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 71, No. 7, July 2019

Arthritis Care & Research

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

Cover design: Sandra Pulmano

♾ This journal is printed on acid-free paper.

Arthritis Care & Research

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

VOLUME 71 • July 2019 • NO. 7

Special Article

Osteoarthritis

Systemic Lupus Erythematosus

Rheumatoid Arthritis

Sjögren's Syndrome

Cover image: The figure on the cover (from Birnbaum et al, page 936) is a skin biopsy specimen taken from the proximal thigh in a patient with a non–length dependent pattern of neuropathic pain.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 853–854
DOI 10.1002/acr.23714

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

How Should We Consider Lupus Without Antinuclear Antibodies?

David I. Daikh[1] and Martin Aringer[2]

Most patients with systemic lupus erythematosus (SLE) are antinuclear antibody (ANA) positive, but some are not (1). Although this difference represents a small number of patients, the size of this minor population is still debated. A fundamental question regarding this subset of patients is whether SLE can develop in the absence of an autoantibody response to nuclear antigens, with the alternative concept that ANA negativity reflects a testing anomaly. Some of the imperfect sensitivity is definitely due to inadequate testing and technical issues (2). Patients may also have autoantibody subsets with preferential or exclusive cytoplasmic staining. Moreover, ANA may be reduced below the detectable threshold by therapy. This possibility may be more likely if positive ANA results are exclusively based on antibodies that correlate with disease activity, such as those to double-stranded DNA (dsDNA) or histones.

In this issue, Choi et al report on a very thorough analysis of ANA in the SLE International Collaborating Clinics inception cohort that attempts to control for all these technical issues (3). The authors used an experienced laboratory to analyze the first available serum of 1,137 patients, where the same technician used the high-quality gold standard method of indirect immunofluorescence to measure ANA. Notably, however, the HEp-2000 cell was used as a substrate in this study. This cell is enriched with Ro antigen and had a relatively high false negative rate in a previous study (2).

In the analysis of Choi et al, 1,049 of the 1,137 patients (92.3%) had positive nuclear staining at a 1:160 titer or higher. An additional 17 patients (1.5%) had isolated cytoplasmic staining, which supports the idea that cytoplasmic staining should be reported. The patients with negative HEp-2000 immunofluorescence in this cohort were largely white, had lower disease activity, and had more commonly received glucocorticoids.

Of the remaining 71 patients without detected indirect immunofluorescence staining, 16 had anti-Ro antibodies and 8 had anti-dsDNA antibodies. Accordingly, despite considerable expertise and an optimized substrate, there are clearly technical laboratory issues that can influence ANA sensitivity. These issues can be far more significant with other technology and in laboratories without adequate immunologic experience (4,5). This problem presumably exists for diagnosing other connective tissue diseases (CTDs) as well, for which ANA is used as a screening test. It is not clear how many of the remaining 47 patients had an initial ANA test after SLE treatment was initiated and therefore may have converted to ANA negative. Regardless, the study confirms that the frequency of true ANA-negative SLE is very low.

Based on this idea, Choi et al propose replacing the well-established term CTD (6) with ANA-associated rheumatic disease (AARD). While we strongly agree with the underlying concept that the development of ANA reflects a fundamental aspect of the pathogenesis of SLE and likely other CTDs, we view the proposed term critically. These diseases were originally referred to as CTDs based on a concept that is not entirely correct from today's point of view. The same could also be said for rheumatoid arthritis (RA), or even for systemic lupus erythematosus. However, CTD in the more narrow sense, not including RA, remains a useful descriptive term, rather than a precise concept. Additional terms, such as mixed connective tissue disease (5) and undifferentiated connective tissue disease have followed, and every rheumatologist still knows what is meant by CTD (7,8).

Would the same be true for AARD, and would this term be more precise to a degree supporting a fundamental name change? This appears unlikely in two respects. First, the authors highlight two important issues in their manuscript that do not support the term, namely the fact that there are ANA-negative patients with SLE and the fact that cytoplasmic staining is of importance. The consequence of the former issue would be referring to ANA-negative ANA-associated rheumatic disease. The latter issue could bring up the suggestion to replace the term ANA by the

[1]David I. Daikh, MD, PhD: San Francisco Veterans Affairs Medical Center and University of California, San Francisco; [2]Martin Aringer, MD: Medical Center and Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany.

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

Address correspondence to Martin Aringer, MD, Professor of Medicine, Division of Rheumatology, Department of Medicine IIII, Medical Center and Faculty of Medicine Carl Gustav Carus at the TU Dresden, Fetscherstrasse 74, 01307 Dresden, Germany. E-mail: Martin.Aringer@uniklinikum-dresden.de.

Submitted for publication July 16, 2018; accepted in revised form July 17, 2018.

term anticellular antibodies to be more accurate, thus leaving rheumatologists with a combination of two incomplete concepts.

Second, while ANA tests are used as a screening test for CTDs, patients with many other diseases that may or may not have major similarities in pathogenesis, such as RA or autoimmune thyroid disease, commonly have positive ANA results. This limitation in specificity in comparison to its high sensitivity has led to a reconsideration of the original position of ANA in a recent SLE classification approach jointly supported by the European League Against Rheumatism (EULAR) and the American College of Rheumatology (9). These novel criteria have not yet been published in full, but validation results were presented at the EULAR congress in June 2018 (10). Rather than providing clarity, lumping diseases together under the banner of such a nonspecific test will only serve to confuse less experienced practitioners and students. The goal of such designations should be to clarify and illuminate conceptually unifying features of disease, for example on the basis of pathogenesis.

The main focus of Choi et al, however, is a different and very valid point. ANA testing requires an appropriate substrate as well as great care and experience. Importantly, we should not miss the additional information obtained by taking HEp-2 (or HEp-2000) cytoplasmic fluorescence into account. In this regard, we would fully concur with the argument of this study (3), which by its design and sample size is an important contribution to better understanding ANA performance in SLE.

AUTHOR CONTRIBUTIONS

Drs. Daikh and Aringer drafted the article, revised it critically for important intellectual content, and approved the final version to be published.

REFERENCES

1. Leuchten N, Hoyer A, Brinks R, Schoels M, Schneider M, Smolen J, et al. Performance of antinuclear antibodies for classifying systemic lupus erythematosus: a systematic literature review and meta-regression of diagnostic data. Arthritis Care Res (Hoboken) 2018;70:428–38.
2. Pisetsky DS, Spencer DM, Lipsky PE, Rovin BH. Assay variation in the detection of antinuclear antibodies in the sera of patients with established SLE. Ann Rheum Dis 2018;77:911–3.
3. Choi MY, Clarke AE, St. Pierre Y, Hanly JG, Urowitz MB, Romero-Diaz J, et al. Antinuclear antibody–negative systemic lupus erythematosus in an international inception cohort. Arthritis Care Res (Hoboken) 2019;71:893–902.
4. Bonilla E, Francis L, Allam F, Ogrinc M, Neupane H, Phillips PE, et al. Immunofluorescence microscopy is superior to fluorescent beads for detection of antinuclear antibody reactivity in systemic lupus erythematosus patients. Clin Immunol 2007;124:18–21.
5. Mariz HA, Sato E, Barbosa SH, Rodrigues SH, Dellavance A, Andrade LE. Pattern on the antinuclear antibody–Hep-2 test is a critical parameter for discriminating antinuclear antibody–positive healthy individuals and patients with autoimmune rheumatic diseases. Arthritis Rheum 2011;63:191–200.
6. Klemperer P. The concept of connective-tissue disease. Circulation 1962;25:869–71.
7. 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.
8. LeRoy EC, Maricq HR, Kahaleh MB. Undifferentiated connective tissue syndromes. Arthritis Rheum 1980;23:341–3.
9. Aringer M, Dorner T, Leuchten N, Johnson SR. Toward new criteria for systemic lupus erythematosus: a standpoint. Lupus 2016;25:805–11.
10. Aringer M, Costenbader KH, Brinks R, Boumpas D, Daikh D, Jayne D, et al. Validation of new systemic lupus erythematosus classification criteria [abstract]. Ann Rheum Dis 2018;77 Suppl 2:60.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 855–864
DOI 10.1002/acr.23716

American College of Rheumatology
Empowering Rheumatology Professionals

Cost-Effectiveness of Diet and Exercise for Overweight and Obese Patients With Knee Osteoarthritis

Elena Losina,[1] Karen C. Smith,[2] A. David Paltiel,[3] Jamie E. Collins,[1] Lisa G. Suter,[4] David J. Hunter,[5] Jeffrey N. Katz,[1] and Stephen P. Messier[6]

Objective. The Intensive Diet and Exercise for Arthritis (IDEA) trial showed that an intensive diet and exercise (D+E) program led to a mean 10.6-kg weight reduction and 51% pain reduction in patients with knee osteoarthritis (OA). The aim of the current study was to investigate the cost-effectiveness of adding this D+E program to treatment in overweight and obese (body mass index >27 kg/m^2) patients with knee OA.

Methods. We used the Osteoarthritis Policy Model to estimate quality-adjusted life-years (QALYs) and lifetime costs for overweight and obese patients with knee OA, with and without the D+E program. We evaluated cost-effectiveness with the incremental cost-effectiveness ratio (ICER), a ratio of the differences in lifetime cost and QALYs between treatment strategies. We considered 3 cost-effectiveness thresholds: $50,000/QALY, $100,000/QALY, and $200,000/QALY. Analyses were conducted from health care sector and societal perspectives and used a lifetime horizon. Costs and QALYs were discounted at 3% per year. D+E characteristics were derived from the IDEA trial. Deterministic and probabilistic sensitivity analyses (PSAs) were used to evaluate parameter uncertainty and the effect of extending the duration of the D+E program.

Results. In the base case, D+E led to 0.054 QALYs gained per person and cost $1,845 from the health care sector perspective and $1,624 from the societal perspective. This resulted in ICERs of $34,100/QALY and $30,000/QALY. In the health care sector perspective PSA, D+E had 58% and 100% likelihoods of being cost-effective with thresholds of $50,000/QALY and $100,000/QALY, respectively.

Conclusion. Adding D+E to usual care for overweight and obese patients with knee OA is cost-effective and should be implemented in clinical practice.

INTRODUCTION

Knee osteoarthritis (OA) is highly prevalent, affecting 14 million Americans, ~50% of whom are obese (1,2). Obesity and knee OA together result in 3.5 quality-adjusted life-years (QALYs) lost per person (1 QALY measures the equivalent of 1 year of perfect health) (2). Weight management and exercise are recommended by OA treatment guidelines, including those of the Osteoarthritis Research Society International and the American Academy of Orthopaedic Surgeons (3,4).

Other treatment options for knee OA have inherent limitations. Pharmacologic treatments, including nonsteroidal antiinflammatory drugs (NSAIDs) and opioids, are moderately efficacious, but their long-term use is frequently limited by side effects such as cardiovascular and gastrointestinal events and opioid addiction (5,6). While total knee arthroplasty (TKA) is both effective and cost-effective, it is generally reserved for patients in the later stages of OA progression (7,8).

Several randomized controlled trials (RCTs) have shown that weight loss is associated with reduced pain in patients with

Supported by the NIH (grants R01-AR-064320 and K24-AR-057827 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases).

[1]Elena Losina, PhD, Jamie E. Collins, PhD, Jeffrey N. Katz, MD, MSc: Orthopaedic and Arthritis Center for Outcomes Research (OrACORe), Policy and Innovation eValuation in Orthopaedic Treatments (PIVOT) Research Center, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [2]Karen C. Smith, BA: Orthopaedic and Arthritis Center for Outcomes Research (OrACORe), Policy and Innovation eValuation in Orthopaedic Treatments (PIVOT) Research Center, Brigham and Women's Hospital, Boston, Massachusetts; [3]A. David Paltiel, PhD, MBA: Yale School of Public Health, Yale University, New Haven, Connecticut; [4]Lisa G. Suter, MD: Yale School of Medicine Yale University, New Haven, Connecticut, and Veterans Affairs Medical Center, West Haven, Connecticut; [5]David J. Hunter, MBBS, MPH, ScD: Institute of Bone and Joint Research, Kolling Institute, University of Sydney and Royal North Shore Hospital, Sydney, Australia; [6]Stephen P. Messier, PhD: Wake Forest University, Winston-Salem, North Carolina.

Dr. Losina is a member of the Samumed Payers Advisory Board and has received consulting fees from TissueGene (less than $10,000 each). Dr. Hunter has received consulting fees from Flexion Therapeutics, Merck Serono, TissueGene, and TLC Biopharmaceuticals (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Elena Losina, PhD, Orthopaedic and Arthritis Center for Outcomes Research, Policy and Innovation eValuation in Orthopaedic Treatments (PIVOT) Research Center, Department of Orthopaedic Surgery, Brigham and Women's Hospital, 75 Francis Street, BTM Suite 5016, Boston, MA 02115. E-mail: elosina@partners.org.

Submitted for publication March 21, 2018; accepted in revised form July 24, 2018.

SIGNIFICANCE & INNOVATIONS

- While previous work has established the clinical efficacy (pain reduction) of diet and exercise programs for the treatment of knee osteoarthritis (OA), this study is the first to confirm that such programs also provide excellent economic value when compared to alternative uses of scarce OA treatment resources.
- We considered multiple willingness-to-pay thresholds ($50,000, $100,000, and $200,000 per quality-adjusted life-year [QALY]). In the base case, the diet and exercise program was cost-effective at all thresholds considered.
- If payers are willing to spend $50,000 per QALY gained, allowing patients to participate in the diet and exercise program for up to 8 years provides the best value. If they are willing to spend $100,000 per QALY gained, allowing patients to participate in the diet and exercise program indefinitely provides the best value.
- Programs to provide patients with knee OA with access to diet and exercise treatment should be implemented in clinical care.

knee OA (9). A meta-analysis of RCTs showed that a weight loss of 10% is expected to have a clinically relevant effect on disability (10). Participants randomized to a diet and exercise (D+E) regimen in the Arthritis, Diet, and Activity Promotion Trial (ADAPT) experienced a 30% reduction in pain over 18 months compared to 17% in the healthy lifestyle control group (11). The intent-to-treat analysis in the Intensive Diet and Exercise (IDEA) trial demonstrated that patients randomized to the D+E group experienced a 51% reduction in pain severity over 18 months compared to a 28% reduction in those randomized to receive exercise alone (the attention control). After 18 months, 38% of participants in the D+E group reported little or no pain compared to 22% of those in the exercise group and 20% of those in the diet group (12).

Economic evaluations of D+E regimens are scarce. A recent systematic review identified only 1 economic evaluation of an OA intervention targeting obesity: a cost-effectiveness analysis of the ADAPT trial (13,14). That analysis did not investigate the cost-effectiveness of D+E regimens in the context of other OA treatments. It was also limited to the 18-month trial time frame; this short-term time horizon is problematic, because OA is a chronic disease. The authors of the review concluded that there is a pressing need for long-term analyses that report cost per QALY outcomes to facilitate comparisons of D+E cost-effectiveness across health care sectors (13). In the current study, we addressed this gap in the literature with a formal cost-effectiveness analysis of adding a D+E regimen to usual care for overweight and obese patients with knee OA.

MATERIALS AND METHODS

Analytic overview. We used the Osteoarthritis Policy (OAPol) model, a validated, published computer simulation model of knee OA (2,7,15,16), to assess the cost-effectiveness of adding a D+E program to usual care for knee OA. OAPol is a state-transition Monte Carlo simulation model that estimates quality-adjusted life expectancy (QALE) and lifetime medical costs for patients with knee OA. A state-transition model characterizes the clinical progress in each patient with knee OA as a sequence of annual transitions between health states. Monte Carlo refers to the process of simulating one hypothetical patient with knee OA at a time and determining that patient's health state transitions with a set of transition probabilities.

Our primary outcome was the incremental cost-effectiveness ratio (ICER), a ratio of the differences in costs and QALYs gained between treatments. We considered a treatment cost-effective if its ICER was below a given willingness-to-pay (WTP) threshold. No single WTP threshold is used to make decisions in the US, and discussion about an appropriate threshold remains unsettled (17). In selecting a threshold, we sought to present how the preferred strategy depends on society's willingness to pay for an additional QALY and to provide comparative guidance on what ICER might be considered an acceptable value. To this end, we included 3 thresholds. A threshold of $50,000/QALY is commonly used in the field of cost-effectiveness analysis, and because some evidence suggests that $50,000/QALY is too low for US health care, we also included thresholds of $100,000/QALY and $200,000/QALY (18).

We conducted analyses from both the health care sector and societal perspectives, as recommended by the Second Panel on Cost-Effectiveness in Health and Medicine (19). The latter includes costs of caregiving and lost productivity due to OA pain and surgery. Supplementary Table 1 (available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract) outlines the cost and quality-of-life components of each perspective. All costs are in 2016 US dollars, and costs and QALYs were discounted by 3% annually.

OA policy model. The OAPol model simulates patients with knee OA based on demographic and clinical characteristics including age, BMI, comorbidities, and knee OA pain and structural severity. Model patients transition annually between health states, defined by obesity, comorbidities, and severity of knee OA. Time in each health state, including all associated costs and health-related quality of life effects, is accounted for from model entry until death.

QALE was estimated using data from the Osteoarthritis Initiative (OAI), a large longitudinal cohort study in patients with knee OA (20). The OAI measured health-related quality of life using the Short Form 12 (SF-12) health survey. We transformed

SF-12 responses into preference-based measures (i.e., utilities) using a previously published conversion algorithm (21). These utilities, stratified by age, comorbidities, obesity, and knee OA pain, were the weights used by the OAPol model to estimate quality-adjusted survival.

Model patients with a BMI ≥30 kg/m^2 are considered obese. Obesity lowers quality of life utility, increases the incidence of cardiovascular disease, diabetes mellitus, and cancer, and increases mortality (2,22,23). The model assumes that patients with a BMI of 18.5–24.99 kg/m^2 and those with a BMI of 25–29.99 kg/m^2 carry similar risks of comorbidities.

Knee OA treatments in the OAPol model include NSAIDs, TKA, and D+E. All treatments can affect quality of life utility by reducing knee OA–related pain. TKA also alters the presence of structural knee OA, and D+E reduces BMI. Each treatment has an associated cost and likelihood of toxicities. Toxicities carry their own costs and quality of life decrements.

OAPol includes all direct medical costs of knee OA treatment as well as non-OA medical costs stratified by age and comorbidities (see Supplementary Table 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract). Costs were adjusted for inflation using the method recommended by the Second Panel on Cost-Effectiveness in Medicine (Supplementary Appendix Section 1a, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract) (19).

Cohort characteristics. We derived cohort characteristics and quality of life utilities from the IDEA trial (Tables 1 and 2). The mean ± SD age of the patients was 66 ± 6 years, 72% were female, and the average BMI was 33.6 kg/m^2. The severity of OA-related pain was assigned using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale (24). The mean ± SD baseline WOMAC pain score was 32.5 ± 15.5 (0–100 scale, with 100 = worst). At the time of initiation of the trial, 13% of the cohort had diabetes mellitus, and 10% had cardiovascular disease. Other comorbidity incidence data were derived from National Health and Nutrition Examination Survey (NHANES) 2011–2014 data, and relative risks of mortality were derived using data from NHANES and Centers for Disease Control and Prevention Life Tables, 2011 (25,26).

Table 1. Cohort characteristics and indirect annual costs associated with OA*

Parameter	Value
Age, mean ± SD years	66 ± 6
Women, %	72
Race, %	
White non-Hispanic	81
African American non-Hispanic	6.3
African American Hispanic	6.3
White Hispanic	6.3
BMI, mean ± SD kg/m^2	33.6 ± 3.7
BMI at year 1, minimum, maximum	27, 41
K/L grade 2, %	50
K/L grade 3, %	50
WOMAC pain score, mean ± SD	
Year 1	32.5 ± 15.5
Subsequent year increase†	2 ± 10
Comorbidity prevalence, %	
Cardiovascular disease	10
Diabetes mellitus	13
Indirect costs, dollars	
Annual OA pain productivity cost‡	1,037
Annual OA caregiving cost‡	1,128
Productivity costs for TKA, year 1	3,311
Productivity costs for revision TKA (year 1)	3,592

* Data for patient characteristics were derived from reference 12. Data for indirect costs were derived from references 15, 43, and 50. BMI = body mass index; K/L = Kellgren/Lawrence (grade range 0–4, with 4 representing most severe). TKA = total knee arthroplasty
‡ Costs incurred by patients with knee osteoarthritis (OA) with a Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain score of >40 in the base case.

