the American College of Physicians (ACP), the British Society for Rheumatology, and the European Alliance of Associations for Rheumatology, uncertainty of the efficacy and safety of many pharmacologic interventions remains (1,4–6). Moreover, due to lack of evidence on comparative efficacy and safety, guidelines do not prioritize between these pharmacologic options (4).

The comparative efficacy between current first-line options, e.g., nonsteroidal antiinflammatory drugs (NSAIDs), corticosteroids, or colchicine, and other pharmacologic interventions, e.g., interleukin-1 (IL-1) inhibitors, remains unclear. Network meta-analysis (NMA) could help improve the precision by combining direct and indirect evidence, an approach that to date has not been performed to assess the comparative efficacy and safety of pharmacologic antiinflammatory interventions for gout flares. We

[1]Linan Zeng, PhD: West China Second University Hospital, Sichuan University, Chengdu, China, and McMaster University, Hamilton, Ontario, Canada; [2]Anila Qasim, HBSc, MSc, Gordon H. Guyatt, MD, Romina Brignardello-Petersen, MD, PhD: McMaster University, Hamilton, Ontario, Canada; [3]Tuhina Neogi, MD, PhD, FRCPC: Boston University, Boston, Massachusetts; [4]John D. Fitzgerald, MD, PhD: University of California, Los Angeles and VA Greater Los Angeles Health Care System, Los Angeles, California; [5]Nicola Dalbeth, MD, FRACP: University of Auckland, Auckland, New Zealand; [6]Ted R. Mikuls, MD, MSPH: University of Nebraska Medical Center and VA Nebraska-Western Iowa Health Care System, Omaha.

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

Address correspondence to Romina Brignardello-Petersen, MD, PhD, 1280 Main Street West, Hamilton, ON, L8S 4L8, Canada. Email: brignarr@mcmaster.ca.

Submitted for publication March 13, 2020; accepted in revised form July 23, 2020.

SIGNIFICANCE & INNOVATIONS

- Despite consistent recommendations of first-line options for gout flare from guidelines, uncertainty of the efficacy and safety of pharmacologic interventions remains.
- This systematic review identifies, in patients with gout flares, a potential advantage of canakinumab versus other antiinflammatory interventions in pain reduction at day 2 and longest follow-up, and in improvement of joint tenderness at day 2.
- Among commonly used interventions, intravenous or intramuscular corticosteroids may be superior to cyclooxygenase 2 (COX-2) highly selective nonsteroidal antiinflammatory drugs (NSAIDs), ibuprofen NSAIDs, colchicine, and oral corticosteroids in pain reduction at day 2. Acetic acid derivative NSAIDs are probably superior to ibuprofen NSAIDs in reducing joint swelling at day 2 and patient global assessment at longest follow-up
- This systematic review highlights the need for further evaluation of the comparative efficacy and safety of interventions used commonly in practice but not yet tested in randomized controlled trials (e.g., colchicine, pyrazolidine derivative NSAIDs, COX-2 selective NSAIDs, and fenamate NSAIDs) and of multiple-drug treatments (e.g., interleukin-1 inhibitor plus acetic acid derivative NSAIDs) for gout flares.

therefore conducted this NMA considering both direct and indirect comparison to address the relative efficacy and safety of pharmacologic antiinflammatory interventions for gout flares based on evidence from randomized controlled trials (RCTs).

MATERIALS AND METHODS

Our systematic review was proposed by the ACR as one of the systematic reviews supporting its 2020 guideline of management of patients with gout (7). We did not register a protocol but followed the methodology established by the ACR to conduct systematic reviews to inform their guidelines. This report adheres to the Preferred Reporting items for Systematic Reviews and Meta-Analyses statement (8).

Data source and searches. A research librarian conducted a single literature search for evidence pertaining to 57 questions in support of the ACR 2020 guidelines simultaneously in Ovid Medline, Embase, and Cochrane library on September 24, 2018. We updated the search for this specific question through December of 2019. The search strategies for each database are outlined (see Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24402/abstract).

Study selection. We made decisions about eligibility criteria for patients, interventions, outcomes, and types of studies based on the needs of the ACR guidelines. We included RCTs that enrolled adult patients with gout flares and compared ≥2 antiinflammatory pharmacologic interventions or compared pharmacologic interventions with placebo. Eligible trials reported at least 1 outcome, including pain, joint tenderness, joint swelling, patient global assessment (PtGA), or serious adverse events (SAE) with any duration of follow-up. Based on input of the guideline panel, we grouped interventions according to pharmacologic mechanism

Table 1. Pharmacologic interventions included in each intervention node*

Category of pharmacologic mechanism, intervention node	Pharmacologic interventions included in each node
Corticosteroids	
Oral	Prednisolone
IM or IV	Compound betamethasone, methylprednisolone, triamcinolone acetonide
Colchicine	Colchicine
ACTH	ACTH
NSAIDs	
Acetic acid derivative	Etodolac, indomethacin, diclofenac
Ibuprofen	Ketoprofen, naproxen, flurbiprofen
Pyrazolidine derivative	Phenylbutazone, azapropazone
Fenamate	Meclofenamate sodium, flufenamic acid
Selective NSAIDs	
COX-2 selective	Meloxicam
COX-2 highly selective	Etoricoxib, celecoxib, rofecoxib, lumiracoxib
IL inhibitors	
Rilonacept	Rilonacept
Canakinumab	Canakinumab
Anakinra	Anakinra
Acetaminophen	Acetaminophen
Combinations	Rilonacept plus indomethacin
IL-1 inhibitors plus acetic acid derivative NSAIDs	

* ACTH = adrenocorticotropic hormone; COX-2 = cyclooxygenase 2; IL-1 = interleukin-1; IM = intramuscular; IV = intravenous; NSAIDs = nonsteroidal antiinflammatory drugs.

of action and route of administration (Table 1). We excluded trials that compared interventions from the same intervention node (e.g., both arms in the trial used ibuprofen NSAIDs) and those not published in the English or as conference abstracts only.

Reviewers, working in pairs, screened titles and abstracts to determine potential eligibility for all guideline questions, and entries identified by at least 1 reviewer proceeded to full-text eligibility review, which was also conducted in duplicate. A pair of reviewers (LZ and AQ) confirmed eligibility of the studies addressing this systematic review question. A third adjudicator (RB-P) helped to resolve any disagreement, through consensus.

Data abstraction. One reviewer (LZ) used standardized forms to extract data of study design, characteristics of participants, regimens of pharmacologic interventions, and relevant outcomes. Another reviewer (AQ) checked the data. A third adjudicator (RB-P) reviewed disagreements, and the 3 reviewers reached consensus through discussion.

The guideline panel prioritized methods for measurement for the outcomes that were endorsed by the Outcome Measures in Rheumatology (9), and time points of interest (day 2 or the day closest to day 2, and longest available follow-up). We abstracted data from the following outcomes:

1. Mean change in pain score. The prioritized instrument was the 100-mm visual analog scale (VAS) (0 mm = no pain, 100 mm = unbearable pain) in which the minimum important difference (MID) for gout patients is a 19-point reduction (10).
2. Mean change in joint tenderness and mean change in joint swelling. The prioritized instrument was the 4-point Likert scale (for joint tenderness: 0 = no pain, 3 = pain, winces, and withdraws; for joint swelling: 0 = no swelling, 3 = bulging beyond the joint margins), where the MID is a 1-point reduction for joint tenderness and a 1-point reduction for joint swelling (10).
3. Mean change in PtGA. The prioritized instrument was the 5-point Likert scale (0 = excellent, 4 = poor). A MID for this 5-point Likert scale has not been established for gout patients.
4. SAE. We counted any adverse event that was classified as serious by the authors. When the authors did not report any SAE, we assumed none had occurred.

When the primary trials did not report the SD, we imputed the SD by using the median of SDs from other included trials that applied the same instrument in similar populations during similar follow-up periods.

Risk of bias and certainty of evidence. One reviewer (LZ) assessed the risk of bias of individual studies using the Cochrane risk of bias tool, and another reviewer (AQ) cross-checked the judgments. A third adjudicator (RB-P) reviewed disagreements not resolved by discussion.

Using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) for NMA, we chose a null effect as a threshold and assessed the certainty that a particular intervention has an effect (i.e., improves a particular outcome) compared with another. The certainty of the evidence can be high, moderate, low, or very low. The assessment of this body of evidence from randomized trials started as high and was rated down based on limitation of risk of bias, inconsistency, indirectness, publication bias, intransitivity, incoherence, and imprecision. The steps of the GRADE assessment for each comparison and outcome included:

1. Rating the certainty of both direct and indirect evidence contributing to the network estimate. For rating certainty in indirect evidence, we focused on the dominant first order loop. The certainty of the indirect evidence depends on the lowest certainty rating of the direct comparisons in the loop and intransitivity (i.e., the extent of similarity of direct comparisons forming the indirect comparison) (11).
2. Rating the certainty of the network estimate. When the network estimate was based on only direct or indirect evidence, the network certainty rating was based on the certainty of that estimate (11). When both direct and indirect estimates were available, the rating of the network estimate was based on the dominant evidence. To determine the final rating, we considered incoherence (i.e., the extent of similarity of direct and indirect estimates) and imprecision (11).

Data synthesis and analysis. To calculate direct estimates of effect for each paired comparison, we performed a frequentist random-effects pairwise meta-analysis using Review Manager 5.3 (Nordic Cochrane Centre; http://ims.cochrane.org/revman/download). For continuous outcomes, we used the standardized mean differences (SMDs) and corresponding 95% confidence intervals (95% CIs). For dichotomous outcomes, considering many trials had 0 events in 1 or 2 arms, we used risk differences (RDs) and corresponding 95% CIs as the measure of effect. We quantified statistical heterogeneity by estimating the variance between trials using chi-square test and I^2 statistic.

We conducted the NMA using a frequentist framework and a random-effects model by the package netmeta in R (version 1.1.463) (12). For continuous outcomes, we first calculated SMDs and corresponding 95% CIs, and then converted the SMDs into MDs in the natural units of prioritized standard scales by multiplying the SMDs by an estimate of the SD associated with the standard scales. We used RDs and 95% CIs for dichotomous outcomes as the measure of pooled effect.

Data interpretation. To make conclusions from the NMA, we used a novel methodology developed by the GRADE working group in which interventions are classified in groups from the most to the least efficacious or safe for each outcome (13). The approach begins by choosing an intervention that has

the most direct comparisons with other interventions as the reference intervention. The next step in the approach is to choose a decision threshold to categorize the interventions as not convincingly different, better, or worse than the reference. We chose a null effect as the decision threshold. Using the same decision threshold, we differentiated among interventions from categories that were better or worse than the reference. We then identified interventions within each category as those with high or moderate certainty relative to the reference standard, and those with low or very low certainty (13).

To facilitate the interpretation of the comparative efficacy and safety of each interventions in relation to the reference, we assumed an effect of the reference and calculated the difference between each intervention when compared to this reference. For continuous outcome, we estimated that the effect of the reference was the weighted average of the mean change from baseline in the reference arm across all studies. For dichotomous outcomes, we used an inverse-variance fixed-effects model and meta-analysis of proportions based on a generalized linear mixed model. We assessed the certainty of evidence by using GRADE for observational studies (treating the single arm from RCT as before–after study).

RESULTS

The initial search for all 57 questions in support of the guidelines yielded 3,337 citations; after reviewing abstracts for the systematic reviews, 466 proved potentially eligible. Twenty-nine RCTs (30 articles) proved eligible for this particular systematic review, which, following full text review, was focused on gout flare management. The updated search, which included dates until December 2019, found 1 new trial. We finally included 30 RCTs (31 articles) with 4,268 patients. We did not provide the specific reasons for exclusion of studies for this systematic review because we simultaneously screened studies for all of the systematic reviews for the broader needs of the full guidelines.

Characteristic of the included studies. The eligible trials studied several antiinflammatory interventions and their combinations for gout flare management, including oral corticosteroids, intravenous or intramuscular corticosteroids, acetic acid derivative NSAIDs, ibuprofen NSAIDs, fenamate NSAIDs, pyrazolidine derivative NSAIDs, cyclooxygenase 2 (COX-2) selective NSAIDs, COX-2 highly selective NSAIDs, adrenocorticotropic hormone, rilonacept, canakinumab, anakinra, colchicine, IL-1 inhibitor plus acetic acid derivative, and a free choice of colchicine, naproxen, or prednisolone (Table 2) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24402/abstract). Risk of bias of individual RCTs was mainly due to inadequate or unclear reporting of random sequence generation (46.7% [14 of 30]) or of allocation concealment (63.3% [19 of 30]), incomplete outcome including high proportion of lost to follow-up or unbalanced proportion of lost to follow-up between groups (43.3% [13 of 30]), and selective reporting including incomplete reporting of important outcomes or of mean or SDs (46.7% [14 of 30]) (see Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24402/abstract).