Treatment strategies. The duration, efficacy, discontinuation rate, and cost of the usual care and D+E regimens are described below. Patients receiving usual care began their treatment sequence at year 1. Patients receiving D+E began both the D+E regimen and the usual care treatment in parallel at year 1.

Usual care. Usual care treatment consisted of a pharmacologic NSAID regimen, followed by TKA in those eligible and willing to undergo surgery and revision TKA in those for whom primary TKA failed. Based on NSAID utilization in the IDEA cohort at baseline (unpublished observations), half of the patients were assumed to begin with the NSAID regimen. The other half used analgesics intermittently, without any long-term efficacy, until they were eligible for TKA (see Supplementary Appendix Section 2c, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract). Usual care does not include steroid injections, because although these injections are efficacious in the short-term, they have not been shown to have long-term efficacy in reducing knee OA–related pain (27).

Intervention duration. Patients could continue on the NSAID regimen for up to 20 years. As we describe below, discontinuation due to minor toxicity or treatment failure caused most patients to end treatment before this maximum was reached.

Efficacy. Data for NSAID efficacy were derived from the Glucosamine/chondroitin Arthritis Intervention Trial (28). TKA structural efficacy rates were derived from data in a study by Paxton and colleagues (29). TKA pain decrements were derived from 2 longitudinal studies in patients with knee OA undergoing TKA: the Adding

Table 2. Quality of life utilities (nonobese/obese)*

	WOMAC pain score (0–100 scale)				
Age group, years	0	1–15	16–40	41–70	71–100
0 comorbidities					
45–54	0.841/0.830	0.816/0.806	0.780/0.769	0.714/0.703	0.656/0.645
55–64	0.847/0.836	0.822/0.812	0.786/0.775	0.720/0.709	0.662/0.651
65–74	0.871/0.860	0.846/0.835	0.810/0.799	0.744/0.733	0.685/0.675
75+	0.854/0.843	0.829/0.818	0.793/0.782	0.727/0.716	0.669/0.658
1 comorbidity					
45–54	0.818/0.807	0.791/0.780	0.755/0.744	0.679/0.668	0.645/0.634
55–64	0.824/0.813	0.797/0.786	0.761/0.750	0.685/0.674	0.651/0.640
65–74	0.848/0.837	0.821/0.810	0.785/0.774	0.708/0.698	0.674/0.664
75+	0.831/0.820	0.804/0.793	0.768/0.757	0.692/0.681	0.658/0.647
2+ comorbidities					
45–54	0.806/0.795	0.794/0.783	0.732/0.721	0.635/0.624	0.500/0.489
55–64	0.812/0.801	0.800/0.789	0.738/0.727	0.641/0.630	0.506/0.495
65–74	0.836/0.825	0.824/0.813	0.762/0.751	0.665/0.654	0.530/0.519
75+	0.819/0.808	0.807/0.796	0.745/0.734	0.648/0.637	0.513/0.502

* Data were derived from references 20 and 21. WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

Value in Knee Arthroplasty (AViKA) study (30) and the Study of Total Knee Arthroplasty Responses (STARs) (31) (see Supplementary Appendix Section 2a and Supplementary Table 3, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

Discontinuation. A total of 11.28% of model patients receiving NSAIDs discontinue treatment in the first year due to minor toxicities. This probability was derived from multiple, large randomized controlled trials of NSAIDs in arthritis (32–36). We assume that all discontinuations due to minor toxicity occur in the first year. In all years of treatment, patients can discontinue NSAIDs due to treatment failure. Discontinuation attributable to treatment failure is considered to occur if patients return to within 9 points of their starting pain score (half of the average initial decrease in pain).

After TKA, patients only move to a subsequent treatment if they require revision surgery. Therefore, we assume that discontinuation due to minor toxicity is 0%. Treatment failure–associated discontinuation can only occur if a patient experiences a structural failure and if pain returns to within 21 points of his or her starting pain score (half of the average initial decrease in pain) (see Supplementary Appendix Section 2b, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

Cost. The cost of NSAID treatment was calculated using an average of the cost of treatment (from Red Book Online [37]) weighted by the utilization of treatments by OA patients (Medicare Current Beneficiary Survey 2009 [38]). The cost also included an annual office visit and laboratory tests. NSAIDs cost $841 in the first year and $810 in subsequent years. TKA costs, derived from Medicare Fee Schedules, were $17,976 (primary) and $24,985 (revision) and $109 each year after the year of surgery (39). Cost derivation methods for NSAIDs and TKA are described in a prior publication (15).

Diet and exercise. Intervention duration. Consistent with the IDEA trial duration (18 months), in the base case, patients could remain on the D+E program for up to 2 years. Sensitivity analyses were performed to evaluate implementing D+E for longer durations. Because data on long-term weight loss are limited, we conservatively assumed that benefits from D+E were not maintained after the program ended (see Supplementary Appendix Section 3a, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

Efficacy. Patients on the D+E regimen were assigned a probability of the percent BMI reduction, derived from the IDEA trial data. The BMI reduction was associated with a percent reduction from baseline in the WOMAC pain score (Tables 3 and 4).

Table 3. Diet and exercise efficacy*

Percent reduction in BMI	Probability of BMI reduction	Percent reduction in WOMAC pain score, mean ± SD	Probability of weight loss failure (subsequent years)
20–25	0.08	52.1 ± 40.3	0.04
15–20	0.11	42.6 ± 53.1	0.04
10–15	0.23	27.5 ± 51.6	0.04
5–10	0.29	27.5 ± 51.6	0.34
0	0.30	11.9 ± 44.6	NA

* Data were derived from the Intensive Diet and Exercise for Arthritis trial data sets. BMI = body mass index; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index; NA = not applicable.

Table 4. Diet and exercise failure and cost*

Parameter	Value
Probability of pain reduction failure (subsequent years)	
Given weight loss success	0.21
Given weight loss failure	0.57
Discontinuation	
Overall discontinuation, %	8 (12, 16)
Discontinuation due to treatment failure, %	0 (50, 100)
Duration of D+E	2 years (3, 5, 8 years, no limit)
Cost, dollars	
Personnel	
First year	328
Subsequent years	281
Meal replacements	
First year	455 (0)
Subsequent years	0
Gym membership, assumption	
First year	600 (0)
Subsequent years	600 (0)

* For overall discontinuation, in the base case, we assumed that the diet and exercise (D+E) program would run for 2 years; thus, overall discontinuation occurred only in the first year. In sensitivity analyses, we tested longer durations of the D+E program and in those instances, the probability of discontinuation is annual. Values in parentheses were assessed by one-way sensitivity analyses. Probability of pain reduction failure, overall discontinuation, and personnel costs were derived from reference 12. Meal replacement costs were derived from references 12 and 40.

Each subsequent year, a patient had a probability of either losing the BMI reduction (termed weight loss failure) or losing the pain reduction (pain failure). Pain and weight loss failures did not necessarily occur together, but weight loss failure increased the probability of pain failure. Probability of weight loss failure was stratified by the amount of BMI reduction. If a patient experienced weight loss or pain failure, they reverted to the weight or pain level that they would have experienced had they not been on the D+E regimen. Likewise, in the year following termination of D+E, all patients revert to the weight or pain levels that they would have experienced had they not received the D+E intervention (see Supplementary Appendix Section 3b, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

We validated OAPol-predicted weight loss from the D+E regimen by comparing mean BMI in the D+E arm of the IDEA trial at the 5-year follow-up (18% of patients reporting data) to OAPol-predicted BMI at 5 years. The BMI in the IDEA trial D+E cohort was 31.65 kg/m^2, and OAPol-predicted BMI for D+E patients was 32.94 kg/m^2.

Discontinuation. Patients had an overall probability of discontinuing the D+E regimen each year. This overall discontinuation was divided into general discontinuation (discontinuation unrelated to treatment efficacy) and treatment failure discontinuation (discontinuation due to failure to maintain weight loss). Because we do not have data on why patients discontinued the D+E program in the IDEA trial, in the base case, we assumed that the 8% discontinuation rate was based entirely on general discontinuation; patients who had lost weight had an equal probability of discontinuing as those who did not. In sensitivity analyses, we varied the percent of discontinuations due to treatment failure (see Supplementary Appendix Section 3c, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

Cost. The cost of the D+E regimen was estimated using 3 components: personnel ($328 in the first year, $281 in subsequent years), meal replacements ($455 in the first year), and gym membership ($600 each year). Personnel costs were derived using 2015–2016 interventionist salaries at Wake Forest University. Meal replacement costs were derived using the current retail cost of the meal replacements (40) and the average number of meal replacement containers used by study patients (unpublished trial data). Gym costs were an assumption. Fixed costs were not included (see Supplementary Appendix Section 3d, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

Process benefits. Exercise may confer additional increases in quality-of-life utility beyond those due to reductions in pain and weight. We refer to these increases as "process benefits," because they appear to be attributable to the process of exercising rather than outcomes (41). Based on published data, we estimated the process benefits from the exercise intervention by increasing a patient's annual utility by 0.026 QALYs (41). The base case analysis did not include process benefits; however, in a one-way sensitivity analysis, they were added to the first year or all years of exercise (see Supplementary Appendix Section 3e, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

Base case analysis. *Health care sector perspective*. In the base case, we used D+E characteristics derived from IDEA trial data, with added costs for a gym membership. We assumed that D+E regimen costs included personnel costs, meal replacements, and a hypothetical $600/year gym membership. Discontinuation was estimated at 8% annually (the IDEA trial had a 12% discontinuation rate over 18 months). Process benefits were not included.

Societal perspective. The societal perspective included all health and quality of life components from the health care sector with added costs accounting for lost productivity and caregiving. Productivity costs reflect work absenteeism among patients with OA in the US labor force (42). Knee OA–related pain costs $1,037 in lost productivity annually. The primary and revision TKA costs were $3,311 and $3,592, respectively, in lost productivity during the year of the surgery (15).

Based on a study by Gupta and colleagues in which it was reported that 52.1% of the indirect costs for patients with OA were related to caregiving (43), we assumed that productivity costs due to knee OA pain represented 47.9% of annual indirect costs. This resulted in a total annual indirect cost of $2,166. In the base case, only patients with a WOMAC pain score >40 incurred productivity costs related to caregiving and OA pain. We varied this threshold in the sensitivity analyses.

Sensitivity analyses. One-way sensitivity analyses varied key D+E parameters. D+E duration was assessed at 2 (base case), 3, 5, and 8 years as well as without limits. The cost of D+E varied only for personnel costs, to 50% more than the base case cost. Overall discontinuations were assessed at 8% (base case), 12%, and 16% annually; the percentages of discontinuations due to treatment failure were assessed at 0% (base case), 50%, and 100%. Process benefits were not included (base case) or received by patients during the first year or all years of D+E. The societal perspective included an additional analysis varying the WOMAC pain threshold for productivity and caregiving costs from 1 point (any pain), to 15, 40, or 70 points.

We used probabilistic sensitivity analyses (PSA) to determine the effect of the uncertainty regarding the D+E regimen parameters (see Supplementary Appendix Table 4, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract). We varied cost, overall discontinuation, the relative reduction in pain severity, and the probability of pain and BMI failure. The societal perspective analysis also varied costs of OA-related caregiving. In all iterations, the duration of D+E was 2 years, discontinuation due to treatment failure was 0%, and process benefits were not included.

Results are shown with acceptability curves, which show the percentage of PSA simulations (of 1,000) for which an intervention was cost-effective at different values of willingness to pay for additional QALYs.

RESULTS

Base case analysis. The D+E regimen led to a gain of 5.4 QALYs for every 100 patients (increasing per-person QALE from 8.909 QALYs to 8.963 QALYs). From the health care sector and societal perspectives, the D+E regimen raised the per-person costs by $1,845 and $1,624, respectively. D+E was cost-effective at all thresholds. D+E had an ICER of $34,100/QALY from the health care sector perspective and an ICER of $30,000/QALY from the societal perspective (Table 5).

The improvement in quality of life from D+E was due to decreases in BMI and the WOMAC pain score. During the first year of the D+E program, the average WOMAC pain score in the D+E cohort was 6.8 (of 100) points lower than that in the usual care cohort, and their average BMI was 2.8 kg/m^2 lower than that in the usual care cohort. During the second year, the average WOMAC pain score in the D+E cohort was 4.9 points lower and their BMI was 2.3 kg/m^2 lower compared to that in the usual care cohort. Due to the assumption that D+E benefits would not extend beyond 2 years, D+E delayed, but did not avert, total knee replacement.

Table 5. Cost-effectiveness of diet and exercise*

	QALE	Lifetime cost, dollars	ICER, dollars/QALY
Health care perspective			
Usual care	8.909	116,200	34,100
Diet and exercise	8.963	118,100	
Societal perspective			
Usual care	8.909	130,700	30,000
Diet and exercise	8.963	132,400	

* Quality-adjusted life expectancy (QALE) and lifetime cost are shown as per-person values and were discounted at 3% per year. ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life-years.

One-way sensitivity analyses from the health care sector perspective. D+E was cost-effective at a WTP threshold of $100,000/QALY in all variations (Figure 1). Because D+E was always cost-effective at a threshold of $100,000/QALY, we do not report specific results for the $200,000/QALY threshold. The ICER for D+E was greater than the $50,000/QALY threshold only when cost was increased to 150% of the base case (ICER = $52,100/QALY). When only personnel costs were included, the ICER was $6,200/QALY, and when personnel and meal replacement costs were included, the ICER was $12,900. The inclusion of process benefits in the first year of D+E lowered the ICER to $22,600/QALY, and including process benefits in all years lowered the ICER further to $15,300/QALY. Varying discontinuations changed the ICER minimally.

When compared to usual care, all D+E durations were cost-effective at a WTP threshold of $50,000/QALY. We also compared D+E programs of different durations incrementally. The 2-year program was dominated by the 3-year program. The 3-year, 5-year, 8-year, and indefinite D+E programs had ICERs of $32,800/QALY, $33,400/QALY, $42,100/QALY, and $79,200/QALY respectively.

One-way sensitivity analyses from the societal perspective. D+E was somewhat more cost-effective when analyzed from the societal perspective than from the health care sector perspective. The highest ICER, $48,700/QALY, occurred when cost was 150% of the base case. Varying the WOMAC pain score threshold for productivity and caregiving costs had a small effect on cost-effectiveness. ICERs ranged from $26,800/QALY (WOMAC pain score >15) to $32,700/QALY (WOMAC pain score >1) (see Supplementary Figure 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

Probabilistic sensitivity analyses. From the health care sector perspective, at WTP thresholds of $50,000/QALY and $100,000/QALY, the likelihood of D+E being cost-effective was

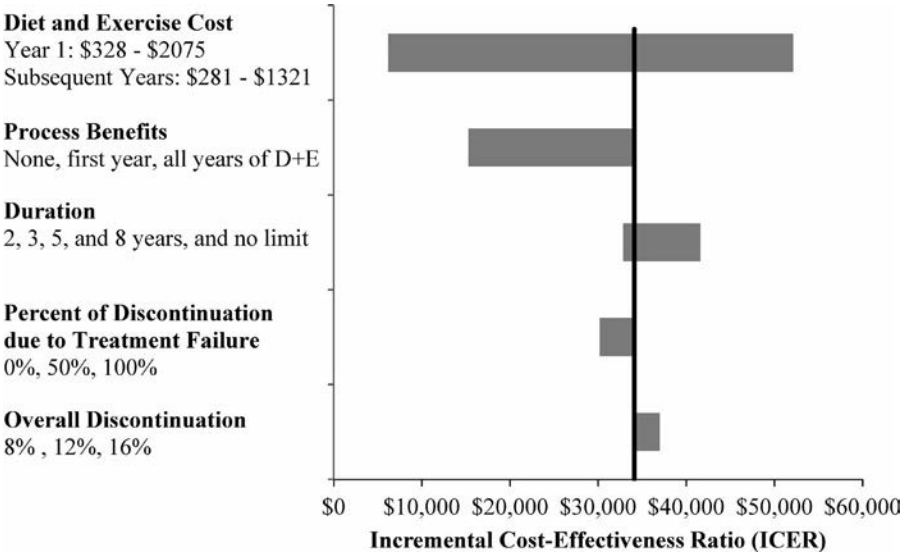


Figure 1. One-way sensitivity analysis of diet and exercise parameters (health care sector perspective). The figure illustrates the incremental cost-effectiveness ratio (ICER) estimated for the diet and exercise regimen under a variety of conditions. In each analysis, all parameters were held at base case values except for the parameters shown on the vertical axis, which varied according to the values listed. The leftmost end of each bar represents the ICER when the parameter of interest is set to its most favorable value; the rightmost end of each bar represents what happens when the parameter assumes its least favorable value. The vertical line shows the base case ICER. Process benefits are the increase in quality of life utility that occurs from the process of exercising (in addition to increases from weight loss and pain reduction). Overall discontinuation refers to the cohort's overall discontinuation rate, and the percentage of discontinuations due to treatment failure is the percentage of discontinuations that occur specifically in patients who did not maintain their weight loss.

58% and 100%, respectively (Figure 2). From the societal perspective, at WTP thresholds of $50,000/QALY and $100,000/QALY, the likelihood of D+E being cost-effective was 68% and 100%, respectively (see Supplementary Figure 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract).

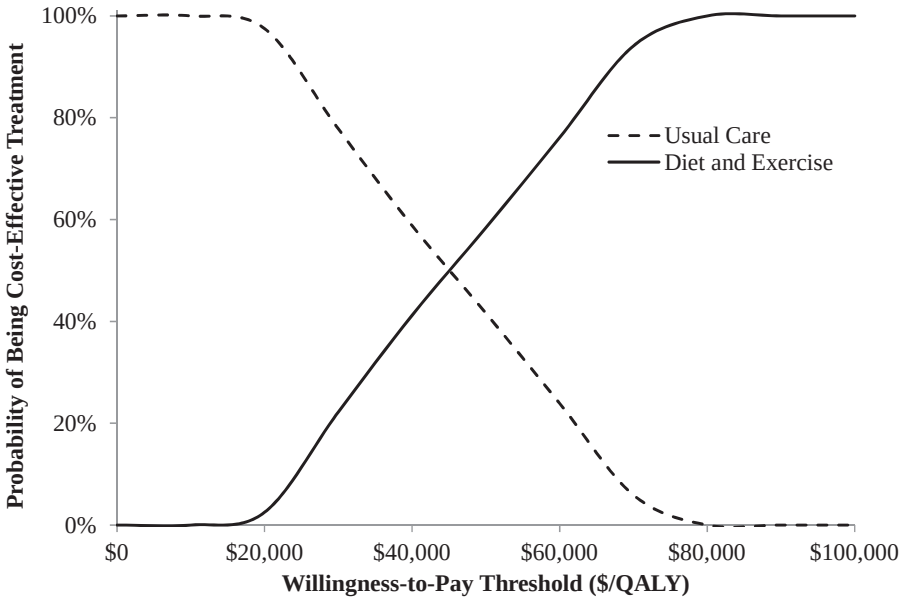


Figure 2. Cost-effectiveness acceptability curve (health care sector perspective). The curves show the percentage of simulations (of 1,000) for which an intervention was cost-effective at a given willingness-to-pay threshold. Each of the 1,000 analyses independently sampled model input parameters from the specified distribution (see Supplementary Table 4, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract). QALY = quality-adjusted life-years.

DISCUSSION

The results of our analysis suggest that incorporating a D+E regimen into usual care treatment for patients with knee OA would be highly cost-effective from both societal and health care sector perspectives. In our base case evaluation and the majority of sensitivity analyses, D+E had an ICER below $50,000/QALY, the more conservative cost-effectiveness threshold. D+E was always cost-effective with a threshold of $100,000/QALY, which is increasingly used in US cost-effectiveness analyses (18).

Both weight loss and D+E are effective for reducing knee OA pain and improving function (10–12,44). Cost-effectiveness analyses of exercise as treatment for knee OA (without weight loss as an explicit goal) have also generally indicated that exercise programs are cost-effective (45,46). A previous analysis of D+E in the ADAPT trial showed that D+E was cost-effective for self-reported function, pain, and stiffness (14). Our findings corroborate this and suggest that D+E would be cost-effective as a program with limited duration or as a program in which patients can continue to participate indefinitely. Our comparison of D+E durations suggests that with a WTP threshold of $50,000/QALY, the 8-year program provides the best value. If the WTP threshold is raised to $100,000/QALY, the program without a limit on duration offers the best value.

Of note, we showed that D+E is cost-effective when the effectiveness measure includes an adjustment for quality of life, which permits comparisons with other treatments. Although the 0.054 difference in QALE between the 2-year D+E program and usual care is small, this is because OA treatments primarily improve quality rather than quantity of life. The improvements in QALE from the 2-year D+E program are similar to those from over-the-counter (OTC) naproxen (0.081 QALYs) (16). The base case D+E ICER is also comparable to other OA treatments. OTC naproxen and TKA, for example, have ICERs of $57,100/QALY and $22,500/QALY, respectively (updated to 2016 US dollars) (7,16,47,48).

Given the number of Americans with OA, implementing a D+E program into usual care may lead to substantial improvement in quality of life on a population level, but funding the program may have a non-trivial effect on payers' budgets. The major components of the cost of D+E are meal replacements and a gym membership. Payers considering coverage for D+E programs could develop strategic partnerships with gyms and meal replacement manufacturers to minimize the budget impact.

We note several limitations. First, the D+E regimen was based on results from a clinical trial D+E regimen, which was led by professional interventionists and required significant patient investment. The outcomes of D+E regimens should also be established in community settings, where patients may not be as strongly motivated. In addition, the trial and model cohort had an average starting K/L grade of 2.5; therefore, the results may not apply to a cohort with more severe OA.

We made several assumptions to project the results of an 18-month clinical trial over a longer duration. Because data on long-term weight loss maintenance are limited, we assumed that patients lost all weight and pain reduction benefits once the D+E program ended, and that the base case D+E program would last for only 2 years. Long-term data on D+E adherence and the sustainability of weight loss and pain reduction are needed to more accurately model D+E treatments. We also did not consider potential correlations between baseline characteristics (e.g., age, pain) and D+E outcomes, which may have biased our point estimates.

Model inputs were derived from a variety of national data sources and published literature (see Supplementary Table 5, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr23716/abstract). Quality of life utilities were derived from the OAI using the SF-12, rather than directly measured in the IDEA cohort. The utility increase from the process of exercising (process benefit) was derived from the Health Survey for England (HSE) using the EuroQol 5-domain instrument. Because the populations and measures in the OAI and HSE differ, the process benefit quality of life values were included only in a one-way sensitivity analysis. Because our costs were trial-based, they may not entirely reflect the cost of implementing D+E outside of a clinical trial, although we added the gym membership cost to more accurately reflect what patients may have to contribute. In conformity with widely accepted guidelines for the conduct of economic evaluation (19,49), we used extensive sensitivity analyses to address uncertainty surrounding our findings. Our estimates were robust to uncertainty from the trial data.

Our findings strongly suggest that implementing D+E in the treatment of knee OA provides good value and should be a priority for clinicians and policy-makers. Further studies should consider how best to implement these programs and make them accessible to patients with knee OA.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Losina, Smith, Paltiel, Collins, Suter, Hunter, Katz, Messier.

Acquisition of data. Losina, Messier.

Analysis and interpretation of data. Losina, Smith, Paltiel, Collins, Suter, Hunter, Katz, Messier.

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. Losina E, Walensky RP, Reichmann WM, Holt HL, Gerlovin H, Solomon DH, et al. Impact of obesity and knee osteoarthritis on morbidity and mortality in older Americans. Ann Intern Med 2011;154:217–26.
3. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014;22: 363–88.
4. Jevsevar DS, Brown GA, Jones DL, Matzkin EG, Manner PA, Mooar P, et al. The American Academy of Orthopaedic Surgeons evidence-based guideline on: treatment of osteoarthritis of the knee, 2nd edition. J Bone Joint Surg Am 2013;95:1885–6.
5. Smith SR, Deshpande BR, Collins JE, Katz JN, Losina E. Comparative pain reduction of oral non-steroidal anti-inflammatory drugs and opioids for knee osteoarthritis: systematic analytic review. Osteoarthritis Cartilage 2016;24:962–72.
6. McCarberg B, Tenzer P. Complexities in the pharmacologic management of osteoarthritis pain. Curr Med Res Opin 2013;29:539–48.
7. Losina E, Walensky RP, Kessler CL, Emrani PS, Reichmann WM, Wright EA, et al. Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume. Arch Intern Med 2009;169:1113–21; discussion 21–2.
8. Skou ST, Roos EM, Laursen MB, Rathleff MS, Arendt-Nielsen L, Simonsen O, et al. A randomized, controlled trial of total knee replacement. N Engl J Med 2015;373:1597–606.
9. Wluka AE, Lombard CB, Cicuttini FM. Tackling obesity in knee osteoarthritis. Nat Rev Rheumatol 2013;9:225–35.
10. Christensen R, Bartels EM, Astrup A, Bliddal H. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: a systematic review and meta-analysis. Ann Rheum Dis 2007;66:433–9.
11. Messier SP, Loeser RF, Miller GD, Morgan TM, Rejeski WJ, Sevick MA, et al. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the Arthritis, Diet, and Activity Promotion Trial. Arthritis Rheum 2004;50:1501–10.
12. 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.
13. Flego A, Dowsey MM, Choong PF, Moodie M. Addressing obesity in the management of knee and hip osteoarthritis: weighing in from an economic perspective. BMC Musculoskelet Disord 2016;17:233.
14. Sevick MA, Miller GD, Loeser RF, Williamson JD, Messier SP. Cost-effectiveness of exercise and diet in overweight and obese adults with knee osteoarthritis. Med Sci Sports Exerc 2009;41:1167–74.
15. Losina E, Paltiel AD, Weinstein AM, Yelin E, Hunter DJ, Chen SP, et al. Lifetime medical costs of knee osteoarthritis management in the United States: impact of extending indications for total knee arthroplasty. Arthritis Care Res (Hoboken) 2015;67:203–15.
16. Katz JN, Smith SR, Collins JE, Solomon DH, Jordan JM, Hunter DJ, et al. Cost-effectiveness of nonsteroidal anti-inflammatory drugs and opioids in the treatment of knee osteoarthritis in older patients with multiple comorbidities. Osteoarthritis Cartilage 2016;24:409–18.
17. Woods B, Revill P, Sculpher M, Claxton K. Country-level cost-effectiveness thresholds: initial estimates and the need for further research. Value Health 2016;19:929–35.
18. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness: the curious resilience of the $50,000-per-QALY threshold. N Engl J Med 2014;371:796–7.
19. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG, editors. Cost-effectiveness in health and medicine. 2nd ed. New York: Oxford University Press; 2017.
20. National Institute of Health: National Institute of Mental Health Data Archive. Osteoarthritis Initiative (OAI). 2013. URL: https://data-archive.nimh.nih.gov/oai/.
21. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care 2004;42:851–9.
22. Lean ME, Han TS, Seidell JC. Impairment of health and quality of life using new US federal guidelines for the identification of obesity. Arch Intern Med 1999;159:837–43.
23. Muller-Nordhorn J, Muckelbauer R, Englert H, Grittner U, Berger H, Sonntag F, et al. Longitudinal association between body mass index and health-related quality of life. PLoS One 2014;9:e93071.
24. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
25. Centers for Disease Control and Prevention. U.S. Department of Health and Human Services. National Center for Health Statistics. National Health and Nutrition Examination Survey 2011-2014. URL: https://www.cdc.gov/nchs/nhanes/index.htm.
26. DeLemos BP, Xiang J, Benson C, Gana TJ, Pascual ML, Rosanna R, et al. Tramadol hydrochloride extended-release once-daily in the treatment of osteoarthritis of the knee and/or hip: a double-blind, randomized, dose-ranging trial. Am J Ther 2011;18:216–26.
27. McAlindon TE, LaValley MP, Harvey WF, Price LL, Driban JB, Zhang M, et al. Effect of intra-articular triamcinolone vs saline on knee cartilage volume and pain in patients with knee osteoarthritis: a randomized clinical trial. JAMA 2017;317:1967–75.
28. Clegg DO, Reda DJ, Harris CL, Klein MA, O'Dell JR, Hooper MM, et al. Glucosamine, chondroitin sulfate, and the two in combination for painful knee osteoarthritis. N Engl J Med 2006;354:795–808.
29. Paxton EW, Namba RS, Maletis GB, Khatod M, Yue EJ, Davies M, et al. A prospective study of 80,000 total joint and 5000 anterior cruciate ligament reconstruction procedures in a community-based registry in the United States. J Bone Joint Surg Am 2010;92 Suppl 2:117–32.
30. Losina E, Collins JE, Wright J, Daigle ME, Donnell-Fink LA, Strnad D, et al. Postoperative care navigation for total knee arthroplasty patients: a randomized controlled trial. Arthritis Care Res (Hoboken) 2016;68:1252–9.
31. Dave AJ, Selzer F, Losina E, Usiskin I, Collins JE, Lee YC, et al. The association of pre-operative body pain diagram scores with pain outcomes following total knee arthroplasty. Osteoarthritis Cartilage 2017;25:667–75.
32. Chan FK, Lanas A, Scheiman J, Berger MF, Nguyen H, Goldstein JL. Celecoxib versus omeprazole and diclofenac in patients with osteoarthritis and rheumatoid arthritis (CONDOR): a randomised trial. Lancet 2010;376:173–9.
33. Silverstein FE, Faich G, Goldstein JL, Simon LS, Pincus T, Whelton A, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. Celecoxib Long-term Arthritis Safety Study. JAMA 2000;284:1247–55.
34. Lisse JR, Perlman M, Johansson G, Shoemaker JR, Schechtman J, Skalky CS, et al. Gastrointestinal tolerability and effectiveness of rofecoxib versus naproxen in the treatment of osteoarthritis: a randomized, controlled trial. Ann Intern Med 2003;139:539–46.
35. Cannon CP, Curtis SP, FitzGerald GA, Krum H, Kaur A, Bolognese JA, et al. Cardiovascular outcomes with etoricoxib and diclofenac in patients with osteoarthritis and rheumatoid arthritis in the Multinational Etoricoxib and Diclofenac Arthritis Long-term (MEDAL) programme: a randomised comparison. Lancet 2006;368:1771–81.

36. Schnitzer TJ, Burmester GR, Mysler E, Hochberg MC, Doherty M, Ehrsam E, et al. Comparison of lumiracoxib with naproxen and ibuprofen in the Therapeutic Arthritis Research and Gastrointestinal Event Trial (TARGET), reduction in ulcer complications: randomised controlled trial. Lancet 2004;364:665–74.
37. Red Book Online. IBM Watson Health information. URL: https://www.micromedexsolutions.com.
38. Centers for Medicare and Medicaid Services. Medicare Current Beneficiary Survey (MCBS) 2009. URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/.
39. Centers for Medicare and Medicaid Services. Medicare Fee Schedules 2015. URL: https://www.cms.gov/medicare/medicare-fee-for-service-payment/feeschedulegeninfo/index.html.
40. GNC Total Lean Lean Shake. General Nutrition Centers. Accessed: March 12, 2018. URL: http://www.gnc.com/weight-management/meal-replacements/?landing=1.
41. Pavey TG, Anokye N, Taylor AH, Trueman P, Moxham T, Fox KR, et al. The clinical effectiveness and cost-effectiveness of exercise referral schemes: a systematic review and economic evaluation. Health Technol Assess 2011;15:i–xii, 1–254.
42. Berger A, Hartrick C, Edelsberg J, Sadosky A, Oster G. Direct and indirect economic costs among private-sector employees with osteoarthritis. J Occup Environ Med 2011;53:1228–35.
43. Gupta S, Hawker GA, Laporte A, Croxford R, Coyte PC. The economic burden of disabling hip and knee osteoarthritis (OA) from the perspective of individuals living with this condition. Rheumatology (Oxford) 2005;44:1531–7.
44. Bliddal H, Leeds AR, Stigsgaard L, Astrup A, Christensen R. Weight loss as treatment for knee osteoarthritis symptoms in obese patients: 1-year results from a randomised controlled trial. Ann Rheum Dis 2011;70:1798–803.
45. Richardson G, Hawkins N, McCarthy CJ, Mills PM, Pullen R, Roberts C, et al. Cost-effectiveness of a supplementary class-based exercise program in the treatment of knee osteoarthritis. Int J Technol Assess Health Care 2006;22:84–9.
46. Pinto D, Robertson MC, Hansen P, Abbott JH. Cost-effectiveness of nonpharmacologic, nonsurgical interventions for hip and/or knee osteoarthritis: systematic review. Value Health 2012;15:1–12.
47. Centers for Medicare and Medicaid Services. National health expenditure accounts: personal health care. URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html.
48. U.S. Bureau of Economic Analysis. Personal consumption expenditures: Services: Health Care (DHLCRC1Q027SBEA). URL: https://fred.stlouisfed.org/series/DHLCRC1Q027SBEA.
49. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making 2012;32:722–32.
50. Consumer Price Index for All Urban Consumers (CPI-U). National Bureau of Labor Statistics. 2016. URL: https://www.bls.gov/cpi/.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 865–874
DOI 10.1002/acr.23725

American College of Rheumatology
Empowering Rheumatology Professionals

Knee Osteoarthritis and the Risk of Medically Treated Injurious Falls Among Older Adults: A Community-Based US Cohort Study

Kamil E. Barbour,[1] Naoko Sagawa,[2] Robert M. Boudreau,[2] Mary E. Winger,[2] Jane A. Cauley,[2] Michael C. Nevitt,[3] Tomoko Fujii,[4] Kushang V. Patel,[5] and Elsa S. Strotmeyer[2]

Objective. The risk of falls among adults with knee osteoarthritis (OA) has been documented, yet, to our knowledge no studies have examined knee OA and the risk of medically treated injurious falls (overall and by sex), which is an outcome of substantial clinical and public health relevance.

Methods. Using data from the Health Aging and Body Composition Knee Osteoarthritis Substudy, a community-based study of white and African American older adults, we tested associations between knee OA status and the risk of injurious falls among 734 participants with a mean ± SD age of 74.7 ± 2.9 years. Knee radiographic OA (ROA) was defined as having a Kellgren-Lawrence grade of ≥2 in at least 1 knee. Knee symptomatic ROA (sROA) was defined as having both ROA and pain symptoms in the same knee. Injurious falls were defined using a validated diagnosis code algorithm from linked Medicare fee-for-service claims. Cox regression modeling was used to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs).

Results. The mean ± SD follow-up time was 6.59 ± 3.12 years. Of the 734 participants, 255 (34.7%) had an incident injurious fall over the entire study period. In the multivariate model, compared with those without ROA or pain, individuals with sROA (HR 1.09 [95% CI 0.73–1.65]) did not have a significantly increased risk of injurious falls. Compared with men without ROA or pain, men with sROA (HR 2.57 [95% CI 1.12–5.91]) had a significantly higher risk of injurious falls. No associations were found for women or by injurious fall type.

Conclusion. Knee sROA was independently associated with an increased risk of injurious falls in older men, but not in older women.

INTRODUCTION

Knee osteoarthritis (OA) is a common and disabling chronic condition among older adults (ages ≥65 years) (1). In a previous study in the US, the prevalence of knee radiographic OA (ROA) and symptomatic ROA (sROA) among adults ages ≥60 years was shown to be 37.4% and 12.1%, respectively; thus, about 1 in 3 adults with knee OA have reported pain (2). The prevalence of knee ROA was significantly higher in women versus men (42.1% versus 31.2% but knee sROA prevalence did not differ by sex (2). Furthermore, the lifetime risk (to age 85 years) of knee sROA was estimated to be about 45% and did not vary by sex (3). Knee OA may lead to reduced quality of life (4) and early retirement (5). Major health outcomes associated with knee OA include expensive joint replacement (620,000 OA-attributable knee replacements in the US in 2010) (6), and possibly an increased risk of mortality (7,8).

Falls are the leading cause of injury-related morbidity and mortality among older adults, with more than 1 in 4 older adults

Presented in part at the American College of Rheumatology Annual Scientific Meeting, Washington, DC, 2016 and the American Society for Bone Research Conference, 2016.

The opinions and assertions contained herein are those of the authors and do not necessarily represent those of the CDC.

Supported by the National Institute on Aging (contracts N01-AG-6-2101, N01-AG-6-2103, N01-AG-6-2106), the National Institute of Nursing Research (grant R01-NR12459), and in part by the NIH (Intramural Research Program). Dr. Strotmeyer's work was supported by the University of Pittsburgh Claude D. Pepper Older Americans Independence Center (pilot grant P30-AG024827) and the National Institute on Aging (grant R01-AG028050). Ms Winger's work was supported by the National Institute on Aging (training grant T32-AG000181).

[1]Kamil E. Barbour, PhD: Centers for Disease Control and Prevention, Atlanta, Georgia, and United States Public Health Service, Commissioned Corps, Rockville, Maryland; [2]Naoko Sagawa, MD, MPH, Robert M. Boudreau, PhD, Mary E. Winger, MPH, Jane A. Cauley, DrPH, Elsa S. Strotmeyer, PhD, MPH: University of Pittsburgh, Pittsburgh, Pennsylvania; [3]Michael C. Nevitt, PhD: University of California, San Francisco; [4]Tomoko Fujii, MD, PhD: University of Pittsburgh, Pittsburgh, Pennsylvania, and The University of Tokyo Hospital, Tokyo, Japan; [5]Kushang V. Patel, PhD: University of Washington, Seattle.

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

Address correspondence to Kamil E. Barbour, PhD, Centers for Disease Control and Prevention, 4770 Buford Highway, MS# F78, Atlanta, GA, 30319. E-mail: iyk1@cdc.gov.

Submitted for publication February 14, 2018; accepted in revised form August 14, 2018.

SIGNIFICANCE & INNOVATIONS

- To our knowledge, no studies have examined the association between knee osteoarthritis (OA) and the risk of medically treated injurious falls, which is an outcome of substantial clinical and public health relevance.
- Compared with men without radiographic OA or pain, men with symptomatic radiographic OA (hazard ratio 2.57 [95% confidence interval 1.12–5.91]) had a significantly higher risk of injurious falls. No association was found for women.
- Nonpharmacologic therapies (e.g., physical activity interventions) may help reduce the risk of falls in older adults, particularly older men with knee OA, by improving physical function. For example, EnhanceFitness, an evidence-based community-delivered physical activity program that is recommended by the Centers for Disease Control for adults with arthritis and disseminated by many YMCA recreational facilities across the US, has been shown to produce substantial improvements in function (e.g., muscle strength and balance) and may reduce the risk of a medically treated injurious fall.

falling each year (9), resulting in direct medical costs of approximately $32 billion in 2015 (10). Major fall-related injuries among older adults, including hip fractures and brain injuries, are associated with a decline in functional abilities and reductions in social and physical activities (11).

In a systematic review of 12 studies, 17% of the falls were attributed to gait/balance disorders or weakness as the most likely cause, which are common characteristics of adults with knee OA (11). A fairly recent study among adults with knee OA showed that lower knee extension muscle strength and lower knee flexion muscle strength were associated with increased falls (12). Moreover, poor or declining physical function is a risk factor for fractures, 95% of which occur because of a fall (13–16). Multiple cross-sectional studies have examined the association between knee OA and falls, yielding primarily null associations (17–19). Several studies have examined the association between baseline knee OA and risk of incident falls, with some indicating an increase in the risk of falls (20–22) and others finding a null result (23,24). Knee OA severity appears to impact the risk of falls, with greater severity linked to a higher rate of falls (25). Other knee OA–related outcomes, such as knee arthroplasty and knee instability, have been examined in regard to falls. Interestingly, adults with knee arthroplasty do not appear to have an increased risk of falls compared with adults without knee arthroplasty (26,27), whereas knee instability has been shown to be associated with a greater prevalence of recurrent falls (28). Yet, injurious falls are an outcome of greater clinical and public health relevance (29). Two cross-sectional studies found a higher prevalence of injurious falls in adults with doctor-diagnosed arthritis versus those without doctor-diagnosed arthritis (2 times greater) (30) and in adults with lower limb arthritis versus those without lower limb arthritis (about 1.3 times greater) (31). A longitudinal study of community-dwelling older adults showed a 40% increased risk of self-reported injurious falls among adults with arthritis or rheumatism (32). However, to our knowledge, the association between knee OA and the risk of incident injurious falls has not been examined. If the risk of injurious falls is higher among older adults with knee OA, targeted therapy/programs that would modify function and/or pain for those with elevated risk would be beneficial in order to reduce morbidity and mortality from falls.