Table 2. Characteristics of included RCTs (n = 30)*

Characteristic	Value
No. of patients randomized, median (range)	91.5 (20–416)
No. of multi-arm trials, %	4 (13.3)
Weeks of treatment duration, median (range)	1.0 (0.1–52.1)
Intervention evaluated	
Acetic acid derivative NSAIDs	1,112/17
COX-2 highly selective NSAIDs	753/11
Corticosteroids, IM or IV	394/7
Corticosteroids, oral	312/3
Canakinumab	270/3
Ibuprofen NSAIDs	367/6
Colchicine	199/1
Rilonacept	75/1
IL-1 inhibitor plus acetic acid derivative NSAIDs	75/1
ACTH	53/2
Acetic acid derivative NSAIDs plus acetaminophen	45/1
Oral corticosteroids plus acetaminophen	45/1
Colchicine, or naproxen, or prednisone	44/1
Pyrazolidine derivative NSAIDs	44/3
COX-2 selective NSAIDs	31/1
Fenamate NSAIDs	13/1
Outcome analyzed, no. of patients analyzed/no. of trials	
Serious adverse events	4,266/30
Pain	3,961/23
Joint tenderness	2,928/17
Joint swelling	2,173/16
Patient global assessment	2,154/15
Methodologic characteristics, no. of trials (%)	
Adequate generation of random sequence	16 (53.3)
Adequate allocation concealment	11 (36.7)
Adequate blinding of outcome assessors	23 (76.7)
Characteristics of patients	
Percentage of men, median (range)	92.1 (68.4–100)
Age, median (range) years	53 (43.8–69.6)
Report of gout duration, no. of trials (%)	10 (33.3)

* Values are the number of patients randomized/number of trials, unless indicated otherwise. ACTH = adrenocorticotropic hormone; COX-2 = cyclooxygenase 2; IL-1 = interleukin-1; IM = intramuscular; IV = intravenous; NSAIDs = nonsteroidal antiinflammatory drugs; RCT = randomized controlled trial.

Effects of the interventions. We chose acetic acid derivative NSAIDs as the reference intervention for all outcomes, as it has the most direct comparisons with other interventions. Because 1 RCT that compared anakinra with a free choice of colchicine, naproxen, or prednisolone did not have interventions connected to the network by any node, we did not include this RCT in the NMA (14). In the results from NMA for the effectiveness outcomes (i.e., pain, joint tenderness, joint swelling, PtGA), a negative number indicates a better result with the intervention (i.e., greater pain reduction, better joint tenderness or joint swelling resolution, better PtGA improvement), whereas a positive number indicates a better result with the comparison. Network plots illustrating

the interventions and whether they have been compared directly in RCTs for each outcome are presented (see Supplementary Figure 2A–I, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24402/abstract).

Pain. Nineteen RCTs (3,560 patients, 9 interventions) described the change in pain from baseline at day 2 (15–32). The reference (i.e., acetic acid derivative NSAIDs) showed an important average reduction in pain from baseline to day 2 (MD –30.67 [95% CI –31.89, –29.45] on a 0 to 100 VAS; very low certainty; MID –19) (Table 3). Of the 36 pairwise comparisons between interventions, direct evidence was available for 12. Canakinumab proved probably the most effective intervention for reducing pain at day 2 (MD relative to acetic acid derivative NSAIDs –41.12 [95% CI –53.36, –29.11]; moderate certainty). Intravenous or intramuscular corticosteroids may be superior to other interventions but inferior to canakinumab (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24402/abstract). Rilonacept was probably better than acetic acid derivative NSAIDs but inferior to intravenous or intramuscular corticosteroids and canakinumab (see Supplementary Table 2). There were no convincing differences between COX-2 highly selective NSAIDs, ibuprofen NSAIDs, acetic acid derivative NSAIDs, colchicine, oral corticosteroids, or IL-1 inhibition plus acetic acid derivative NSAIDs (see Supplementary Table 2).

The NMA for change in pain at the longest follow-up (median 7 days, range 3–28 days) included 16 RCTs (2,384 patients, 9 interventions) (16–19,21–26,28–32). Of the 36 pairwise comparisons between interventions, direct evidence was available for 11. Acetic acid derivative NSAIDs showed an important average reduction in pain from baseline to the longest follow-up (MD –40.09 [95% CI –42.25, –39.61], very low certainty). Canakinumab was probably the most effective intervention at the longest follow-up (MD relative to acetic acid derivative NSAIDs –12.84 [95% CI –20.76, –4.91], moderate certainty). There were no convincing differences between acetic acid derivative NSAIDs, COX-2 highly selective NSAIDs, ibuprofen NSAIDs, colchicine, intravenous or intramuscular corticosteroids, oral corticosteroids, or rilonacept or IL-1 inhibition plus acetic acid derivative NSAIDs (see Supplementary Table 2).

Joint tenderness. Eight RCTs (1,308 patients; 6 interventions) reported on the change of joint tenderness from baseline on day 2 (16,18,20,25,31–33). The reference (i.e., acetic acid derivative NSAIDs) showed an important average improvement relative to baseline on joint tenderness at day 2 (MD –1.29 [95% CI –1.38, –1.21] on a 0 to 3 scale, very low certainty; MID –1) (Table 3). Of the 15 pairwise comparisons between interventions, direct evidence proved available for 6. Canakinumab was probably the most effective intervention at day 2 (MD relative to acetic acid derivative NSAIDs –0.67 [95% CI –1.03, –0.30], moderate certainty). However, the difference between canakinumab and acetic acid derivative NSAIDs was unimportant to gout patients (smaller than the MID of 1 point reduction). There were no convincing differences between COX-2 highly selective NSAIDs, ibuprofen NSAIDs, intravenous or intramuscular corticosteroids, oral corticosteroids, and the reference standard, acetic acid derivative NSAIDs (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24402/abstract).