To address this substantial knowledge gap, we examined the association between knee OA (for both ROA and sROA) and treated incident injurious falls among community-dwelling white and African American older men and women, from the Health Aging and Body Composition (ABC) Knee Osteoarthritis substudy.

Furthermore, we also performed a priori secondary analyses, stratified by sex and by type of injurious fall (fracture versus nonfracture). Because of sex differences in pain threshold, when reporting pain, the impact of pain on behavior, and overall risk behavior, it was important to examine these associations separately in men and women. Moreover, it is important to stratify by sex because there is evidence that sex hormones influence the development of OA and osteoporosis/fracture risk, practically via the reduction of estrogen levels in postmenopausal women (33).

PATIENTS AND METHODS

Study population. The parent Health ABC study enrolled 3,075 women and men, ages 70–79, from 2 field centers, Pittsburgh, Pennsylvania and Memphis, Tennessee at visit 1 (baseline, 1997–1998). Participants had to report no difficulty walking at least one-fourth of a mile and/or climbing a flight of stairs to be eligible to participate. White participants were identified from a random sample of white Medicare beneficiaries. African American participants were identified as age-eligible community residents from designated zip code areas surrounding Pittsburgh, Pennsylvania and Memphis, Tennessee. Exclusion criteria included reported difficulty performing basic activities of daily living, obvious cognitive impairment, inability to communicate with the interviewer, intention of moving within 3 years, or participation in a trial involving a lifestyle intervention. There were 3,044 enrollees that remained at visit 2 (1998–1999).

The Health ABC Knee Osteoarthritis substudy included 1,123 participants from visits 2 (1998–1999) or 3 (1999–2000) (Figure 1). Participants were included in the substudy at visit 2 if they had qualifying knee pain and knee radiograph. Cases with qualifying knee pain were identified if they had "knee pain, aching, or stiffness on most days for at least 1 month" at some point over the previous year or if they reported moderate or worse knee pain during the previous month in association with at least 1 activity on the Western Ontario and McMaster

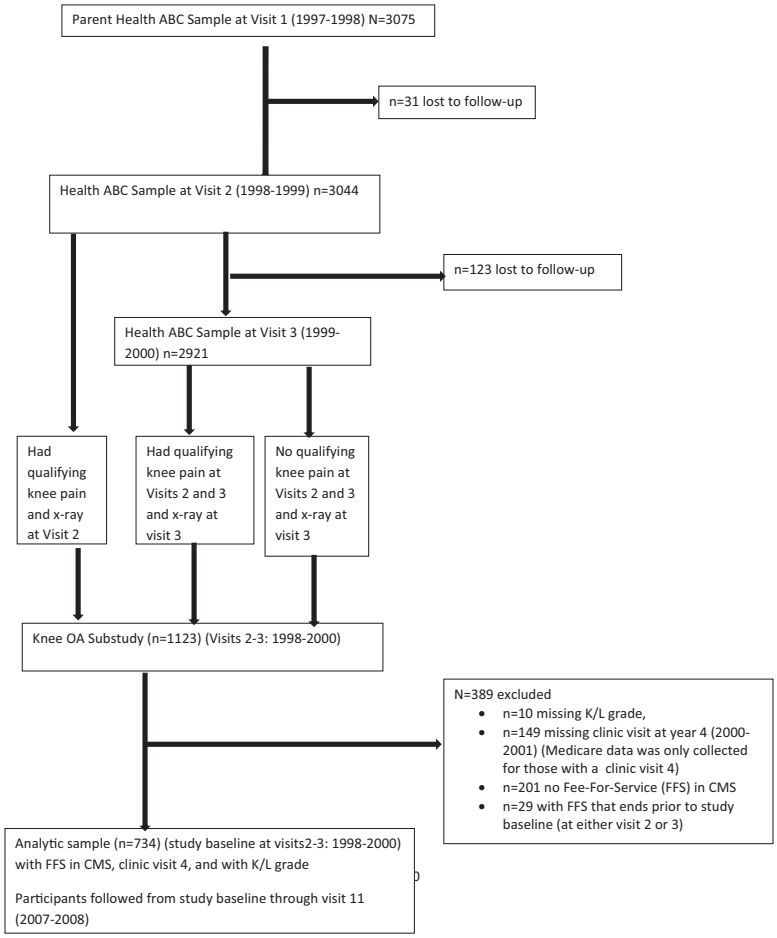


Figure 1. Flow chart for creating the analytic sample. OA = osteoarthritis; K/L = Kellgren/Lawrence; CMS = Centers for Medicare and Medicaid Services.

Universities Arthritis Index knee pain scale (34). Additional participants with qualifying pain were added at visit 3 if they had qualifying knee pain at both visits 2 and 3, and knee radiograph at visit 3. Finally, a random set of 270 controls (with no qualifying pain at either visit 2 or 3) with knee radiographs at visit 3 were selected from 1,798 participants for inclusion in the substudy (35).

Participants included in the analytic sample (n = 734) were those that were followed after study baseline (i.e., visits 2 or 3 in 1998–2000) with a clinic visit 4 (Medicare data was only collected for those with a clinic visit 4 [2000–2001]), who did not have a missing Kellgren-Lawrence (K/L) grade reading of their knee radiograph and had Medicare fee-for-service (FFS) in the Centers for Medicare and Medicaid Services (to ascertain status of injurious falls) that extended beyond study baseline enrollment date (1998–2000) (Figure 1). Participants in the analytic study were followed from study baseline (1998–2000) until the occurrence of an injurious fall, the loss of FFS, loss to follow-up, death, or through visit 11 (2007–2008) (when Medicare claims were last collected for this study). Injurious fall status was assessed during this time period of FFS in the Centers for Medicare and Medicaid Services data set. The institutional review board at each center approved the study protocol, and written informed consent was obtained from all the participants.

Exposure variable of knee OA and pain. At both visits 2 and 3, expert readers assessed posteroanterior and skyline projection knee Radiolov to assess K/L grade based on individual radiographic features (joint space narrowing, osteophytes, subchondral attrition, cysts, and sclerosis) and scored using the Osteoarthritis Research Society International atlas in the medial and lateral compartment of the tibiofemoral joint and the patellofemoral joint (36). Participants had radiographs taken at either visits 2 or 3. Follow-up began at either visit 2 or 3, depending on when the radiographs were taken. Interrater reliability was excellent (weighted kappa 0.87 for K/L grade). Knee ROA was defined as a K/L grade ≥2.

Adults were categorized into 4 mutually exclusive groups, including knee sROA, knee ROA without pain, knee pain without ROA, and no ROA or pain. During the clinic visit when knee imaging was completed, participants were asked if they had "knee pain on most days in the past 30 days." Knee pain was defined as

having pain symptoms (during the majority of the last 30 days) in at least 1 knee. Knee ROA was defined as having a K/L grade of ≥2 in at least 1 knee. Knee sROA was defined as having both ROA and pain symptoms in the same knee.

Injurious falls. Incident injurious falls were ascertained from outpatient and inpatient Medicare claims and defined using a diagnosis code algorithm from linked Medicare claims. All injuries captured in Medicare claims are included, which includes any outpatient care billed by any type of provider. Any unique event with an International Classification of Diseases, Ninth Revision fall code (E880–888) plus nonfracture injury, a vertebral fracture code (805–806) with a fall code, or any nonvertebral fracture code (800–804, 807–829) with/without a fall code was considered an injurious fall, because an estimated 80% of nonvertebral fractures are attributed to falls (37). All traumatic (e.g., motor vehicle accidents), intentional, and pathologic injuries were not considered to meet the definition of an injurious fall.

In the adjudication of our diagnoses code algorithm as previously detailed (38), a subset of Medicare fall injuries were compared to the self-reported fall injuries with medical records. The injuries adjudicated were included due to potential uncertainty of particular diagnoses codes to classify a primary injurious fall; these included concurrent stroke code, fall code with uncertain injury, fracture code that was not in the first or second positions (i.e., not listed as the first or second reason for visit/hospitalization billing), and vertebral column/rib fracture without a fall code. Overall, the injurious fall adjudication showed an excellent agreement, except for vertebral fractures, where only 50% of vertebral fractures were confirmed. We revised the initial diagnoses code algorithm to exclude vertebral fractures without concurrent fall codes.

Covariates evaluated for inclusion in models. All covariates evaluated for inclusion in models were measured once at either visit 1 or 2. Potential covariates associated with the exposure or outcome at $P < 0.1$ were included in the full multivariate adjusted model. If covariates were available at both visits, the visit 1 measurement was used. These potential covariates were selected based on documented associations with knee OA and falls. Demographic variables included self-report of age, sex, race (white or African American), and education (less than high school [HS], HS graduate, or postsecondary), and study site (Memphis or Pittsburgh). Weight was measured on a standard balance beam scale to the nearest 0.1 kg. Height was measured by a stadiometer to the nearest 0.1 cm. The anthropometric measure of body mass index (BMI) was calculated using the formula weight (kg)/height (m^2).

Lifestyle factors included self-report of smoking (never, past, or current smoker) and physical activity (kcal/kg/week). Physical activity was determined using the caloric expenditure in the past week for self-reported duration of walking, climbing stairs, and exercise (39).

Several medical characteristics were considered for the analysis. Participants self-reported their current health status (fair/poor/very poor versus good/very good/excellent) and history of falls in the past 12 months, depression, poor vision, myocardial infarction, and stroke. Diabetes was defined using fasting glucose (≥126 mg/dl), self-report, or hypoglycemic medication use. Diagnosed and/or treated hypertension was defined via self-report or antihypertensive medication use. To assess supplementary intake for vitamin D and calcium, and medication use such as nonsteroidal antiinflammatory drugs (NSAIDs), statins, steroids, and antidepressants, participants were asked to bring all prescription and over-the-counter medications, which were coded based on the Iowa Drug Information System (40). The total number of other medications was assessed using the number of other prescription medications (excluding steroids and antidepressants).

Statistical analysis. Chi-square tests were used to evaluate proportion differences for incident injurious falls across study covariates. Two-sample *t*-tests (Wilcoxon rank sum test for non-normally distributed data) were used to examine mean differences in continuous covariates by incident injurious falls status. In order to compare baseline knee OA status with study covariates, chi-square tests were performed to assess proportion differences. Fisher's exact test was performed for all tests of proportion if the expected value for any cell was <5. In order to compare mean differences in continuous covariates by baseline knee OA status, analysis of variance (Kruskal-Wallis test for non-normally distributed data) was performed.

Cox proportional hazards models were used to estimate hazard ratios (HRs) with 95% confidence intervals (95% CIs) and compare the time from the visits 2 or 3 (depending on when radiograph was taken) to incident injurious falls by baseline knee OA status (knee sROA, knee pain without ROA, knee ROA without pain groups compared with reference group [no ROA or pain]), while controlling for potential confounders. Participants were right-censored if they did not have the event of interest by the time they were lost to follow-up, their follow-up ended, or by the time of their death. Individuals with missing covariate data were dropped from the multivariate analyses. Furthermore, we performed a priori secondary analyses, stratified by sex and injurious falls type (fracture versus nonfracture). The proportional hazards assumption was assessed by first testing the interaction of knee OA indicator variables (sROA, ROA without pain, or pain without ROA, versus no ROA/no pain) with log(time) and then the knee OA indicator variables using the Supremum test based on Schoenfeld residuals, in full multivariate adjusted models. The proportional hazards assumption was considered violated if $P < 0.05$ for the knee OA indicator variables *log(time) interaction terms or knee OA indicator variables. Knee OA did not violate the proportional hazards assumption in the full multivariate adjusted overall, sex-stratified, and injury type–stratified models. All analyses were performed

Table 1. Distributions of baseline knee OA status and incident injurious falls by baseline characteristics (n = 734)*

	Knee sROA (n = 249)†	Knee ROA without pain (n = 32)‡	Pain without knee ROA (n = 306)	No knee pain or ROA (n = 147)	*P* by knee OA status	Injurious fall (n = 255)	*P* by injurious falls status
Overall	33.9	4.4	41.7	20.0	–	34.7	–
Demographic characteristics							
Sex					0.32		<0.01
Men	31.3	4.2	41.3	23.3		26.7	
Women	35.7	4.5	41.9	17.9		39.9	
Race					<0.01		<0.01
White	29.3	2.1	47.4	21.2		42.7	
African American	40.3	7.4	33.9	18.4		23.9	
Site					0.01		0.10
Pittsburgh	40.3	3.1	36.8	19.8		38.1	
Memphis	29.1	5.3	45.4	20.2		32.2	
Education					0.28		<0.01
<High school	37.4	6.6	37.4	18.5		26.1	
High school graduate	29.9	3.6	45.7	20.8		32.6	
Postsecondary	34.8	3.3	41.5	20.4		42.8	
Lifestyle characteristics							
Smoking					0.19		0.04
Never	37.2	4.3	40.1	18.5		37.2	
Past smoker	33.0	4.4	43.2	19.4		35.0	
Current smoker	20.6	4.8	44.4	30.2		20.6	
Health status					0.04		0.04
Fair/poor/very poor	41.9	5.7	40.3	12.1		26.6	
Good/very good/excellent	32.3	4.1	42.0	21.6		36.4	
Medical characteristics							
History of falls last 12 months					0.02		0.22
Yes	31.9	2.7	50.8	14.6		38.4	
No	34.6	4.8	38.8	21.9		33.5	
Diabetes					0.59		0.08
Yes	33.3	5.3	44.4	17.0		29.2	
No	34.1	4.1	40.9	21.0		36.4	
Hypertension					0.02		0.45
Yes	38.1	4.8	40.4	16.8		33.5	
No	29.1	3.8	43.2	23.8		36.2	
Stroke					0.20		0.89
Yes	25.0	8.3	33.3	33.3		33.3	
No	34.2	4.2	42.1	19.5		34.8	
Myocardial infarction					0.42		0.02
Yes	32.4	7.4	36.8	23.5		22.1	
No	34.3	4.1	42.2	19.5		36.1	
Depression					0.61		0.34
Yes	39.1	2.2	45.7	13.0		41.3	
No	33.5	4.6	41.5	20.4		34.4	
Poor vision					0.33		0.22
Yes	35.2	3.6	42.9	18.3		36.6	
No	32.3	5.3	40.1	22.3		32.3	

(continued)

Table 1. *(Cont'd)*

	Knee sROA (n = 249)†	Knee ROA without pain (n = 32)‡	Pain without knee ROA (n = 306)	No knee pain or ROA (n = 147)	*P* by knee OA status	Injurious fall (n = 255)	*P* by injurious falls status
Calcium supplement use					0.38		<0.01
Yes	29.7	3.5	47.1	19.8		45.9	
No	35.3	4.5	40.1	20.1		31.2	
Vitamin D supplement use					0.81		0.02
Yes	35.6	2.3	40.2	21.8		46.0	
No	33.8	4.5	42.0	19.8		33.1	
Antidepressant use					0.50		0.03
Yes	26.1	4.4	56.5	13.0		56.5	
No	34.2	4.2	41.3	20.2		34.0	
Statin use					0.84		0.25
Yes	30.6	5.1	43.9	20.4		39.8	
No	34.5	4.1	41.5	19.9		33.9	
NSAID use					<0.01		0.09
Yes	46.4	4.1	43.8	5.7		39.7	
No	29.5	4.3	41.0	25.2		32.8	
Steroid use					0.04		0.01
Yes	23.8	4.8	61.9	9.5		54.8	
No	34.5	4.3	40.5	20.7		33.5	
Knee OA status					–		0.15
sROA	–	–	–	–		35.3	
ROA	–	–	–	–		28.1	
Pain without ROA	–	–	–	–		38.2	
No pain or ROA	–	–	–	–		27.9	

* Values are the percentage of participants unless indicated otherwise. Data from the Health Aging and Body Composition (Health ABC) study, a US cohort study of 3,075 women and men, ages 70–79 years. OA = osteoarthritis; sROA = symptomatic radiographic osteoarthritis; NSAID = nonsteroidal antiinflammatory drug.
† Kellgren-Lawrence grade ≥2 + symptoms.
‡ Kellgren-Lawrence grade ≥2.

using the SAS, version 9.3. Statistical significance for all analyses was determined at the $\alpha < 0.05$ level.

RESULTS

Overall. The mean ± SD follow-up time was 6.59 ± 3.12 years. The mean ± SD age of the participants was 74.7 ± 2.9 years. There were 249 participants with sROA, 32 had ROA without pain, 306 had pain without ROA, and 147 did not have pain or ROA in either knee. For the entire study period, 255 of the 734 participants (34.7%) had an incident injurious fall. The average annual incidence rate of injurious falls across 11 years (from 1998 through 2008) was 4.84 per 100 person-years. The incidence rate per 100 person years of injurious falls, by year, was 0.7 in 1998, 3.0 in 1999, 4.0 in 2000, 5.8 in 2001, 5.2 in 2002, 3.7 in 2003, 4.8 in 2004, 5.4 in 2005, 3.2 in 2006, 6.8 in 2007, and 10.3 in 2008. The distributions of baseline knee OA status and incident injurious falls by baseline characteristics are shown in Table 1. Baseline knee OA status differed significantly by race, site, health status, history of falls, hypertension status, NSAID use, or steroid use (Table 1). Women, white subjects, those who had higher education, never smoked, better health, or no heart disease, or those who took calcium or vitamin D supplements, antidepressants, or steroids were significantly more likely to have an incident injurious fall by the end of follow-up (Table 1). Mean age and BMI varied significantly by knee OA status ($P < 0.01$ for both). The baseline mean age was higher in adults with an incident injurious fall by the end of follow-up compared to those without an injurious fall by the end of follow-up (75.1 versus 74.5 years, $P = 0.01$). The baseline mean BMI was lower in those with an incident injurious fall by the end of follow-up versus those without such a fall by the end of follow-up (27.3 versus 28.3 kg/m^2; $P = 0.01$). Median physical activity did not significantly differ by knee ROA status ($P = 0.35$). Baseline median physical activity was 2.4 kcal/kg/week for those with an incident injurious fall and 2.2 kcal/kg/week for those without an incident injurious fall by the end of follow-up, $P = 0.09$.

Number of other prescription medications (excluding steroids and antidepressants) differed by knee OA status (knee sROA = 3.64, knee ROA without pain = 2.81, pain with knee ROA = 3.74, no knee pain or ROA = 2.51; all $P < 0.01$) and was borderline significantly higher among adults with an injurious fall versus those without an injurious fall (3.61 versus 3.18; $P = 0.05$).

Men. There were a total of 77 injurious falls out of 288 men (26.7%), and the cumulative incidence varied by knee OA status. The cumulative incidence of injurious falls by knee OA group was 33.3% for sROA, 16.7% for ROA without pain, 29.4% for pain without ROA, and 14.9% for no ROA or pain.

Women. There were 178 injurious falls out of 446 participants (39.9%), and the cumulative incidence did not vary by knee OA status. The cumulative incidence of injurious falls by knee OA group was 36.5% for sROA, 35.0% for ROA without pain, 43.9% for pain without ROA, and 38.8% for no ROA or pain.