For the longest follow-up (median 7 days, range 5–14 days), the NMA included 10 RCTs (1,731 patients, 6 interventions) (16–18,21,23,26,27,31–33). From the 15 pairwise comparisons between interventions, direct comparisons proved available for 6. Acetic acid derivative NSAIDs showed an important average improvement relative to baseline on joint tenderness at the longest follow-up (MD –1.77 [95% CI –1.83, –1.71], very low certainty; MID –1). There were no convincing differences between any of the interventions and the reference standard, acetic acid derivative NSAIDs (see Supplementary Table 2).

Joint swelling. Seven RCTs (969 patients, 6 interventions) described the change of joint swelling from baseline on day 2 (16,18,25,31–33). The reference (i.e., acetic acid derivative NSAIDs) showed an important average improvement relative to baseline on joint swelling at day 2 (MD –0.89 [95% CI –1.02, –0.76] on a 0–3 scale, very low certainty; MID –1) (Table 3). Of the 15 pairwise comparisons between interventions, direct evidence proved available for 6. Canakinumab was the only intervention that may be better than acetic acid derivative NSAIDs for improving joint swelling at day 2 (MD –0.61 [95% CI –1.01, –0.21], low certainty; MID –1), but the difference between canakinumab and acetic acid derivative NSAIDs was unimportant (smaller than the MID of 1 point reduction). Acetic acid derivative NSAIDs were probably superior to ibuprofen NSAIDs in joint swelling at day 2 (MD –0.29 [95% CI –0.56, –0.02], moderate certainty). There were no convincing differences between intravenous or intramuscular corticosteroids, oral corticosteroids, COX-2 highly selective NSAIDs, and the reference standard, acetic acid derivative NSAIDs (see Supplementary Table 2).

The NMA for change in joint swelling at the longest follow-up (median 7 days, range 5–14 days) comprised 11 RCTs (1,741 patients, 6 interventions) (16–18,23,25–27,31–33), including direct evidence for 6 of 15 pairwise comparisons. Acetic acid derivative NSAIDs showed an important average improvement relative to baseline on joint swelling at the longest follow-up (MD –1.63 [95% CI –1.70, –1.56], very low certainty; MID –1). There were no convincing differences between the reference standard and any of the other interventions (see Supplementary Table 2).

PtGA. Three RCTs reported PtGA of change from baseline at day 2 (16,18,20). The reference (i.e., acetic acid derivative NSAIDs) showed an important average improvement relative to baseline on PtGA at day 2 (MD –1.47 [95% CI –1.60, –1.34] on a 0–4 scale, very low certainty) (Table 3). The NMA for change in PtGA at day 2 included 3 RCTs (460 patients, 3 interventions).

Table 3. Most and least efficacious or safe treatment for all the outcomes*

Intervention	Pain score†		Joint tenderness‡		Joint swelling§		Patient global assessment¶		SAE RD (95% CI)
	Day 2	Longest follow-up	Day 2	Longest follow-up	Day 2	Longest follow-up	Day 2	Longest follow-up	Longest follow-up
Change from baseline or baseline risk in ref. group#									
Acetic acid derivative NSAIDs	−30.67 (−31.89, −29.45)**	−40.09 (−42.25, −39.61)**	−1.29 (−1.38, −1.21)**	−1.77 (−1.83, −1.71)††	−0.89 (−1.02, −0.76)‡‡	−1.63 (−1.70, −1.56)††	1.47 (−1.60, −1.34)††	−1.64 (−1.74, −1.53)††	0.025 (0.018, 0.035)**
Relative effect in relative to ref.§§									
Canakinumab	−41.12 (53.36, −29.11)¶¶	−12.84 (−20.76, −4.91)¶¶	−0.67 (−1.03, −0.3)¶¶	−0.42 (−0.86, 0.03)††	−0.61 (−1.01, −0.21)††	−0.28 (−0.71, 0.16)††	–	–	0.03 (−0.01. 0.06)**
Corticosteroids: IM or IV	−30.72 (−40.89, −20.79)‡‡	−5.71 (−12.36, 0.79)**	−0.33 (−0.68, 0.01)**	0 (−0.33, 0.33)††	−0.3 (−0.67, 0.08)‡‡	−0.03 (−0.44, 0.37)††	–	–	0 (−0.03, 0.02)**
COX-2 highly selective NSAIDs	1.85 (−2.31, 6.01)**	0.32 (−3.01, 3.65)**	−0.05 (−0.18, 0.08)**	−0.01 (−0.1, 0.08)††	0.1 (−0.23, 0.43)‡‡	−0.07 (−0.19, 0.05)††	−0.01 (−1, 0.98)††	0.095 (−0.08, 0.27)††	0 (−0.01, 0)**
Corticosteroids: oral	4.62 (−1.39, 10.63)**	−0.32 (−4.91, 4.12)**	−0.19 (−0.48, 0.1)**	−0.03 (−0.14, 0.08)††	−0.1 (−0.45, 0.25)‡‡	−0.21 (−0.56, 0.12)††	–	–	−0.03 (−0.05, −0.01)††
Ibuprofen NSAIDs	6.24 (−2.08, 14.78)**	3.8 (−4.12, 11.73)**	0.16 (−0.08, 0.41)**	−0.19 (−0.08, 0.46)††	0.29 (0.02, 0.56)##	−0.04 (−0.36, 0.29)††	0.21 (−0.56, 0.98)††	0.44 (0.02, 0.86)##	−0.02 (−0.04, 0.01)**
Rilonacept	−11.78 (−23.56, 0)‡‡	−3.17 (−10.94, 4.6)**	–	–	–	–	–	–	0 (−0.03, 0.03)**
IL-1 inhibition + acetic acid derivative NSAIDs	−6.47 (−18.02, 5.31)**	−1.59 (−9.35, 6.18)**	–	–	–	–	–	–	0.04 (−0.01, 0.09)**
Colchicine	10.63 (−2.54, 24.02)**	4.91 (−5.39, 15.37)**	–	–	–	–	–	–	−0.02 (−0.04, 0.01)**
Pyrazolidine derivative NSAIDs	–	–	–	–	–	–	–	–	0 (−0.04, 0.03)**
ACTH	–	–	–	–	–	–	–	–	0 (−0.05, 0.05)**
COX-2 selective NSAIDs	–	–	–	–	–	–	–	–	0 (−0.08, 0.08)**
Fenamate NSAIDs	–	–	–	–	–	–	–	–	0 (−0.11, 0.11)**