Multivariate analyses. In the multivariate model with men and women combined, compared with those without ROA or pain, individuals with sROA (HR 1.09 [95% CI 0.73–1.65]), ROA without pain (HR 1.01 [95% CI 0.46–2.20]), and pain without ROA (HR 1.08 [95% CI 0.74–1.57]) did not have a significantly increased risk of injurious falls (Table 2). Among men only, and compared with men without ROA or pain, those with sROA (HR 2.57 [95% CI 1.12–5.91]) had a significantly higher risk of injurious falls (Table 2). No significant association existed between knee OA and injurious falls in women. The 4*2 interaction term predicting injurious falls between knee OA and sex was not statistically significant ($P > 0.05$); however, interpretation of this result should be viewed with caution as this analysis was likely underpowered as a result of the low number of adults with ROA and no pain. The association between knee OA and injurious falls did not differ by injurious fall type (fracture versus nonfracture) (Table 3).

DISCUSSION

To our knowledge, this is the first study to examine knee OA and the risk of treated incident injurious falls. Our study demonstrated that knee sROA was associated with an increased risk of injurious falls among older community-dwelling men, independent of many potential confounders. Knee OA was not a predictor of injurious falls overall, among women, or by injurious fall type (fracture versus nonfracture). Our findings suggest that knee sROA is a risk factor for injurious falls in men, but not in women. Fall prevention efforts that target men with knee sROA are needed in order to reduce injurious falls risk.

Men with sROA had a 2.6-fold increased hazard rate of injurious falls compared with men who had no pain or ROA in either knee. In contrast, there was no significant association in women. It is unclear why this association was observed in men only. One potential explanation for this observation is that men are more likely to fall than women under similar conditions of health (e.g., OA) and balance (41). This may be the result of several factors, for instance, women who report pain are more likely to limit their activity than men (42), and men with similar conditions of health are more likely to put themselves in hazardous situations than women (41). Additionally, women may have a lower pain threshold and men are more likely to only report severe pain (43). A recent longitudinal 3-year study using data from the Swedish National Study on Aging

Table 2. Adjusted risk of injurious falls overall and by sex*

	No. w/ injurious falls	Knee sROA HR (95% CI)†	Knee ROA w/out pain HR (95% CI)†	Knee pain w/out ROA HR (95% CI)†
Overall				
Age-adjusted (n = 734)	255	1.05 (0.72–1.52)	0.91 (0.44–1.87)	1.15 (0.80–1.64)
Full MV model (n = 714)	249	1.09 (0.73–1.64)	1.11 (0.51–2.44)	1.05 (0.72–1.54)
Men				
Age-adjusted (n = 288)	77	1.86 (0.91–3.83)	0.93 (0.20–4.25)	1.73 (0.85–3.50)
Full MV model (n = 278)	75	2.59 (1.13–5.98)‡	1.19 (0.25–5.68)	2.03 (0.94–4.37)
Women				
Age-adjusted (n = 446)	178	0.76 (0.49–1.18)	0.83 (0.36–1.88)	0.91 (0.60–1.38)
Full MV model (n = 436)	174	0.88 (0.54–1.43)	1.13 (0.44–2.87)	0.86 (0.56–1.34)

* Values were adjusted for age, sex, race, education, body mass index, physical activity, smoking, health status, history of falls in past 12 months, diabetes, hypertension, myocardial infarction, use of steroids, nonsteroidal antiinflammatory drugs, antidepressants, calcium supplements, vitamin D supplements, and total number of other prescription medications (excluding steroid and antidepressant use). sROA = symptomatic radiographic osteoarthritis; HR = hazard ratio; 95% CI = 95% confidence interval; MV = multivariate model.

† Reference group comprises participants without ROA or pain in a knee.

‡ Significant.

Table 3. Adjusted risk of fracture and nonfracture injurious falls overall*

Injurious falls	No. with injurious falls	Knee sROA HR (95% CI)†	Knee ROA without pain HR (95% CI)†	Knee pain without ROA HR (95% CI)†
Fracture				
Age-adjusted (n = 667)	188	1.08 (0.70–1.69)	1.19 (0.56–2.51)	1.18 (0.78–1.81)
Full MV model (n = 649)	184	1.18 (0.73–1.90)	1.70 (0.75–3.83)	1.06 (0.68–1.66)
Nonfracture‡				
Age-adjusted (n = 546)	67	0.99 (0.49–2.01)	–	1.18 (0.61–2.30)
Full MV model (n = 530)	65	1.08 (0.49–2.38)	–	1.22 (0.59–2.52)

* Values were adjusted for age, race, sex, education, body mass index, physical activity, smoking, health status, history of falls in past 12 months, diabetes, hypertension, myocardial infarction, use of steroids, nonsteroidal antiinflammatory drugs, antidepressants, calcium supplements, vitamin D supplements, and total number of other prescription medications (excluding steroid and antidepressant use). See Table 2 for definitions.
† The reference group comprises participants without ROA or pain in a knee.
‡ No nonfracture injuries occurred among participants with knee ROA without pain.

and Care in Kungsholmen showed that among men with the presence of pain or pain that limited their daily activities, there was an increased risk of injurious falls, but for women these associations were null (44). To our knowledge, previous longitudinal studies have not examined knee OA using objectively measured radiographic data and related these to injurious falls treated in Medicare.

However, a few studies have examined radiographically measured knee OA and incident self-reported falls by sex, or in 1 sex (21,23,24). One study among adults who were followed <3 years showed that knee OA was a significant predictor of falls in women, but not in men (21), whereas 2 other studies found no association in either sex (23,24). Further studies using radiographically measured knee OA and injurious falls are needed, based on our initial findings. Comparing these studies is arduous due to the heterogeneity of the study populations and exposures and outcomes.

The mechanism regarding the association between knee OA and injurious falls in men appears to be perceived pain. Among men with knee sROA or pain without ROA, the increased risk of injurious falls when compared with men without ROA or pain was more than 2-fold. Men with ROA but without pain had a similar incidence of injurious falls as men without ROA or pain. This is not surprising because pain has been shown to be associated with falls in a meta-analysis of 21 studies (45).

Having a history of falls has been shown to be a predictor of incident falls (46). In our study, a history of falls varied significantly by knee OA status, with 50% of adults with pain without ROA having a history of a fall in the last 12 months. However, history of falls did not significantly predict incidence injurious falls in both the univariate and multivariate analyses, suggesting that injurious falls (a proxy for severe falls) may have nonoverlapping risk factors when compared with the outcome of falls.

In consideration of the heterogeneity in the association of knee OA with fall outcomes across studies, it is prudent to consider nonpharmacologic therapies that might help reduce the risk of falls in older adults, particularly older men with knee OA, by improving physical function. For example, EnhanceFitness, an evidence-based community-delivered physical activity program that is recommended by the CDC for adults with arthritis and disseminated by many YMCA recreational facilities across the US, has been shown to produce substantial improvements (18% to 35%) in function (e.g., muscle strength and balance) (47,48). Additionally, the program has shown that consistent EnhanceFitness users had a 26% reduced risk of falls requiring medical care (49). Increased implementation of this intervention or other physical activity interventions (50) may reduce the risk of injurious falls among adults with knee OA, though further studies are needed.

Our study has notable strengths, including being the first to examine radiographic knee OA and the risk of incident injurious falls. Furthermore, fall injuries were determined from both outpatient and inpatient Medicare claims, which allowed outcomes to be collected even if participants did not attend subsequent Heath ABC clinic visits. Medicare includes all potentially relevant health services provided for injurious falls for adults ≥65 years, though it would not capture health services provided by Veterans Affairs. which for this age group may affect missingness of health services data in men more than women. These adjudicated injurious falls from Medicare claims may provide a more complete assessment and time frame of injurious falls versus relying solely on self-reported injurious falls, which are likely subject to recall bias. Moreover, we examined both nonfracture and fracture fall injuries. Finally, we adjusted for many potential confounders, and for a long follow-up period (median of >6.5 years).

Our study does, however, have several potential limitations. First, we measured knee OA at baseline only, and radiographic and pain changes may occur over time. Second, self-report of certain potential confounders may bias findings (e.g., physical activity). Third, the low prevalence of adults with knee ROA and no pain in this sample may have reduced our power to detect associations with injurious falls in this group. Fourth, the 389 participants taken from the Knee Osteoarthritis substudy and excluded from

the analytic sample varied slightly from the analytic sample, which may impact the generalizability of the findings. Although knee OA, age, sex, and BMI did not differ by group, those excluded were more likely to be African American and slightly more educated. In addition, our sample comes from a nondisabled well-functioning population at baseline, which may also affect generalizability. Finally, we adjusted for many potential confounders, but residual confounding is a limitation of all observational studies.

In summary, in a cohort of older men and women, knee sROA was independently associated with a 2.6- fold increased risk of incident injurious falls in men only. More studies are needed to confirm this initial finding and explore why this association was limited to men. Studies with a larger cohort of participants with radiographic evidence of knee OA but no pain are needed to better understand the independent impact of knee OA without pain on injurious falls.

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. Barbour 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. Barbour, Sagawa, Boudreau, Cauley, Nevitt, Strotmeyer.

Acquisition of data. Strotmeyer.

Analysis and interpretation of data. Barbour, Sagawa, Boudreau, Winger, Cauley, Nevitt, Fujii, Patel, Strotmeyer.

REFERENCES

1. Allen KD, Golightly YM. State of the evidence. Current Opin Rheumatol 2015;27:276–83.
2. Dillon CF, Rasch EK, Gu Q, Hirsch R. Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94. J Rheumatol 2006;33:2271–9.
3. Murphy L, Schwartz TA, Helmick CG, Renner JB, Tudor G, Koch G, et al. Lifetime risk of symptomatic knee osteoarthritis. Arthritis Rheum 2008;59:1207–13.
4. Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clin Exp Res 2005;17:255–63.
5. Karpansalo M, Manninen P, Kauhanen J, Lakka TA, Salonen JT. Perceived health as a predictor of early retirement. Scand J Work Environ Health 2004:287–92.
6. Steiner C, Andrews R, Barrett M, Weiss A. HCUP Projections: mobility/orthopedic procedures 2011 to 2012. HCUP projections report 2012;3:2012–03.
7. Liu Q, Niu J, Huang J, Ke Y, Tang X, Wu X, et al. Knee osteoarthritis and all-cause mortality: the Wuchuan Osteoarthritis Study. Osteoarthritis Cartilage 2015;23:1154–7.
8. Kluzek S, Sanchez-Santos M, Leyland K, Judge A, Spector T, Hart D, et al. Painful knee but not hand osteoarthritis is an independent predictor of mortality over 23 years follow-up of a population-based cohort of middle-aged women. Ann Rheum Dis 2016;75:1749–56.
9. Bergen G. Falls and fall injuries among adults aged≥ 65 years—United States, 2014. MMWR Morb Mortal Wkly Rep 2016;65:993–8.
10. Burns ER, Stevens JA, Lee R. The direct costs of fatal and non-fatal falls among older adults—United States. J Safety Res 2016;58:99–103.
11. Rubenstein LZ, Josephson KR. Falls and their prevention in elderly people: what does the evidence show? Med Clin North Am 2006;90:807–24.
12. De Zwart AH, van der Esch M, Pijnappels MA, Hoozemans MJ, van der Leeden M, Roorda LD, et al. Falls associated with muscle strength in patients with knee osteoarthritis and self-reported knee instability. J Rheumatol 2015;42:1218–23.
13. Barbour KE, Lui LY, McCulloch CE, Ensrud KE, Cawthon PM, Yaffe K, et al. Trajectories of lower extremity physical performance: effects on fractures and mortality in older women. J Gerontol A Biol Sci Med Sci 2016;71:1609–15.
14. Cawthon PM, Fullman RL, Marshall L, Mackey DC, Fink HA, Cauley JA, et al. Physical performance and risk of hip fractures in older men. J Bone Miner Res 2008;23:1037–44.
15. Cummings SR, Nevitt MC, Browner WS, Stone K, Fox KM, Ensrud KE, et al. Risk factors for hip fracture in white women. N Engl J Med 1995;332:767–74.
16. Lewis CE, Ewing SK, Taylor BC, Shikany JM, Fink HA, Ensrud KE, et al. Predictors of non-spine fracture in elderly men: the MrOS study. J Bone Miner Res 2007;22:211–9.
17. Mat S, Ng CT, Tan MP. Influence of hip and knee osteoarthritis on dynamic postural control parameters among older faller. J Rehabil Med 2017;49:258–63.
18. Muraki S, Akune T, Oka H, En-yo Y, Yoshida M, Nakamura K, et al. Prevalence of falls and the association with knee osteoarthritis and lumbar spondylosis as well as knee and lower back pain in Japanese men and women. Arthritis Care Res (Hoboken) 2011;63:1425–31.
19. Foley S, Lord SR, Srikanth V, Cooley H, Jones G. Falls risk is associated with pain and dysfunction but not radiographic osteoarthritis in older adults: Tasmanian older adult cohort study. Osteoarthritis Cartilage 2006;14:533–9.
20. Doré AL, Golightly YM, Mercer VS, Shi XA, Renner JB, Jordan JM, et al. Lower-extremity osteoarthritis and the risk of falls in a community-based longitudinal study of adults with and without osteoarthritis. Arthritis Care Res (Hoboken) 2015;67:633–9.
21. Scott D, Blizzard L, Fell J, Jones G. Prospective study of self-reported pain, radiographic osteoarthritis, sarcopenia progression, and falls risk in community-dwelling older adults. Arthritis Care Res (Hoboken) 2012;64:30–7.
22. Tsonga T, Michalopoulou M, Malliou P, Godolias G, Kapetanakis S, Gkasdaris G, et al. Analyzing the history of falls in patients with severe knee osteoarthritis. Clin Orthop Surg 2015;7:449–56.
23. Arden NK, Crozier S, Smith H, Anderson F, Edwards C, Raphael H, et al. Knee pain, knee osteoarthritis, and the risk of fracture. Arthritis Rheum 2006;55:610–5.
24. Muraki S, Akune T, Ishimoto Y, Nagata K, Yoshida M, Tanaka S, et al. Risk factors for falls in a longitudinal population-based cohort study of Japanese men and women: the ROAD Study. Bone 2013;52:516–23.
25. Khalaj N, Abu Osman NA, Mokhtar AH, Mehdikhani M, Wan Abas WA. Balance and risk of fall in individuals with bilateral mild and moderate knee osteoarthritis. PloS one 2014;9:e92270.
26. Riddle DL, Golladay GJ. A longitudinal comparative study of falls in persons with knee arthroplasty and persons with or at high risk for knee osteoarthritis. Age Ageing 2016;45:794–800.
27. Smith T, Pearson M, Latham S. Are people following hip and knee arthroplasty at greater risk of experiencing a fall and fracture? Data from the Osteoarthritis Initiative. Arch Orthop Trauma Surg 2016;136:865–72.

28. Nevitt MC, Tolstykh I, Shakoor N, Nguyen US, Segal NA, Lewis C, et al. Symptoms of knee instability as risk factors for recurrent falls. Arthritis Care Res (Hoboken) 2016;68:1089–97.
29. Nevitt MC, Cummings SR, Hudes ES. Risk factors for injurious falls: a prospective study. J Gerontol 1991;46:164–70.
30. Barbour KE, Stevens JA, Helmick CG, Luo YH, Murphy LB, Hootman JM, et al. Falls and fall injuries among adults with arthritis—United States, 2012. MMWR Morb Mortal Wkly Rep 2014;63:379–83.
31. Sturnieks DL, Tiedemann A, Chapman K, Munro B, Murray SM, Lord SR. Physiological risk factors for falls in older people with lower limb arthritis. J Rheumatol 2004;31:2272–9.
32. O'Loughlin JL, Robitaille Y, Boivin JF, Suissa S. Incidence of and risk factors for falls and injurious falls among the community-dwelling elderly. Am J Epidemiol 1993;137:342–54.
33. Linn S, Murtaugh B, Casey E. Role of sex hormones in the development of osteoarthritis. PM R 2012;4:S169–S73.
34. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
35. Conroy MB, Kwoh CK, Krishnan E, Nevitt MC, Boudreau R, Carbone LD, et al. Muscle strength, mass, and quality in older men and women with knee osteoarthritis. Arthritis Care Res (Hoboken) 2012;64:15–21.
36. Altman RD, Hochberg M, Murphy Jr W, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage 1995;3 Suppl A:3–70.
37. Schwartz AV, Nevitt MC, Brown Jr BW, Kelsey JL. Increased falling as a risk factor for fracture among older women: the study of osteoporotic fractures. Am J Epidemiol 2005;161:180–5.
38. Sagawa N, Marcum ZA, Boudreau RM, Hanlon JT, Albert SM, O'Hare C, et al. Low blood pressure levels for fall injuries in older adults: the Health, Aging and Body Composition Study. Eur J Ageing 2018;15:321–30.
39. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 2000;32 Suppl 1:S498–504.
40. Pahor M, Chrischilles E, Guralnik J, Brown S, Wallace R, Carbonin P. Drug data coding and analysis in epidemiologic studies. Eur J Epidemiol 1994;10:405–11.
41. Pereira CL, Baptista F, Infante P. Men older than 50 yrs are more likely to fall than women under similar conditions of health, body composition, and balance. Am J Phys Med Rehabil 2013;92:1095–103.
42. Davison MJ, Ioannidis G, Maly MR, Adachi JD, Beattie KA. Intermittent and constant pain and physical function or performance in men and women with knee osteoarthritis: data from the osteoarthritis initiative. Clin Rheumatol 2016;35:371–9.
43. Fillingim RB. Sex, gender, and pain: women and men really are different. Curr Pain Headache Rep 2000;4:24–30.
44. Welmer AK, Rizzuto D, Calderón-Larrañaga A, Johnell K. Sex differences in the association between pain and injurious falls in older adults: a population-based longitudinal study. Am J Epidemiol 2017;186:1049–56.
45. Stubbs B, Binnekade T, Eggermont L, Sepehry AA, Patchay S, Schofield P. Pain and the risk for falls in community-dwelling older adults: systematic review and meta-analysis. Arch Phys Med Rehabil 2014;95:175–87.
46. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. Epidemiology 2010;21:658–68.
47. Wallace JI, Buchner DM, Grothaus L, Leveille S, Tyll L, LaCroix AZ, et al. Implementation and effectiveness of a community-based health promotion program for older adults. J Gerontol A Biol Sci Med Sci 1998;53:M301–6.
48. Belza B, Shumway-Cook A, Phelan EA, Williams B, Snyder SJ, LoGerfo JP. The effects of a community-based exercise program on function and health in older adults: The EnhanceFitness Program. J Appl Gerontol 2006;25:291–306.
49. Greenwood-Hickman MA, Rosenberg DE, Phelan EA, Fitzpatrick AL. Participation in older adult physical activity programs and risk for falls requiring medical care, Washington State, 2005–2011. Prev Chronic Dis 2015;12.
50. Gillespie LD, Robertson MC, Gillespie WJ, Sherrington C, Gates S, Clemson LM, et al. Interventions for preventing falls in older people living in the community. Cochrane Database Syst Rev CD007146;2012.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 875–884
DOI 10.1002/acr.23698

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

What Is the Evidence to Support the Association Between Metabolic Syndrome and Osteoarthritis? A Systematic Review

Shanshan Li[1] and David T. Felson[2]

Objective. There is conflicting evidence on the association between metabolic syndrome (MetS) with the risk of osteoarthritis (OA). We aimed to systematically summarize the empirical evidence and discuss challenges in research methodologies in addressing this question.

Methods. We performed a systematic literature review based on PubMed, Embase, Web of Science, and the Cochrane Database of Systematic Reviews on published epidemiologic studies that examined the association between MetS and the risk of OA. We included cross-sectional studies, case–control studies, and cohort studies with appropriate covariate adjustments. We extracted information on prevalence, incidence, crude and adjusted effect estimates, and the 95% confidence intervals from the articles, or this information was provided by the authors. We listed the main methodologic issues existing in current literature and provided recommendations for future research on this topic.

Results. We identified 7 eligible studies on knee OA, 3 on hip OA, and 3 on hand OA. In studies that adjusted for body mass index or weight, MetS was not significantly associated with the risk of knee OA. No significant associations were reported for hip OA. For hand OA, the data were sparse and insufficient to reach a conclusion. Studies were mostly cross-sectional, exposure included only 1 time measurement, few studies had incident outcomes, and covariate adjustment was often insufficient.