* Values are the mean difference (MD) (95% confidence interval [95% CI]) unless indicated otherwise. MDs in the natural units of standard scales for continuous outcomes and risk differences (RDs) for dichotomous outcome are shown. Effectiveness outcomes included pain score, joint tenderness, joint swelling, and patient global assessment, and safety outcomes included serious adverse events (SAEs). ACTH = adrenocorticotropic hormone; COX-2 = cyclooxygenase 2; IL-1 = interleukin-1; IM = intramuscular; IV = intravenous; MID = minimum important difference; NSAIDs = nonsteroidal antiinflammatory drugs; ref. = reference; VAS = visual analog scale.
† Mean change, 100-mm VAS for pain, where 0 mm = no pain, 100 mm = unbearable pain; MID = –19.
‡ Mean reduction, 4-point standard Likert scale for joint tenderness, where 0 = no pain, 3 = pain, winces, and withdraws; MID = –1.
§ Mean reduction, 4-point standard Likert scale for joint swelling, where 0 = no swelling, 3 = bulging beyond the joint margins; MID = –1.
¶ Mean change, 5-point standard Likert scale for patient global assessment, where 0 = excellent, 4 = poor.
# The reference was acetic acid derivative NSAIDs for all the outcomes. For continuous outcomes, the effect of the reference was the change from baseline at a particular time point in the acetic acid derivative NSAIDs arm across trials; for dichotomous outcomes, the effect was the risk of the outcome in the acetic acid derivative NSAIDs arm across trials (the baseline risk).
** Interventions showing least effectiveness and safety.
†† Interventions showing most effectiveness and safety.
‡‡ Interventions showing intermediate effectiveness and safety.
§§ Cell values represent the effect of the treatment in each row when compared to the reference (e.g., canakinumab resulted in pain reduction 41.12 units greater than acetic acid derivative NSAIDs, or a 71.79-unit reduction from baseline).
¶¶ Interventions showing good patient outcomes.
Interventions showing inferior patient outcomes, including SAEs.

Of the 4 pairwise comparisons between intervention, direct evidence proved available for only 1 comparison. There were no convincing differences between any of the interventions (see Supplementary Table 2).

The NMA for change in PtGA at the longest follow–up (median 7 days, range 5–8 days) included 5 RCTs (638 patients, 3 interventions) (16–18,23,26) including direct evidence for 1 of 3 pairwise comparisons. Acetic acid derivative NSAIDs showed an important average improvement relative to baseline on PtGA at the longest follow-up (MD –1.64 [95% CI –1.74, –1.53], very low certainty). Ibuprofen NSAIDs were probably worse than acetic acid derivative NSAIDs (MD 0.44 [95% CI 0.02, 0.86], moderate certainty). There were no convincing differences between COX-2 highly selective NSAIDs and acetic acid derivative NSAIDs (see Supplementary Table 2).

SAE. The NMA for SAEs included 29 RCTs (4,248 patients, 13 interventions) (15–23,26–44) and 78 paired estimates, of which 15 had both direct and indirect evidence and 58 had only indirect evidence. The median duration of available follow-up was 8 days (range 5–365 days). Oral corticosteroids were the only intervention that may be safer than acetic acid derivative NSAIDs (RD –0.03 [95% CI –0.05, –0.01], very low certainty). There were no convincing differences between any of the other interventions (see Supplementary Table 2).

The only SAE reported in the oral corticosteroids group was a case of low potassium associated with prednisolone. The main SAEs associated with acetic acid derivative NSAIDs were gastrointestinal events, including gastric or gastroduodenal ulcers, abdominal pain, and vomiting. SAEs reported in the COX-2 highly selective NSAIDs group were mainly in the urinary system and included renal calculi, uronephrosis, and renal failure. Serious infections and cardiovascular events were reported in the canakinumab group. However, the causality between the SAE and canakinumab was not reported. Among the 3 canakinumab trials, 2 trials found increased risk of infection associated with canakinumab during a 6-month follow-up (incidence of infection of 18.8% and 22.1% in canakinumab groups, 8.8% and 15.7% in triamcinolone groups), while the other small trial failed to find any difference in a follow-up of 8 weeks (incidence of infection of 7% in both groups) (25,33,44).

One trial not included in the NMA reported no significant difference between anakinra versus a free choice of colchicine, naproxen, or prednisolone in pain reduction, joint tenderness improvement, joint swelling improvement, PtGA, or SAE (14) (for details on the effects of each intervention see Supplementary Table 2).

DISCUSSION

The results of this NMA highlight a potential advantage of canakinumab versus other antiinflammatory interventions for gout flares in pain reduction at day 2 and the longest follow-up (moderate certainty). Canakinumab also showed larger effects on joint tenderness and joint swelling at day 2 (moderate certainty; low certainty), but the differences were unimportant (smaller than the MIDs) (Table 3). Among the commonly used therapies for gout flares (i.e., NSAIDs, colchicine, and corticosteroids), intravenous or intramuscular corticosteroids may be more effective than COX-2 highly selective NSAIDs, ibuprofen NSAIDs, acetic acid derivative NSAIDs, and oral corticosteroids on pain reduction at short-term (low certainty) (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24402/abstract). Ibuprofen NSAIDs were probably worse than acetic acid derivative NSAIDs in joint swelling at day 2 and PtGA at the longest follow-up (moderate certainty) (Table 3). For the safety evaluation, oral corticosteroids may cause fewer SAEs than acetic acid derivative NSAIDs (very low certainty) (Table 3). Results showed no convincing differences in safety among the other pharmacologic interventions.

Our study has several strengths. Using rigorous NMA methods, we incorporated direct and indirect evidence of the comparative efficacy and safety of antiinflammatory treatment for gout flares. We used the GRADE approach to assess the certainty of evidence informing the estimates. The outcomes evaluated in this review are important from both patient and provider points of view (45). For enhancing the interpretability of results, we converted the SMDs from NMA into MDs in the natural units of standard instruments and compared the MDs to the MIDs. We estimated the efficacy or baseline risk of the reference group (i.e., acetic acid derivative NSAIDs), facilitating the interpretation of comparative efficacy and safety of other pharmacologic interventions in relation to the reference. Moreover, the approach of making a conclusion from an NMA enabled a transparent, straightforward process of classifying interventions according to their relative benefit and harm. Our review also includes recently published studies that were not included in prior reviews and summarizes all the available RCT evidence.

In terms of limitations of the present review, in order to deal with the large number of interventions and relatively small number of trials for each intervention, we created clusters of interventions, taking the risk that effects would differ across treatments within clusters. Another limitation is that the degree to which the apparent improvement is due to natural history or placebo effects is uncertain, because the effect of the reference treatment was based on a before–after comparison in the included RCTs. A third limitation is that 3 of the RCTs that enrolled patients with difficult-to-treat gouty arthritis might cause heterogeneity and intransitivity (24,25,33,44). Additionally, we planned to conduct subgroup analyses based on the number of joints involved, pain levels, duration of the flare at presentation, duration of antiinflammatory therapy, and dose of the agent. Few trials, however, assessed differences in the relative effects of the interventions by patient characteristics. Information to inform subgroup analysis based on patient characteristics was therefore unavailable. As there were multiple interventions in

some categories, we are unable to compare efficacy and safety between different dosing. Furthermore, evaluation of rare event AEs would be underpowered in RCTs.

Previous systematic reviews that evaluated only direct estimates did not report important differences in pain reduction between canakinumab and intravenous or intramuscular corticosteroids versus other pharmaceutical interventions (46–48). The difference is likely due to the enhanced precision of estimates that this NMA provides, through inclusion of more studies and consideration of both direct and indirect evidence.

A Cochrane systematic review and a systematic review in support of the ACP guidelines found no difference in pain relief between NSAIDs and oral glucocorticoids (48). The Cochrane systematic review also indicated no difference between conventional NSAIDs and selective COX-2 inhibitor in pain relief, swelling, and global improvement (49). In the present systematic review, we categorized NSAIDs into subgroups according to the pharmacologic mechanism of action, which enabled the comparison within NSAIDs and the comparison between subcategory of NSAIDs and other interventions. We found consistent results that NSAIDs were not different with oral glucocorticoids in effectiveness outcomes (see Supplementary Table 2). However, ibuprofen NSAIDs were inferior to acetic acid derivative NSAIDs in resolution of joint swelling at day 2 and improvement of PtGA at longest follow-up (Table 3). Another Cochrane systematic review of colchicine for acute gout identified no studies comparing colchicine to any other active treatment (50). In our NMA, colchicine compared indirectly with other interventions, although ibuprofen NSAIDs were shown to be inferior to canakinumab, rilonacept, and intravenous or intramuscular corticosteroids, but showed no difference with other interventions (see Supplementary Table 2).

Cost or financial barriers to medications are not considered in this systematic review. Although our review highlights potential advantages of canakinumab in terms of effectiveness, cost and the administration route have limited its use (51). Inherent delays with prior authorization requirements likely limit the practical use of canakinumab for management of gout flare. These issues have been explicitly considered and addressed in the 2020 ACR Guideline for the Management of Gout (7). In the present review, among the 3 canakinumab trials, 2 trials found increased risk of infection associated with canakinumab while the other 1 small trial failed to find any difference (25,33,44). Future RCTs and observational studies are needed to evaluate the safety of canakinumab in this regard.

Future studies need to evaluate the comparative efficacy and safety of pharmacologic interventions used commonly in practice but not yet tested in RCTs (e.g., colchicine, pyrazolidine derivative NSAIDs, COX-2 selective NSAIDs, and fenamate NSAIDs). RCTs are also needed to evaluate IL-inhibitors other than canakinumab. Experts writing in prior guidelines have suggested evaluating the efficacy and safety of combination-drug treatments for gout flares (e.g., IL-1 inhibitor plus acetic acid derivative) (6). Future studies should report data for relevant patient subgroups (e.g., those with polyarticular gout or subgroups based on flare severity), thus enabling subgroup analysis of patients with different characteristics in subsequent systematic reviews.

In summary, the present systematic review provides a current, comprehensive summary of the comparative efficacy and safety of pharmacologic interventions used in clinical practice for antiinflammatory treatment in patients with gout flare. Canakinumab may be superior to other alternatives and intravenous or intramuscular corticosteroids may be the second-best treatment in terms of pain reduction at day 2. Acetic acid derivative NSAIDs may be superior to ibuprofen NSAIDs on the improvement of joint swelling at day 2 and PtGA at the longest follow-up.

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. Brignardello-Petersen 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. Neogi, Fitzgerald, Dalbeth, Mikuls, Guyatt, Brignardello-Petersen.

Acquisition of data. Zeng, Qasim, Brignardello-Petersen.

Analysis and interpretation of data. Zeng, Qasim, Neogi, Fitzgerald, Dalbeth, Mikuls, Guyatt, Brignardello-Petersen.

REFERENCES

1. Qaseem A, Harris RP, Forciea MA. Management of acute and recurrent gout: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017;166:58–68.
2. Bardin T, Voshaar MA, van de Laar MA. The human and economic burden of difficult-to-treat gouty arthritis. Joint Bone Spine 2015;82(Suppl 1):eS2–8.
3. Dalbeth N, Merriman TR, Stamp LK. Gout. Lancet 2016;388:2039–52.
4. Richette P, Doherty M, Pascual E, Barskova V, Becce F, Castañeda-Sanabria J, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis 2017;76:29–42.
5. Hui M, Carr A, Cameron S, Davenport G, Doherty M, Forrester H, et al. The British Society for Rheumatology guideline for the management of gout. Rheumatology 2017;56:e1–20.
6. Khanna D, Fitzgerald JD, Khanna PP, Bae S, Singh MK, Neogi T, et al. 2012 American College of Rheumatology guidelines for management of gout. Part 1: systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. Arthritis Care Res (Hoboken) 2012;64:1431–46.
7. FitzGerald J, Dalbeth N, Mikuls T, Brignardello-Petersen R, Guyatt G, Abeles AM, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Care Res (Hoboken) 2020;72:744–760.
8. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med 2015;162:777–84.
9. Schumacher HR, Taylor W, Edwards L, Grainger R, Schlesinger N, Dalbeth N, et al. Outcome domains for studies of acute and chronic gout. J Rheumatol 2009;36:2342–5.
10. Taylor WJ, Redden D, Dalbeth N, Schumacher HR, Edwards NL, Simon LS, et al. Application of the OMERACT filter to measures

of core outcome domains in recent clinical studies of acute gout. J Rheumatol 2014;41:574–80.