Conclusion. Our review was unable to reach a definitive conclusion due to insufficient data, although the data suggest that knee and hip OA are not associated with MetS. Future longitudinal studies with incident OA cases, repeated measurement of MetS, and appropriate covariate adjustment are needed.

INTRODUCTION

Worldwide prevalence of osteoarthritis (OA) is increasing due to aging and the obesity epidemic. Understanding the etiology of OA and subsequent systemic consequences of OA is important, because it confers a significant public health burden. Metabolic syndrome (MetS), a cluster of several cardiometabolic risk factors and a common accompaniment of obesity, is defined as central obesity, dyslipidemia, impaired fasting glucose, and hypertension. Substantial evidence suggests that MetS is associated with an increased risk of cardiovascular disease, type 2 diabetes mellitus, and cancer. Emerging evidence also links MetS with the risk of OA. Obese individuals are at high risk of developing OA, not only in the knee but also in non–weight-bearing joints such as the hand. The association between obesity and hand OA has suggested that the loading conferred by obesity on weight-bearing joints is not the sole explanation for the high risk of OA among obese individuals. Evidence has emerged that chronic inflammation, insulin resistance, and production of abnormal adipocytokines from adipose tissues (such as tumor necrosis factor, interleukin [IL]-1, IL-6, leptin, and adiponectin) may play a role in the etiology of OA. However, the potential mechanisms underlying this association are unclear.

Early reports on the link between MetS and OA appeared in 1990, and the number of articles focusing on this issue has increased dramatically over the past 15 years (see Supplemen-

[1]Shanshan Li, ScD: Boston University School of Medicine, Boston, Massachusetts; [2]David T. Felson, MD, MPH: Boston University School of Medicine, Boston, Massachusetts, and Arthritis Research UK Epidemiology Unit and National Institute for Health Research Biomedical Research Centre, University of Manchester, Manchester, UK.

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

Address correspondence to Shanshan Li, ScD, Sloan Epidemiology Unit, Boston University School of Medicine, 650 Albany Street, X-200, Boston, MA 02118. E-mail: shl607@bu.edu.

Submitted for publication February 12, 2018; accepted in revised form July 10, 2018.

SIGNIFICANCE & INNOVATIONS

- Articles published on the association between metabolic syndrome (MetS) and osteoarthritis (OA) have focused mostly on potential biologic mechanisms. There has been little critical examination of the evidence, including quality of study design and statistical analyses. Whether a link between MetS and OA really exists is still under debate.
- For knee and hip OA, after adjustment for body mass index or weight, most studies showed a null association.
- We found that most existing evidence on the association of MetS and OA risk was of limited quality. Either large high-quality longitudinal studies in the future or a pooling of studies will permit a further examination of the association of MetS, especially with hand OA. Clarifying the relationship could offer opportunities for prevention of OA and have important public health and policy implications.

tary Figure 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23698/abstract). Review articles published on this topic have focused mostly on the potential biologic mechanisms underlying this link. There has been little critical examination of the evidence, including quality of study design and statistical analyses. Whether a link between MetS and OA really exists is still under debate (1). There are at least 2 major study design concerns that are highly relevant to studies of MetS and OA. First, since the loading conferred by obesity is likely to be an independent cause of OA (especially of knee and hip OA), studies examining metabolic factors with OA need to include an adjustment for weight or body mass index (BMI). Second, most published studies are cross-sectional and therefore provide only limited evidence for causality. The strongest evidence for causal relations comes from high-quality longitudinal studies with incident OA as outcomes (2). Thus, the goal of our review was to provide a systematic summary of evidence, discuss challenges in epidemiologic study design and issues regarding the study of an association between MetS and OA, discuss the difficulty in analyses and evaluate strengths of the existing evidence, and suggest future directions.

MATERIALS AND METHODS

Data sources and searches. We conducted a systematic literature review on epidemiologic studies using PubMed, Web of Science, the Cochrane Database of Systematic Reviews, and Embase. In addition to the articles found, we searched references of all identified articles. The search strategies for the PubMed search are shown in Supplementary Appendices 1–3, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23698/abstract. All searches were conducted for published literature up to May 16, 2018. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement and the guidelines for performing a meta-analysis and systematic review of observational studies in epidemiology.

Study selection. Both authors (SL and DTF) independently evaluated each study's eligibility and study quality. Discrepancies were resolved by consensus. To be included, studies had to: 1) report data from an original, peer-reviewed study; 2) be of cross-sectional, case–control, or prospective cohort design using a noninstitutionalized adult population (age >18 years); 3) be a study of humans with and without OA; 4) characterize participants as to whether they had MetS or not; 5) define OA as clinical OA, knee or hip replacement due to OA, or symptomatic OA or radiographic OA, with the latter 2 including imaging evidence of OA; 6) report an association between the 2 conditions; and 7) have adjustment for confounding factors, including adjustment for BMI or weight. For studies published in languages other than English, we reviewed the English abstract and if the full article was needed, we asked a native speaker to translate the article into English. For multiple articles published from the same study, we reviewed all but presented details only from the most recent qualified article.

Data extraction and quality assessment. Information from each selected study was extracted by both authors independently. We evaluated each study based on study design, study population, exposure and outcome definitions, confounding control, bias assessment, and statistical methods, as well as the study-defined effect estimates.

Data synthesis and analysis. The odds ratio or hazard ratio was reported in eligible studies. Due to the limited numbers of longitudinal studies with adequate quality and the heterogeneity of these studies, there was not sufficient data for a meta-analysis. To be consistent with most eligible studies, we presented results on MetS as a binary variable (yes or no).

RESULTS

We found 506 studies on the topic of knee/hip/hand OA and MetS from PubMed, 702 studies from Web of Science, 0 from the Cochrane Database of Systematic Reviews, and 555 studies from Embase up to May 16, 2018. More than 1 article was published based on data from the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) cohort (3). We used the one with the most detailed analyses on numbers of MetS components that was a prospective cohort study with multivariable adjustment (3).

MetS with knee OA. For knee OA, 7 studies met our inclusion criteria, of which 2 were cross-sectional (4,5), 1 was

Table 1. Study characteristics of knee osteoarthritis studies on the association between metabolic syndrome and knee osteoarthritis (OA)*

Type	Longitudinal	Longitudinal	Longitudinal	Longitudinal
Author, year (ref.)	Hellevik et al, 2018 (7)	Niu et al, 2017 (8)	Monira Hussain et al, 2014 (9)	Engstrom et al, 2009 (10)
Study	Nord-Trøndelag Health Study 2 linked to the Norwegian Arthroplasty Register	Framingham Heart Study Offspring cohort	Melbourne Collaborative cohort study	Malmo Diet and Cancer Study
Design	Prospective cohort	Prospective cohort	–	Prospective cohort
Country	Norway	US	Australia	Sweden
Baseline	1995–1997	1992–1995	2003–2007	1991–1994
No.	62,661	991 with no OA at baseline	20,430: 660 TKR, 19,208 no joint replacement	5,171
Mean age	49 years	54.2 years	68 years	57.5 years
Follow-up	15.4 years	10 years	6.8 years	12.4 years
Exclusion	956 excluded due to previous joint replacement of hip or knee (n = 796), missing date of operation (n = 158), or emigration during baseline period (n = 2)	Excluded knees with prevalent OA at baseline	Missing anthropometric measurements	Excluded participants with history of OA
Exposure	Joint Interim Statement definition of MetS: presence of ≥3 of the following: waist circumference ≥102 cm for men and ≥88 cm for women, systolic blood pressure ≥130 mm Hg or diastolic blood pressure ≥85 mm Hg, or treatment of previously diagnosed hypertension, triglycerides ≥1.7 mmol/liter, HDL cholesterol men <1.1 mmol/liter, women <1.3 mmol/liter, and glucose ≥5.6 mmol/liter or self-reported type 2 diabetes mellitus	Assessment of MetS in 1990–1993, NCEP-ATP III criteria for MetS; ≥3 of the following: abdominal obesity (waist circumference, men ≥102 cm, women ≥88 cm), high triglyceride levels (≥150 mg/dl), low HDL cholesterol (men <40 mg/dl, women <50 mg/dl), high blood pressure (systolic ≥130 mm Hg or diastolic ≥85 mm Hg, or treatment for high blood pressure), high fasting glucose (≥110 mg/dl or diagnosis of diabetes mellitus); MetS was also defined using the modified ATP III criteria (high fasting glucose based on the cut point of ≥100 mg/dl)	International Federation of Diabetes definition of MetS: central obesity (waist circumference men ≥94 cm, women ≥80 cm) and any 2 of the following: raised serum triglyceride ≥1.7 mmol/liter, reduced serum HDL cholesterol (men <1.03 mmol/liter, women <1.29 mmol/liter, or specific treatment for lipid abnormalities), raised blood pressure (systolic ≥130 mm Hg or diastolic ≥85 mm Hg, or treatment of previously diagnosed hypertension), and impaired fasting glucose ≥5.6 mmol/liter or previously diagnosed type 2 diabetes mellitus	NCEP-ATP III criteria for MetS: presence of any 3 components: high waist circumference (men ≥102 cm, women ≥88 cm), low HDL (men <1.03 mmol/liter, women <1.29 mmol/liter, or lipid-lowering medication), hypertension (≥130/85 mm Hg or treatment for hypertension), hyperglycemia (fasting plasma glucose ≥5.6 mmol/liter), hypertriglyceridemia (≥1.7 mmol/liter or treatment)

(continued)

Table 1. *(Cont'd)*

Outcome	Incident of total knee replacement: 1,111 total	Incident radiographic OA (men: n = 35; women: n = 31) for K/L grade ≥2; incident symptomatic OA (men: n = 18, women: n = 32) present when a knee developed new combination of radiographic OA and knee pain	660 total knee replacements	89 knee OA, defined as a first knee arthroplasty or high tibia osteotomy, in combination with a diagnosis of OA
Covariates	Sex, smoking, physical activity, education, BMI	Age, sex, education, current smoking, physical activity, alcohol consumption, BMI (kg/m^2)	Age, sex, country of birth, level of education, physical activity, BMI	Age, sex, BMI, smoking, CRP level, physical activity
Statistical analysis	Cox regression model stratified by age group (<50, 50–69.9, ≥70 years)	Sex-specific analysis; generalized estimating equation	Cox proportional hazards models	Cox proportional hazards model
Results	HR 0.89 (95% CI 0.63–1.26) for age <50 years; HR 1.16 (95% CI 0.96–1.41) for age 50–69.9 years; HR 1.27 (95% CI 0.85–1.90) for age ≥70 years	Radiographic OA: ATP III criteria: men RR 1.4 (95% CI 0.8–2.5), women RR 0.8 (95% CI 0.5–1.4); modified ATP III criteria: men RR 1.2 (95% CI 0.7–2.0), women RR 0.8 (95% CI 0.5–1.3); symptomatic OA: ATP III criteria: men RR 0.8 (95% CI 0.4–1.6), women RR 1.2 (95% CI 0.7–2.1); modified ATP III criteria: men RR 1.3 (95% CI 0.7–2.5), women RR 1.1 (95% CI 0.6–1.9)†	HR 1.92 (95% CI 1.59–2.32) before adjusting for BMI; HR 1.24 (95% CI 1.02–1.52) after adjusting for BMI	MetS (yes vs. no) for all: RR 2.1 (95% CI 1.3–3.3) before adjusting for BMI, RR 1.1 (95% CI 0.7–1.8) after adjusting for BMI; MetS (yes vs. no) men RR 1.4 (95% CI 0.6–1.3) before adjusting for BMI, RR 0.6 (95% CI 0.2–1.5) after adjusting for BMI; MetS (yes vs. no) women RR 2.5 (95% CI 1.5–4.4) before adjusting for BMI, RR 1.4 (95% CI 0.8–2.6) after adjusting for BMI

(continued)

Table 1. *(Cont'd)*

Type	Case–control	Cross-sectional	Cross-sectional
Author, year (ref.)	Askari et al, 2017 (6)	Visser et al, 2015 (5)	Shin, 2014 (4)
Study	Fasa Osteoarthritis Study	Netherlands Epidemiology of Obesity study	Fifth Korean National Health and Nutrition Examination Survey (2010)
Design	Case–control, matched for sex	Cross-sectional analyses of baseline measurements	Cross-sectional analyses
Country	Iran	The Netherlands	Korea
Baseline	2013	2008	2010
No.	131 OA patients, 262 matched controls	5,002 participants with a self-reported BMI ≥27 kg/m^2 in the greater area of Leiden; 1,671 inhabitants from Leiderdorp were invited irrespective of BMI	2,363
Mean age, years	OA group 52.9 years, controls 55.5 years	45–65 years	63.4 years
Follow-up	NA	NA	NA
Exclusion	Controls who had radiographic complication caused by OA in the knee and hip were not included	–	–
Exposure	NCEP-ATP III definition of MetS: presence of ≥3 of the following: systolic blood pressure >130 mm Hg and/or diastolic blood pressure >85 mm Hg; triglyceride ≥150 mg/dl; high density lipoprotein cholesterol (men <40 mg/dl, women <50 mg/dl); fasting blood sugar >100; waist circumference women >88 cm, men >102 cm	NCEP-ATP III definition of MetS: presence of ≥3 of the following: elevated waist circumference (men ≥102 cm, women ≥88 cm), elevated triglycerides (≥1.7 mmol/liter or treatment for elevated triglycerides), reduced HDL cholesterol (men <1.03 mmol/liter, women <1.3 mmol/liter, or treatment for reduced HDL cholesterol), elevated blood pressure (systolic ≥130 mm Hg, diastolic ≥85 mm Hg or antihypertensive medication), elevated fasting glucose (≥5.6 mmol/liter or glucose lowering medication)	NCEP-ATP III definition of MetS: presence of ≥3 of the following: waist circumference men ≥90 cm, women ≥85 cm; triglyceride level ≥150 mg/dl or medication use; HDL cholesterol men <40 mg/dl, women <50 mg/dl or medication use; blood pressure of 130/85 mm Hg or greater or medication use; fasting glucose ≥100 mg/dl or medication use
Outcome	131 OA patients, 262 controls matched for sex; diagnosis of OA based on K/L score >1; controls were from referrals to orthopedic clinic of the hospital without symptoms of pain and/or stiffness in knee or hip	Clinical OA defined according to ACR clinical criteria; presence of a knee prosthesis was considered knee OA	Presence of radiographic knee OA defined as a K/L grade ≥2
Covariates	Sex, age, BMI	Age, sex, height, smoking, education, ethnicity	Age, sex, income, smoking, alcohol consumption, physical activity, BMI
Statistical analysis	Multiple logistic regression model	Logistic regression for cross-sectional analyses	Logistic regression
Results	MetS (yes vs. no): RR 6.8 (95% CI 4.1–11.4) before adjusting for age, sex, and BMI; RR 10.9 (95% CI 5.5–21.8) after adjusting for age, sex, and BMI	RR 1.56 (95% CI 1.24–1.97) before adjusting for weight; RR 1.08 (95% CI 0.85–1.39) after adjusting for weight	RR 1.49 (95% CI 1.23–1.79) before adjusting for BMI; RR 0.92 (95% CI 0.74–1.13) after adjusting for BMI; RR 1.49 (95% CI 1.23–1.79) before adjusting for weight; RR 1.04 (95% CI 0.84–1.27) after adjusting for weight

* Studies were presented by the order of study design and publication year. ref. = reference; TKR = total knee replacement; MetS = metabolic syndrome; HDL = high-density lipoprotein; NCEP-ATP III = National Cholesterol Education Program Adult Treatment Panel III revised guideline; K/L = Kellgren/Lawrence; BMI = body mass index; CRP = C-reactive protein; HR = hazard ratio; 95% CI = 95% confidence interval; RR = risk ratio; NA = not applicable; ACR = American College of Rheumatology.
† For both radiographic and symptomatic OA, results from Niu et al did not report RR before adjustment for BMI. We were therefore unable to present RRs with and without BMI adjustment.

Table 2. Study characteristics of hip osteoarthritis (OA) studies on the association between metabolic syndrome and hip OA*

Type	Longitudinal	Longitudinal	Longitudinal
Author, year (ref.)	Hellevik et al, 2018 (7)	Monira Hussain et al, 2014 (9)	Engstrom et al, 2009 (10)
Study	Nord-Trøndelag Health Study 2 linked to the Norwegian Arthroplasty Register	Melbourne Collaborative cohort study	Malmo Diet and Cancer Study
Design	Prospective cohort	Prospective cohort	Prospective cohort
Country	Norway	Australia	Sweden
Baseline	1995–1997	2003–2007	1991–1994
No.	62,661	562 primary hip replacement, 19,208 with no joint replacement	5,171
Mean age	49 years	68 years	57.5 years
Follow-up	15.4 years	6.8 years	12 years
Exposure	Joint Interim Statement definition of MetS: presence of ≥3 of the following: waist circumference men ≥102 cm, women ≥88 cm; systolic blood pressure ≥130 mm Hg or diastolic blood pressure ≥85 mm Hg, or treatment of previously diagnosed hypertension; triglycerides ≥1.7 mmol/liter; HDL cholesterol men <1.1 mmol/liter, women <1.3 mmol/liter; glucose ≥5.6 mmol/liter or self-reported type 2 diabetes mellitus	International Federation of Diabetes definition of MetS: central obesity (waist circumference men ≥94 cm, women ≥ 80 cm), and any 2 of the following: raised serum triglyceride ≥1.7 mmol/liter, reduced serum HDL cholesterol (men <1.03 mmol/liter, women <1.29 mmol/liter, or specific treatment for lipid abnormalities), raised blood pressure (systolic ≥130 mm Hg or diastolic ≥85 mm Hg, or treatment of previously diagnosed hypertension), and impaired fasting plasma glucose ≥5.6 mmol/liter or previously diagnosed type 2 diabetes mellitus	NCEP-ATP III definition of MetS: presence of ≥3 components: waist circumference (men ≥102 cm, women ≥88 cm), low HDL (men <1.03 mmol/liter, women <1.29 mmol/liter, or lipid-lowering medication), hypertension (≥130/85 mm Hg or treatment for hypertension), hyperglycemia (fasting plasma glucose ≥5.6 mmol/liter) and hypertriglyceridemia (≥1.7 mmol/liter or treatment)
Outcome	Incident total hip replacement	Total hip replacement (n = 562)	Arthroplasty due to severe hip OA (n = 120)
Covariates	Sex, smoking, physical activity, education, BMI	Age, sex, country of birth, level of education, physical activity, BMI	Age, sex, smoking, physical activity, CRP, BMI
Statistical analysis	Cox regression model stratified by age group (<50, 50–69.9, ≥70 years)	Cox proportional hazards model	Cox proportional hazards model
Results	HR 0.58 (95% CI 0.40–0.83) for age <50 years; HR 0.93 (95% CI 0.79–1.10) for age 50–69.9 years; HR 0.83 (95% CI 0.65–1.14) for age ≥70 years	MetS not significantly associated with hip replacement before and after adjusting for BMI; HR 1.19 (95% CI 0.95–1.49) before adjusting for BMI; HR 1.00 (95% CI 0.78 1.27) after adjusting for BMI	For both men and women, HR 1.00 (95% CI 0.60–1.50) before adjusting for BMI, HR 0.7 (95% CI 0.4–1.2) after adjusting for BMI; for men only, HR 0.9 (95% CI 0.4–1.8) before adjusting for BMI, HR 0.7 (95% CI 0.3–1.6) after adjusting for BMI; for women only, HR 1.0 (95% CI 0.6–1.7) before adjusting for BMI, HR 0.7 (95% CI 0.4–1.3) after adjusting for BMI

* Studies were presented by the order of study design and publication year. See Table 1 for abbreviations.

a case–control study (6), and 4 were cohort studies (7–10) (Table 1). Most of the reported effect estimates suggested a null association of MetS with knee OA after adjustment for BMI or weight (5,7,8,10). Despite meeting all of our selection criteria, the quality of the following studies and results needs to be interpreted with caution. The cross-sectional study performed by Shin using the fifth Korean National Health and Nutrition Examination Survey demonstrated a 1.49-fold increased risk for knee OA, and results became nonsignificant after further adjustment for body weight or BMI (4). This study did not account for sampling weights in the statistical analysis, even though the study, a nationwide survey, had a stratified, multistage probability sampling design. Contrary to other literature, the small Fasa Osteoarthritis Study showed that the odds ratio between MetS and OA paradoxically increased from 6.8 to 10.9 after adjusting for age, sex, and BMI (6).