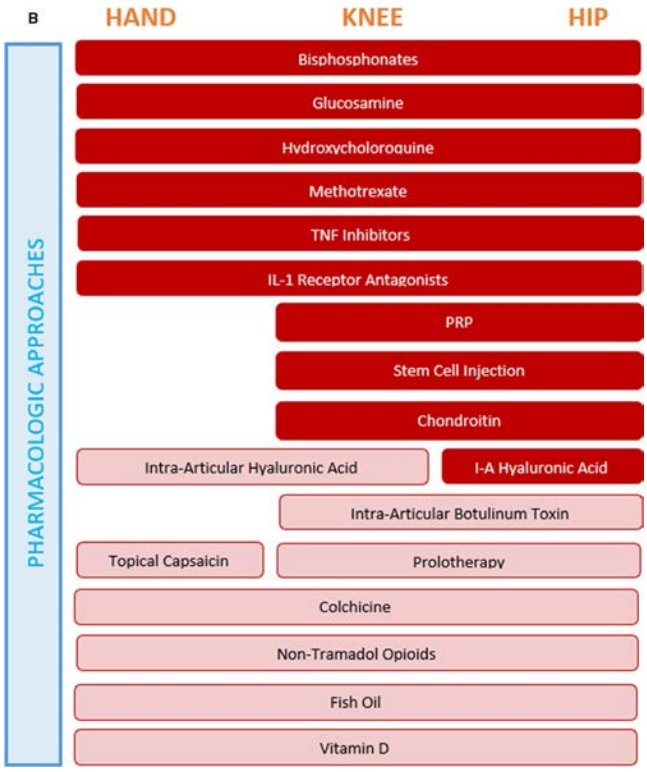
11. Brignardello-Petersen R, Bonner A, Alexander PE, Siemieniuk RA, Furukawa TA, Rochwerg B, et al. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. J Clin Epidemiol 2018;93:36–44.
12. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2005. URL: http://www.R-project.org.
13. Brignardello-Petersen R, Florez I D, Izcovich A, Santesso N, Hazlewood G, Alhazanni W, et al. GRADE approach to drawing conclusions from a network meta-analysis using a minimally contextualised framework BMJ 2020;371:m3900.
14. Janssen CA, Oude Voshaar MA, Vonkeman HE, Jansen TL, Janssen M, Kok MR, et al. Anakinra for the treatment of acute gout flares: a randomized, double-blind, placebo-controlled, active-comparator, non-inferiority trial. Rheumatology 2019;58:1344–52.
15. Janssens HJ, Janssen M, Van de Lisdonk EH, Van Riel PL, Van Weel C. Use of oral prednisolone or naproxen for the treatment of gout arthritis: a double-blind, randomised equivalence trial. Lancet 2008;371:1854–60.
16. Lederman R. A double-blind comparison of etodolac (Lodine) and high doses of naproxen in the treatment of acute gout. Adv Ther 1990;7:344–54.
17. Li T, Chen SL, Dai Q, Han XH, Li ZG, Wu DH, et al. Etoricoxib versus indometacin in the treatment of Chinese patients with acute gouty arthritis: a randomized double-blind trial. Chin Med J 2013;126:1867–71.
18. Maccagno A, Di Giorgio E, Romanowicz A. Effectiveness of etodolac ('Lodine') compared with naproxen in patients with acute gout. Curr Med Res Opin 1991;12:423–9.
19. Man CY, Cheung IT, Cameron PA, Rainer TH. Comparison of oral prednisolone/paracetamol and oral indomethacin/paracetamol combination therapy in the treatment of acute goutlike arthritis: a double-blind, randomized, controlled trial. Ann Emerg Med 2007;49:670–7.
20. Navarra S, Rubin BR, Yu Q, Smugar SS, Tershakovec AM. Association of baseline disease and patient characteristics with response to etoricoxib and indomethacin for acute gout. Curr Med Res Opin 2007;23:1685–91.
21. Rainer TH, Cheng CH, Janssens HJ, Man CY, Tam LS, Choi YF, et al. Oral prednisolone in the treatment of acute gout: a pragmatic, multicenter, double-blind, randomized trial. Ann Intern Med 2016;164:464–71.
22. Roddy E, Clarkson K, Blagojevic-Buchnall M, Mehta R, Oppong R, Avery A, et al. Open-label randomised pragmatic trial (CONTACT) comparing naproxen and low-dose colchicine for the treatment of gout flares in primary care. Ann Rheum Dis 2020;79:276–84.
23. Rubin BR, Burton R, Navarra S, Antigua J, Londoño J, Pryhuber KG, et al. Efficacy and safety profile of treatment with etoricoxib 120 mg once daily compared with indomethacin 50 mg three times daily in acute gout: a randomized controlled trial. Arthritis Rheum 2004;50:598–606.
24. So A, De Meulemeester M, Pikhlak A, Yücel AE, Richard D, Murphy V, et al. Canakinumab for the treatment of acute flares in difficult-to-treat gouty arthritis: results of a multicenter, phase II, dose-ranging study. Arthritis Rheum 2010;62:3064–76.
25. Schlesinger N, Alten RE, Bardin T, Schumacher HR, Bloch M, Gimona A, et al. Canakinumab for acute gouty arthritis in patients with limited treatment options: results from two randomised, multicentre, active-controlled, double-blind trials and their initial extensions. Ann Rheum Dis 2012;71:1839–48.
26. Schumacher HR Jr, Boice JA, Daikh DI, Mukhopadhyay S, Malmstrom K, Ng J, et al. Randomised double blind trial of etoricoxib and indometacin in treatment of acute gouty arthritis. BMJ 2002;324:1488–92.
27. Schumacher HR, Berger MF, Li-Yu J, Perez-Ruiz F, Burgos-Vargas R, Li C. Efficacy and tolerability of celecoxib in the treatment of acute gouty arthritis: a randomized controlled trial. J Rheumatol 2012;39:1859–66.
28. Terkeltaub RA, Schumacher HR, Carter JD, Baraf HS, Evans RR, Wang J, et al. Rilonacept in the treatment of acute gouty arthritis: a randomized, controlled clinical trial using indomethacin as the active comparator. Arthritis Res Ther 2013;15:R25.
29. Willburger RE, Mysler E, Derbot J, Jung T, Thurston H, Kreiss A, et al. Lumiracoxib 400 mg once daily is comparable to indomethacin 50 mg three times daily for the treatment of acute flares of gout. Rheumatology (Oxford) 2007;46:1126–32.
30. Xu Z, Zhang R, Zhang D, Yao J, Shi R, Tang Q, et al. Peptic ulcer hemorrhage combined with acute gout: analyses of treatment in 136 cases. Int J Clin Exp Med 2015;8:6193–9.
31. Xu L, Liu S, Guan M, Xue Y. Comparison of prednisolone, etoricoxib, and indomethacin in treatment of acute gouty arthritis: an open-label, randomized, controlled trial. Med Sci Monit 2016;22:810–7.
32. Zhang YK, Yang H, Zhang JY, Song LJ, Fan YC. Comparison of intramuscular compound betamethasone and oral diclofenac sodium in the treatment of acute attacks of gout. Int J Clin Pract 2014;68:633–8.
33. Schlesinger N, De Meulemeester M, Pikhlak A, Yücel AE, Richard D, Murphy V, et al. Canakinumab relieves symptoms of acute flares and improves health-related quality of life in patients with difficult-to-treat gouty arthritis by suppressing inflammation: results of a randomized, dose-ranging study. Arthritis Res Ther 2011;13:R53.
34. Alloway JA, Moriarty MJ, Hoogland YT, Nashel DJ. Comparison of triamcinolone acetonide with indomethacin in the treatment of acute gouty arthritis. J Rheumatol 1993;20:111–3.
35. Altman RD, Honig S, Levin JM, Lightfoot RW. Ketoprofen versus indomethacin in patients with acute gouty arthritis: a multicenter, double blind comparative study. J Rheumatol 1988;15:1422–6.
36. Axelrod D, Preston S. Comparison of parenteral adrenocorticotropic hormone with oral indomethacin in the treatment of acute gout. Arthritis Rheum 1988;31:803–5.
37. Butler RC, Goddard DH, Higgens CS, Hollingworth F, Pease CT, Stodell MA, et al. Double-blind trial of flurbiprofen and phenylbutazone in acute gouty arthritis. Br J Clin Pharmacol 1985;20:511–3.
38. Cheng TT, Lai HM, Chiu CK, Chen YC. A single-blind, randomized, controlled trial to assess the efficacy and tolerability of rofecoxib, diclofenac sodium, and meloxicam in patients with acute gouty arthritis. Clin Ther 2004;26:399–406.
39. Douglas G, Thompson M. A comparison of phenylbutazone and flufenamic acid in the treatment of acute gout. Rheumatology 1970;10:275–80.
40. Eberl R, Dunky A. Meclofenamate sodium in the treatment of acute gout. Results of a double-blind study. Arzneimittelforschung 1983;33:641–3.
41. Fraser RC, Davis RH, Walker FS. Comparative trial of azapropazone and indomethacin plus allopurinol in acute gout and hyperuricaemia. J R Coll Gen Pract 1987;37:409–11.
42. Siegel LB, Alloway JA, Nashel DJ. Comparison of adrenocorticotropic hormone and triamcinolone acetonide in the treatment of acute gouty arthritis. J Rheumatol 1994;21:1325–7.
43. Smyth CJ, Percy JS. Comparison of indomethacin and phenylbutazone in acute gout. Ann Rheum Dis 1973;32:351–3.
44. Hirsch JD, Gnanasakthy A, Lale R, Choi K, Sarkin AJ. Efficacy of canakinumab vs. triamcinolone acetonide according to multiple gouty arthritis-related health outcomes measures. Int J Clin Pract 2014;68:1503–7.