MetS with hip OA. For hip OA, 3 studies were eligible, all of which were cohort studies (7,9,10) (Table 2). All studies had large sample sizes and a long duration of follow-up. Results before and after BMI adjustment were consistently null. Overall, there was no association between MetS with hip OA.

MetS with hand OA. For hand OA, the only longitudinal study showed a null association (11). However, this study was of small sample size and studied only whites. We did not have sufficient data to reach a definitive conclusion in this review (5,11,12) (Table 3). The 2016 study by Tomi et al focused on patients with HIV and may not be generalizable to the general population (12). Although not included in our review, the Netherlands Epidemiology of Obesity study provided important evidence on adiposity, particularly visceral fat, associated with a 1.3-fold elevated risk for hand OA in men (13).

DISCUSSION

In our current review, most evidence pointed to a null association of MetS with knee and hip OA. For hand OA, the data were limited and conflicting and were not sufficient to allow us to reach a definitive conclusion. Our systematic review showed that the strongest evidence came from a few longitudinal studies (7–10). More rigorous longitudinal evidence is needed.

Overall, we found methodologic deficiencies in most studies examining MetS and OA. Few studies used longitudinal data with sufficient sample sizes to assess associations between MetS and OA (7–10), and even fewer studies excluded prevalent OA cases, or prevalent joint replacement at baseline (7,8,10). The adequacy of control for confounding varied considerably across studies. The definition of MetS and its relevant exposure window were not clear, with existing studies having only 1 time measurement of MetS in adulthood. Since obesity, sedentary lifestyle, and MetS may increase as a consequence of knee or hip OA, cross-sectional studies examining MetS and OA in these joints are limited in their ability to make causal inferences. Evidence was mainly from the US, Europe, and Asia, and future studies from black, Hispanic, and other minority populations are needed.

Of cross-sectional and longitudinal studies of MetS and knee or hip OA, several showed associations unadjusted for BMI or weight (4,5,9,10). In unadjusted analyses from all of these studies, there was a significantly increased risk of OA, and in all, this association diminished greatly and became nonsignificant in all but 1 when analyses were adjusted for BMI (4,5,9,10). In the Fasa Osteoarthritis Study, the odds ratio paradoxically increased from 6.8 to 10.9 after adjustment for age, sex, and BMI (6). Despite consistent findings from the literature that age is a strong risk factor for OA, this study showed a paradoxically reduced risk of OA with advanced age (6).

Inferring causality from observational studies is challenging and is based on multiple assumptions. Hill criteria include strength, consistency, specificity, temporality, biologic gradient, plausibility, coherence, experimental evidence, and analogy (2). Among all components, temporality is the most important consideration. Evidence from cross-sectional studies contributes less when compared with longitudinal studies, with exposure preceding outcome and confounder control (14). The evidence is further strengthened if there is a longitudinal study with control for baseline confounders (14). When there is a potential feedback between the exposure and outcome (for example, hip and knee OA leading to obesity) over time, cross-sectional studies are subject to reverse causation and cannot be used for drawing inferences. When reverse causation is highly unlikely, cross-sectional studies provide some evidence. If study investigators have a clear rationale on the temporal ordering of the exposure and outcome, and the confounding variables are likely to temporally precede the exposure and outcome, then cross-sectional evidence can be helpful (14).

Knee and hip OA can lead to changes in lifestyle, such as reduced physical activity level and weight gain, that may increase the subsequent risk for MetS. Whether there is a bidirectional relationship between MetS and OA is currently unclear. For studies on MetS as a risk factor for OA, differentiating incident from prevalent knee OA is therefore important.

Prevalent hand OA cases may be less subject to reverse causation, because the hand is a non–weight-bearing joint, and hand OA might not be likely to lead to dramatic lifestyle changes when compared with knee or hip OA. However, an absence of evidence does not prove a null association, and more data on the longitudinal association of MetS with hand OA are needed. Currently, only 1 study investigated this association longitudinally (11) (Table 2).

The current evidence focused on incident OA as defined by a Kellgren/Lawrence grade or total joint replacement, which

Table 3. Study characteristics of hand osteoarthritis studies on the association between metabolic syndrome and hand osteoarthritis (OA)*

Type	Longitudinal	Cross-sectional	Cross-sectional
Author, year (ref.)	Strand et al, 2018 (11)	Tomi et al, 2016 (12)	Visser et al, 2015 (5)
Study	Framingham OA Study	Metabolic Syndrome and Fibrosis-Osteoarthritis study (METAFIB-OA)	Netherlands Epidemiology of Obesity (NEO) study
Design	Longitudinal analyses	Cross-sectional single-center study on patients with HIV infection	Cross-sectional analyses of baseline measurements
Country	US	France	The Netherlands
Baseline	1992–1995	2011	2008
No.	586	458 with HIV infection, ages 45–64 years, from January 2011 to December 2012 from outpatient clinics in France	5,002 participants with self-reported BMI ≥27 kg/m^2 in the greater area of Leiden; 1,671 inhabitants from Leiderdorp were invited irrespective of their BMI
Mean age	50–75 years	45–64 years	45–65 years
Follow-up	2 years	NA	NA
Exclusion	Excluded hand OA at baseline	–	–
Exposure	American Heart Association/National Heart, Lung, Blood Institute clinical/laboratory MetS definition: presence of ≥3 of the following: central obesity (waist circumference men ≥102 cm, women ≥88 cm), hypertension (systolic blood pressure ≥130 mm Hg, diastolic blood pressure ≥85 mm Hg, and/or antihypertensive treatment), diabetes mellitus (fasting blood glucose ≥100 mg/dl and/or anti-diabetes treatment), elevated triglycerides (≥150 mg/dl), and low HDL (men <40 mg/dl, women <50 mg/dl, and/or cholesterol-lowering treatment); we created an alternative definition based on clinical and laboratory measurements only (i.e., excluding treatment-related information), since individuals with MetS who are adequately treated theoretically may have a lower risk of OA	International Diabetes Federation criteria: central obesity (waist circumference men ≥94 cm, women ≥80 cm [in Europe] or with ethnicity-specific values for other groups); plus any 2 of the following: triglycerides ≥1.7 mmol/liter, or specific treatment for this lipid abnormality; HDL cholesterol men <1 mmol/liter, women <1.3 mmol/liter, or specific treatment for this lipid abnormality; increased blood pressure with systolic ≥130 or diastolic ≥85 mm Hg, or treatment of previously diagnosed hypertension; increased fasting plasma glucose ≥100 mg/dl (5.6 mmol/liter), or previously diagnosed type 2 diabetes mellitus	ATP III criteria for MetS: presence of ≥3 of the following: elevated waist circumference (men ≥102 cm, women ≥88 cm), elevated triglycerides (≥1.7 mmol/liter or treatment for elevated triglycerides), reduced HDL cholesterol (men <1.03 mmol/liter men, women <1.3 mmol/liter, or treatment for reduced HDL cholesterol), elevated blood pressure (systolic ≥130 mm Hg, diastolic ≥85 mm Hg or antihypertensive medication), elevated fasting glucose (≥5.6 mmol/liter or glucose lowering medication)
Outcome	Incident radiographic hand OA (n = 56); ≥2 distal or proximal interphalangeal joints with K/L grade ≥2	Any patient with ≥1 finger joint scored K/L grade ≥2 was considered to have radiographic hand OA	Clinical OA defined according to ACR clinical criteria; bony and soft swellings as well as deformities of distal interphalangeal, proximal interphalangeal, metacarpophalangeal, carpometacarpal, and wrist joints were assessed
Covariates	Age, sex, BMI	Age, sex, previous hand trauma, HIV duration, CD4 level, HIV viral load, duration of exposure to protease inhibitors	Age, sex, height, smoking, education, ethnicity; analyses on surrogates for systemic process were adjusted for weight
Statistical analysis	Logistic regression	Logistic regression	Weighted analyses; logistic regression for cross sectional analyses
Results	MetS definition including treatment RR 0.98 (95% CI 0.66–1.46) before adjusting for BMI; RR 0.91 (95% CI 0.58–1.44) after adjusting for BMI; MetS clinical/laboratory definition: RR 1.00 (95% CI 0.67–1.50) before adjusting for BMI; RR 0.92 (95% CI 0.58–1.49) after adjusting for BMI	RR 2.18 (95% CI 1.26–3.96) for hand OA	RR 1.62 (95% CI 1.23–2.13) before adjusting for weight; RR 1.46 (95% CI 1.06–2.02) after adjusting for weight

* Studies were presented by the order of study design and publication year. See Table 1 for abbreviations.

is already late in the pathogenesis of joint pathology. For studies using joint replacement as outcomes, excluding prevalent joint replacement at baseline is important. Hellevik et al (7) excluded prevalent joint replacement at baseline and reported effect estimates for incident joint replacement. Future studies examining earlier stages of disease (possibly with magnetic resonance imaging) and joint pain are needed.

Except for the Melbourne Collaborative cohort study (9) (562 hip replacement, 660 knee replacement) and the Nord-Trøndelag Health Study (7) (1,840 hip replacement, 1,111 knee replacement), current evidence was based on a limited number of incident OA events. In general, studies may not have been large enough to give a definite answer and had low power to detect potential effect modification. Combining data from studies to leverage existing cohort resources and examine potential effect modification, such as by sex and race, may be needed.

Currently there is no consistent definition of MetS. In our review, we did not exclude studies based on the definition of MetS, and we listed all definitions of MetS that were used in the original studies. The heterogeneity in MetS definitions has contributed to published studies with different definitions and criteria for MetS. Consistency of the research findings across different MetS definitions would increase the robustness and generalizability of the results.

Current literature mainly focused on MetS as a one-time measurement, thus studying a fixed prevalent exposure. There is a need for studies examining changes in MetS with risk of OA. Questions regarding how MetS changes and how MetS in childhood or early life is related to OA risk in later life is unknown. Further, according to the developmental origins of chronic disease theory, obesity, cardiovascular disease, and type 2 diabetes mellitus develop during intrauterine exposure and fetal programming. Whether this programming is also true for OA etiology is currently unclear. Future study on the intergenerational effects of MetS on OA risk using a life course approach may be of interest.

Whether there is a potential effect modification by sex on the association between MetS and OA is currently unclear. Only 2 studies have tested for this potential effect modification. In the Framingham OA study, Niu et al (8) found no sex-specific findings. Engstrom et al (10) also found MetS and OA risk to be similar by sex. Due to the limited numbers of studies presenting sex-specific results, we cannot have a definite answer regarding effect modification by sex.

Current evidence is subject to uncontrolled and unmeasured confounding. Socioeconomic factors, dietary information, and lifestyle factors contribute to the etiology of both MetS and OA. However, existing studies have not fully considered nor were able to account for influences of these factors. Each study would have been strengthened if the authors had presented additional analyses on these as potential confounders of the association between MetS and OA and had presented information on how observed results would be changed by these potential confounders. Bias analysis methods have been developed to evaluate the robustness of evidence from observational studies (15). None of the published studies included in our search performed sensitivity or bias analyses to assess the influence of unmeasured confounding. We encourage future studies to perform sensitivity analyses and demonstrate the robustness of study findings.

The Korean National Health and Nutrition Examination Study showed a nonsignificant association between MetS and OA after further adjustment for weight or BMI. Despite a stratified, multistage probability sampling design, this study used a logistic regression model without adjusting for sampling weights, which can bias results. It would be helpful for future studies to provide results with sampling weights adjustment.

While this systematic review focused on overall MetS, many of the articles included also gave data on individual components of MetS and their relation with OA. As with MetS, there is a burgeoning literature on the potential relation of each of these with OA. Further, for each, there are unique biologic reasons to suspect a relationship. Our review is insufficient to grapple comprehensively with the evidence linking such elements of MetS as hypertension, diabetes mellitus, lipid abnormalities, and visceral adiposity with OA. Our suggestions about study design are highly relevant to the study of these issues as well.

The number of reviews on MetS and OA has grown exponentially in recent years without a critical look at the quality of the evidence. Our review focused on published studies and thus may be subject to publication bias. Further, our review focused on MetS, and future studies are needed to look into each individual component of MetS.

In conclusion, there was insufficient data from large high-quality studies on the association of MetS with OA, especially hand OA. However, the preponderance of high-quality evidence suggests that there is no association of MetS with either knee or hip OA. Future evidence from large high-quality longitudinal studies is needed. We suggest either larger longitudinal studies or a pooling of studies to permit a further examination of this association, as well as a focus on examining earlier stages of disease. If every future epidemiologic study could address the methodologic considerations mentioned above, the potential for prevention of OA may be substantial, and the results could have important public health and policy implications.

ACKNOWLEDGMENTS

The authors thank David Flynn at Boston University School of Medicine for his help with the literature search and Dr. Michael LaValley for advice. The authors also thank the reviewers for their insightful comments.

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. Li 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. Li, Felson.
Acquisition of data. Li, Felson.
Analysis and interpretation of data. Li, Felson.

REFERENCES

1. Appleton CT, Hawker GA, Hill CL, Pope JE. "Weighing in" on the Framingham Osteoarthritis Study: measuring biomechanical and metabolic contributions to osteoarthritis [editorial]. Arthritis Rheumatol 2017;69:1127–30.
2. Rothman KJ. Modern epidemiology. 3rd ed. Philadelphia: Wolters Kluwer Health; 2008.
3. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, et al. Accumulation of metabolic risk factors such as overweight, hypertension, dyslipidaemia, and impaired glucose tolerance raises the risk of occurrence and progression of knee osteoarthritis: a 3-year follow-up of the ROAD study. Osteoarthritis Cartilage 2012;20:1217–26.
4. Shin D. Association between metabolic syndrome, radiographic knee osteoarthritis, and intensity of knee pain: results of a national survey. J Clin Endocrinol Metab 2014;99:3177–83.
5. Visser AW, de Mutsert R, le Cessie S, den Heijer M, Rosendaal FR, Kloppenburg M. The relative contribution of mechanical stress and systemic processes in different types of osteoarthritis: the NEO study. Ann Rheum Dis 2015;74:1842–7.
6. Askari A, Ehrampoush E, Homayounfar R, Arasteh P, Naghizadeh MM, Yarahmadi M, et al. Relationship between metabolic syndrome and osteoarthritis: the Fasa Osteoarthritis Study. Diabetes Metab Syndr 2017;11 Suppl 2:S827–32.
7. Hellevik AI, Johnsen MB, Langhammer A, Baste V, Furnes O, Storheim K, et al. Metabolic syndrome as a risk factor for total hip or knee replacement due to primary osteoarthritis: a prospective cohort study (the HUNT study and the Norwegian Arthroplasty Register). Clin Epidemiol 2018;10:83–96.
8. Niu J, Clancy M, Aliabadi P, Vasan R, Felson DT. Metabolic syndrome, its components, and knee osteoarthritis: the Framingham Osteoarthritis Study. Arthritis Rheumatol 2017;69:1194–203.
9. Monira Hussain S, Wang Y, Cicuttini FM, Simpson JA, Giles GG, Graves S, et al. Incidence of total knee and hip replacement for osteoarthritis in relation to the metabolic syndrome and its components: a prospective cohort study. Semin Arthritis Rheum 2014;43:429–36.
10. Engstrom G, Gerhardsson de Verdier M, Rollof J, Nilsson PM, Lohmander LS. C-reactive protein, metabolic syndrome and incidence of severe hip and knee osteoarthritis: a population-based cohort study. Osteoarthritis Cartilage 2009;17:168–73.
11. Strand MP, Neogi T, Niu J, Felson DT, Haugen IK. Association between metabolic syndrome and radiographic hand osteoarthritis: data from a community-based longitudinal cohort study. Arthritis Care Res (Hoboken) 2018;70:469–74.
12. Tomi AL, Sellam J, Lacombe K, Fellahi S, Sebire M, Rey-Jouvin C, et al. Increased prevalence and severity of radiographic hand osteoarthritis in patients with HIV-1 infection associated with metabolic syndrome: data from the cross-sectional METAFIB-OA study. Ann Rheum Dis 2016;75:2101–7.
13. Visser AW, Ioan-Facsinay A, de Mutsert R, Widya RL, Loef M, de Roos A, et al. Adiposity and hand osteoarthritis: the Netherlands Epidemiology of Obesity study. Arthritis Res Ther 2014;16:R19.
14. VanderWeele TJ, Jackson JW, Li S. Causal inference and longitudinal data: a case study of religion and mental health. Soc Psychiatry Psychiatr Epidemiol 2016;51:1457–66.
15. Li S, Stampfer MJ, Williams DR, VanderWeele TJ. Association of religious service attendance with mortality among women. JAMA Intern Med 2016;176:777–85.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 885–892
DOI 10.1002/acr.23719

American College of Rheumatology
Empowering Rheumatology Professionals

Most Important Frequently Asked Questions From Patients With Hip or Knee Osteoarthritis: A Best-Worst Scaling Exercise

Aniek A. O. M. Claassen,[1] Keetie C. A. L. C. Kremers–van de Hei,[2] Frank H. J. van den Hoogen,[3] Willemijn H. van der Laan,[1] Wim H. C. Rijnen,[4] Sander Koëter,[2] Joris Botman,[5] Vincent J. J. F. Busch,[1] Henk J. Schers,[4] and Cornelia H. M. van den Ende[3]

Objective. To collect and prioritize the frequently asked questions (FAQs) that patients with hip or knee osteoarthritis (OA) and health care professionals consider to be the most important; to identify informational needs that go beyond guideline recommendations.

Methods. FAQs were collected among health care professionals and from the arthritis helpline of the Dutch Arthritis Foundation. After deleting overlapping FAQs, the remaining FAQs were prioritized by patients and health care professionals using a maximum difference scaling method. A hierarchical Bayesian method was used to calculate relative importance scores. Differences between health care professionals and patients were analyzed using independent *t*-tests.

Results. A total of 28 health care professionals and the arthritis helpline provided 192 FAQs. After deleting overlapping FAQs, 60 FAQs were prioritized by 94 patients (57 [60.6%] women, mean age 67.3 years) and 122 health care professionals (67 [54.9%] women, mean age 45.7 years). The FAQ "What can I do myself to decrease symptoms and to prevent the OA from getting worse?" was prioritized as the most important by both patients and professionals. FAQs that were highly prioritized by patients but significantly different from professionals were more directed toward treatment options offered by health care professionals, whereas highly prioritized FAQs of professionals were more often focused on treatment options involving self-management.

Conclusion. The health care professionals' perspective on informational needs differs from that of OA patients. These differences are important to address in order to achieve more active involvement of patients in their own treatment process.

INTRODUCTION

Patient education is a cornerstone in the management of chronic conditions like osteoarthritis (OA) (1). Providing relevant disease-related and self-management–related information helps patients become actively involved in their own care process (2). Moreover, research has shown that the need for information among OA patients is high (3–5).

A number of informational sources, including health professionals such as general practitioners (GPs) and physiotherapists, health-related web sites, patient information leaflets, and family and friends, are available for patients. When patients obtain information from more than 1 source, they may encounter conflicting information (4). Receiving conflicting information has been found to be associated with undesirable outcomes, such as reduced medication adherence in patients with vasculitis or arthritis and in pregnant women (4–6). Moreover, receiving conflicting expert opinions may be perceived as incompetence of the experts, which in turn has been associated with lower intentions to pursue health behaviors that are known to be beneficial (7).

National and international guidelines for hip or knee OA recommend the provision of accurate information about the condi-

Supported by Coöperatie VGZ.

[1]Aniek A. O. M. Claassen, MSc, PT, Willemijn H. van der Laan, MD, PhD, Vincent J. J. F. Busch, MD, PhD: Sint Maartenskliniek, Nijmegen, The Netherlands; [2]Keetie C. A. L. C. Kremers–van de Hei, MSc, Sander Koëter, MD, PhD: Canisius Wilhelmina Hospital, Nijmegen, The Netherlands; [3]Frank H. J. van den Hoogen, MD, PhD, Cornelia H. M. van den Ende, PhD, PT: Sint Maartenskliniek and Radboud University Medical Center, Nijmegen, The Netherlands; [4]Wim H. C. Rijnen, MD, PhD, Henk J. Schers, MD, PhD: Radboud University Medical Center, Nijmegen, The Netherlands; [5]Joris Botman, PT: Stichting Gezondheidscentrum De Kroonsteen–De Vuursteen, Nijmegen, The Netherlands.