45. Dalbeth N, Zhong CS, Grainger R, Khanna D, Khanna PP, Singh JA, et al. Outcome measures in acute gout: a systematic literature review. J Rheumatol 2014;41:558–68.
46. Sivera F, Wechalekar MD, Andrés M, Buchbinder R, Carmona L. Interleukin-1 inhibitors for acute gout. Cochrane Database Syst Rev 2014:CD009993.
47. Janssens HJ, Lucassen PL, Van de Laar FA, Janssen M, Van de Lisdonk EH. Systemic corticosteroids for acute gout. Cochrane Database Syst Rev 2008:CD005521.
48. Shekelle PG, Newberry SJ, FitzGerald JD, Motala A, O'Hanlon CE, Tariq A, et al. Management of gout: a systematic review in support of an American College of Physicians clinical practice guideline. Ann Intern Med 2017;166:37–51.
49. Van Durme CM, Wechalekar MD, Buchbinder R, Schlesinger N, van der Heijde D, Landewé RB. Non-steroidal anti-inflammatory drugs for acute gout. Cochrane Database Syst Rev 2014:CD010120.
50. Van Echteld I, Wechalekar MD, Schlesinger N, Buchbinder R, Aletaha D. Colchicine for acute gout. Cochrane Database Syst Rev 2014:CD006190.
51. National Institute of Clinical Excellence. Gouty arthritis: canakinumab. Evidence summary ESNM23 2013. URL: https://www.nice.org.uk/advice/esnm23/chapter/Keypoints-from-the-evidence

Errata

DOI 10.1002/acr.24615

In the article by Kolasinski et al in the February 2020 issue of *Arthritis Care & Research* (2019 American College of Rheumatology/Arthritis Foundation Guideline for the Management of Osteoarthritis of the Hand, Hip, and Knee [pages 149–162]), there was an error in Figure 2B: Chondroitin should not have been included among the pharmacologic therapies recommended against for hand osteoarthritis. The corrected Figure 2B is shown below.



We regret the error.

DOI 10.1002/acr.24614

In the article by Buckley et al published in the February 2021 issue of *Arthritis Care & Research* (pages 215-220), acknowledgement of study funding was inadvertently omitted. The following statement should have been included on the first page of the article: "Supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases Training Program in Rheumatic Diseases (T-32-AR007442)."

We regret the error.

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.