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

Address correspondence to Cornelia H. M. van den Ende, PhD, PT, Sint Maartenskliniek, Department of Rheumatology, PO Box 9011, 6500 GM, Nijmegen, The Netherlands. E-mail: e.vandenende@maartenskliniek.nl.

Submitted for publication December 21, 2017; accepted in revised form July 24, 2018.

SIGNIFICANCE & INNOVATIONS

- This study provides informational needs of patients with hip or knee osteoarthritis (OA) that go beyond guideline recommendations.
- The question "What can I do myself to decrease symptoms and to prevent the OA from getting worse?" is considered the most important by patients and health care professionals for all hip and knee OA patients.
- The perspective on information needs differs between patients and health care professionals.
- Patients consider questions oriented toward treatment options offered by health care professionals important, whereas frequently asked questions that are highly prioritized by health care professionals are more directed toward self-management.

tion and its management, to counter misunderstandings for all patients (8–10). Despite these recommendations, patients indicate that they do not always receive the information they need to manage their disease adequately (11). Barriers related to health care professionals can contribute to the lack of provision of consistent and sufficient information. In a systematic review, Egerton et al (12) identified barriers for primary care health care professionals in providing the recommended management of OA to patients: OA was not always seen as a serious condition, but rather as a part of normal aging or as less important than other conditions; health care professionals felt underprepared, because of the lack of clarity and specificity of guidelines or because of their own lack of knowledge about OA treatments; personal beliefs about recommended treatments on effectiveness and patient adherence varied among health care professionals; and health care professionals were challenged by patients' expectations that were other than their own views. These barriers underpin the importance of finding consensus among multiple health care professionals about the content and phrasing of information for OA patients and to formulate this information from a common perspective to make it consistent and clear for information transfer to the patient.

In a recent study, French et al (13) used a multistage consensus process to identify key messages that are essential for patients to know, extracted from multiple guidelines about OA. After optimizing the wording of the key messages, an overall ranking of the messages averaged across all panel members was determined. These messages can be used in patient educational material because they are a translation of evidence-based information. However, although these statements are identified as being essential for patients, they do not necessarily cover patients' needs and preferences for information on topics that go beyond what is covered in guidelines. Because the preferences and the needs of patients are important in their decision-making for treatments, such items are important to investigate (14). Porcheret et al (15) found that patients considered information about the biomedical approach important for an OA consultation in primary care, while current psychosocial and behavioral approaches are recommended in guidelines (10). Therefore, there might be differences in what health care professionals consider important information and what patients want to know and what patients in different stages of their disease consider important (3).

To make an inventory of the informational needs of OA patients that goes beyond guideline recommendations, and to evaluate whether those needs are perceived differently by health care professionals, the current study aimed to answer the following research questions: 1) What are the most important frequently asked questions (FAQs) of patients with hip or knee OA? 2) Are there differences in rating of importance of FAQs between patients and professionals? 3) What is the difference in informational needs of patients and those perceived by professionals? and 4) Are there differences in informational needs among patients in different treatment settings?

PATIENTS AND METHODS

The setting and context for the research was the region of Nijmegen in The Netherlands, where a collaboration among health care providers involved in the care of hip or knee OA set up a conjoint educational program for patients in several communities to harmonize care in the region. A 2-step approach was followed to make an inventory of the most important FAQs.

Step 1: inventory. *Participants.* The following stakeholders and health care professionals were invited by e-mail or newsletter to provide FAQs for the inventory: all orthopedic surgeons (n = 25; specialized in hip or knee surgery) from 3 hospitals in the Nijmegen area, i.e., Radboud University Medical Center (Radboudumc), Canisius Wilhelmina Hospital (CWZ), and Sint Maartenskliniek Hospital (SMK); rheumatologists from the Radboudumc and SMK (n = 29); nurse practitioners from CWZ (n = 2); GPs involved in this project, as well as colleague GPs working in their general practice (n = 24); primary care physiotherapists involved in this project or connected to a local OA network for health professionals specialized in rheumatic diseases (n = 188); and the Dutch Arthritis Foundation, which provided FAQs from OA patients made to the professionals on their telephone helpline.

Procedure. The Dutch College of General Practitioners (Nederlands Huisartsen Genootschap) recently launched the web site www.thuisarts.nl (i.e., home doctor), which provides information about OA (among other diagnoses). Questions and topics about OA covered on this web site include: What is OA?, What are symptoms of OA?, What causes OA?, How is OA diagnosed?, Medication and OA, and Exercise therapy and OA. The information on this web site is based on national and international guidelines. Our aim was not to restrict to FAQs on basic information about hip or knee OA, because these top-

ics are usually covered on many web sites and in educational material. Therefore, we asked participating health care professionals to record 5–10 FAQs they often get from their patients about OA that are not covered on the web site www.thuisarts.nl. Two researchers evaluated all collected FAQs to reduce the total number when possible. First, duplicates were deleted. Next, the researchers individually identified FAQs that were similar in formulation, and based on discussion, we decided which could be combined. Last, FAQs that could be answered with the content of the web site mentioned above were also deleted. A total of 60 FAQs were included for the prioritization step.

Step 2: prioritizing. *Participants.* To prioritize the FAQs, a survey was developed and distributed among patients with hip or knee OA and health care professionals working in the field of OA. GPs from 2 local medical centers were asked to invite patients with an OA diagnosis by mail. These patients were selected through the GP's information system. Inclusion criteria for patients were a diagnosis of hip or knee OA, age >18 years, ability to communicate well in Dutch, basic computer skills, having an email address, and a willingness to participate in the study and sign an informed consent. A total of 398 patients were sent an information letter about the study, with a reply card.

The same health care professionals who were asked for the inventory step were asked to participate in the prioritizing step. In addition, GPs from the Radboudumc Practice Based Research Network, Department of Primary and Community Care (n = 420) were invited to participate. To ensure diversity with regard to the type of discipline and setting, corresponding to the Dutch health care system, we aimed at including health care professionals in the following occupational groups: 35 primary care physiotherapists or exercise therapists, 35 GPs, 20 orthopedic surgeons (including physicians in training to be a specialist and nurse practitioners specialized in orthopedics), and 10 rheumatologists (including physicians in training to be a specialist and physician assistants specialized in rheumatology).

Overall we aimed at including 100 patients and 100 professionals. No guidance is provided in the literature regarding the minimal sample size for a desired statistical power for best-worst scaling methods (16). Sample sizes of previous studies evaluated in a review ranged between 15 and 1,296 participants (16).

Survey development and procedure. We developed an online survey consisting of 2 parts. In the first part, we assessed demographic characteristics of all respondents: age (years), sex (male/female), and education level (low/high). Patients were asked to answer additional questions on the affected joint (hip/knee/both), years since diagnosis (<1, 1–5, 5–10, or >10), and setting (primary care, secondary care, or postsurgery) based on health care use ("Did you visit an orthopedic surgeon in the past?" [yes/no] and "Did you already have joint replacement surgery?" [yes/no]). In the questionnaire for health care providers, we assessed the occupation (GP, physiotherapist, rheumatologist, orthopedic surgeon, or other) and years in practice of professionals.

In the second part of the survey, we prioritized the FAQs from the inventory according to relative importance by means of a maximum difference scaling (MaxDiff) exercise (also known as best-worse scaling). In this methodology, participants are shown a subset of possible items and asked to indicate (among this subset) the most and least important item. Participants complete a number of these sets, an exercise in which each set contains a different subset of items (17). The MaxDiff method has been used successfully for research questions in rheumatology (18,19). In our study, all of the 60 FAQs were presented twice in subsets of 5 FAQs (20), resulting in 24 subsets for each participant. For each subset, participants were asked to indicate the most and least important FAQ that should be answered for all OA patients. In the current study, Sawtooth Software's SSI Web platform (http://www.sawtoothsoftware.com/products/maxdiff-software) was used to develop the online questionnaire with the MaxDiff exercise. The software creates the optimal design of subsets based on 1,000 iterations. A total of 300 versions was created to ensure a variety of combinations of FAQs and a randomized order among participants, to avoid higher importance being given to the first FAQ mentioned. An open link was created to be disseminated to patients and professionals.

Statistical analysis. Descriptive analyses were used to describe demographic characteristics of participants. The choices made by respondents in the MaxDiff exercise were analyzed using the hierarchical Bayesian (HB) method to estimate relative importance (RI) scores (Sawtooth Software I). The HB method allowed us to estimate the individual level of importance by combining information from individuals' specific choices with the distribution of importance across participants, computing individual-level weights under the logit rule. Raw scores were generated by iteration on an interval scale. To facilitate interpretation, the scores were subsequently rescaled to a standardized 0–100 ratio scale; the higher the score, the more important the FAQ. Furthermore, a FAQ with an RI of 5 is twice as important as a FAQ with a RI of 2.5. All RIs sum to 100 for each individual. Thus, the RIs represent the relative importance of an FAQ in relation to all other FAQs. The HB analysis provides a root likelihood (RLH) for random responders. Based on the number of items shown per set (5 in the current study) an RLH >0.269 indicates that the responses appear thoughtful and consistent (21). The HB analyses were performed for patients and health care professionals separately, using the Sawtooth Software platform. Analyses were performed only on data of participants who completed the exercise. The software generated raw scores and RIs for each FAQ per individual respondent.

Potential differences in RIs between the patients' and professionals' top 5 FAQs were analyzed using independent *t*-tests, performed with Stata 13 software. For all analyses, a

significance level of *P* less than or equal to 0.05 was assumed. Additionally, FAQs that differed by ≥1.67 in RI between patients and health care professionals were addressed. We considered a difference of 1.67 to be relevant, because this is the average score per FAQ when the total points (100) are distributed over the 60 FAQs. Differences in RIs in different settings were explored descriptively.

Ethics approval. This study protocol (no. 2017-3184) was presented to the Medical Research Ethics Committee, region Arnhem-Nijmegen, The Netherlands. An exemption was obtained, because this type of study does not require ethics approval according to Dutch law. All participants of the prioritizing step provided online informed consent.

RESULTS

Inventory. A total of 28 health care professionals (11 rheumatologists, 7 orthopedic surgeons, 1 nurse practitioner, 6 GPs, and 3 physiotherapists) took part, and the Dutch Arthritis Foundation provided 192 FAQs. From these FAQs, 104 were deleted because they were duplicates, they could be combined with another FAQ (n = 10), or they could be answered with the information on www.thuisarts.nl (n = 13). Another 5 FAQs were deleted because they were unclearly formulated (n = 3) or addressed other joints than the hip or knee (n = 2). The remaining 60 FAQs were used in the prioritizing step.

Prioritizing. *Participants.* A total of 107 patients started the online questionnaire, and 99 completed the MaxDiff exercise. One patient had an RLH <0.269 and was excluded. Four patients were excluded because they reported that their symptoms were caused by something other than OA. Characteristics of 94 participants (response rate of 24%) are shown in Table 1. Half of the patients had OA symptoms for <5 years. Approximately half of the patients had already had a joint replacement as treatment for their OA.

A total of 140 health care professionals started the online questionnaire, of whom 124 finished the MaxDiff exercise. Two health care professionals were excluded because they did not fit the occupational categories, resulting in data of 122 health care professionals (response rate of 18%) usable for this study. The median years in practice of participating health care professionals was 15.5 years (interquartile range 6–25) (Table 1).

Most important and least important FAQs. Figure 1 shows the top 5 most important FAQs ranked by patients and health care professionals. The overall prioritizing of all 60 FAQs by patients and health care professionals is shown in Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23719/abstract. The highest ranked FAQ both for patients and for health care professionals was "What can I do myself to decrease symptoms and to prevent the OA from getting worse?" Another FAQ that was prioritized in the top 5 for both groups but had significantly different RIs was "What is the natural course of OA?" FAQs that were in the patients' top 5 but were not ranked in the top 5 of the health care professionals were: "What are the newest treatment options?"; "Is there any medication that can either slow down or stop OA?"; and "What are the latest research results concerning OA?" Three FAQs that were in the professionals' top 5 but not in the patients top 5 were: "What can or can I not do in terms of exercise and physical activity?"; "I'm young and I have OA. What changes should I make to my life and what should or shouldn't I do anymore?"; and "Can exercise or being physically active be harmful to my joints?"

Seven FAQs that differ by at least 1.67 in RI score between patients and health care professionals are shown in Table 2. Two FAQs that were scored considerably higher by patients than by health care professionals were "What are the latest research

Table 1. Demographic characteristics of patients and health care professionals participating in the prioritizing step*

Characteristics	Values
Patients (n = 94)	
Age, mean ± SD years	67.3 ± 8.1
Women	57 (60.6)
Education level low	45 (47.9)
Affected joint	
Hip	41 (43.6)
Knee	46 (48.9)
Hip and knee	7 (7.5)
Time since diagnosis, years	
<1	8 (8.5)
1–5	42 (44.7)
5–10	19 (20.2)
>10	25 (26.6)
Setting	
Primary care	25 (26.6)
Secondary care	26 (27.7)
Post joint replacement	43 (45.7)
Professionals (n = 122)	
Age, mean ± SD years	45.7 ± 10.3
Women	67 (54.9)
Occupation	
Physiotherapist	42 (34.4)
General practitioner	49 (40.2)
Orthopedic surgeon	18 (14.8)
Rheumatologist	13 (10.6)
Years in practice, median (IQR), years	15.5 (6–25)

* Values are the number (%) unless indicated otherwise. IQR = interquartile range.

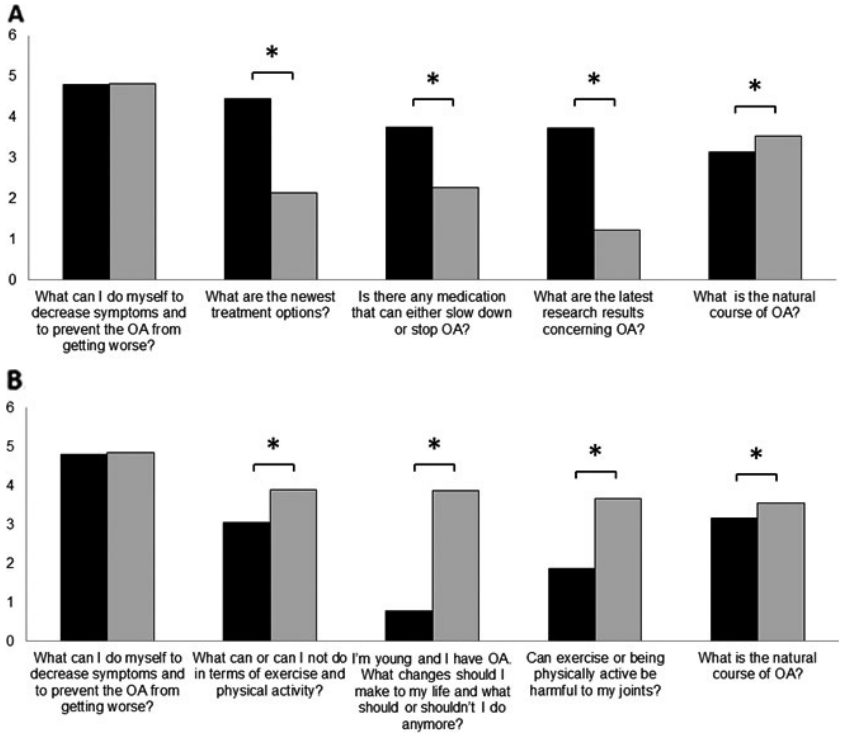


Figure 1. **A**, Relative importance scores (RIs) of patients (solid bars) and health care professionals (shaded bars) of the top 5 most important frequently asked questions (FAQs) as prioritized by patients. **B**, RIs of patients (solid bars) and health care professionals (shaded bars) of the top 5 most important FAQs as prioritized by health care professionals. OA = osteoarthritis; * = significant differences ($P < 0.05$).

results concerning OA?" and "What are the newest treatment options?" The other 5 FAQs were scored higher by health care professionals than by patients (Figure 1).

Exploring subgroup differences. Exploration of subgroup differences between primary care, secondary care, and postsurgery patients did not reveal differences in RIs ≥1.67, indicating the absence of large differences in RI of FAQs between patient groups (Table 2).

DISCUSSION

The FAQ "What can I do myself to decrease symptoms and to prevent the OA from getting worse?" was prioritized as most important by both patients and professionals in the current study. Other FAQs that were highly prioritized by patients were more directed toward treatment options offered by different health care professionals, whereas highly prioritized FAQs of professionals were more often focused on treatment options involving self-management.

The highest ranked FAQ by both patients and health care professionals, "What can I do myself to decrease symptoms and to prevent the OA from getting worse?" highlights the importance of patients' need for information about OA and to feel confident in managing their condition (22,23). This need is in concordance with the guideline recommendation that self-management is important

Table 2. Frequently asked questions (FAQs) with a relative importance absolute difference of ≥1.67 between patients and health care professionals*

FAQ	Patients	Professionals
Can exercise or being physically active be harmful to my joints?	1.87	3.66
Can being overweight be harmful for my joints?	0.94	2.86
Can I continue doing my job or do I need to make certain adaptations to my working environment because of my OA?	0.71	2.71
Why do I have to try all these other treatment options, when surgery is also an option?	1.16	3.17
What are the latest research results concerning OA?	3.74	1.23
What are the newest treatment options?	4.47	2.14
I'm young and I have OA. What changes should I make to my life and what should or shouldn't I do anymore?	0.78	3.88

* OA = osteoarthritis.

for patients with hip or knee OA (8). It is also in line with the nature of the key messages that French et al (13) reported to be essential for hip or knee OA patients, because 15 of those 21 statements covered treatment options and management in which the patients have an active role. This concordance between patients and health care professionals provides a good starting point for education and communication about treatment options. However, a recent review by Chou et al (24) showed that patients have a need for specific guidance. This need can also been seen in our results, where the majority of collected FAQs were very specific, i.e., "What sports can I still do? At what frequency and intensity?" and "What can the orthopedic surgeon or rheumatologist do for me?" Concrete recommendations about self-management, for instance about the type of exercise or how to navigate the health care system, should be handed to patients more directly (25). Providing this information not only in the consultation room but especially in an educational program provided by multiple health care professionals can be a good option to encourage patients toward the principles of self-management (26).

The importance of an active role by the patient is clearly found in the health care professionals' top 5 choices. For instance "What can or can I not do in terms of exercise and physical activity?" and "Can exercise or being physically active be harmful to my joints?" can be seen as questions that illustrate the importance of conservative treatment options in which the patient plays an active role. This concern shows that health care professionals need to put effort into explaining the important active role patients can have in alleviating their symptoms and controlling their OA. Such an effort may be a challenge, because health care professionals themselves do not always have confidence in the outcome of conservative treatment options and in the willingness and capability of patients to play an active role (12,27) . A lack of communication between health care professionals may contribute to low confidence in conservative treatment options (27). Collaboration among health care professionals to answer FAQs may improve insight among professionals on each other's role and perspective on OA treatment. In addition, a joint endeavor of multiple disciplines to answer FAQs offers the opportunity to provide consistent knowledge about OA, which is important because patients report that they receive unclear and inconsistent information (24). The FAQs from our study offer a starting point for discussion to achieve consensus on the content of information and to improve patient education.

Two FAQs that were scored considerably higher by patients than by professionals were "What are the latest research results concerning OA?" and "What are the newest treatment options?" These findings are in line with previous research showing that patients have an interest in recent developments and experimental treatments for their condition (28,29). Active information-seeking behavior of patients has developed in the past decades, and access to information on the internet may contribute to this behavior (30). Although information on experimental treatments is not applicable to all patients, for health care professionals to explicitly address this informational need could be worthwhile. Because effective management of OA requires actively involved patients, delivery of patient-centered care seems essential in this respect (31). After learning about the lack of new treatment options or the unknown effects of experimental treatments, patients may be more open to information about current conservative treatment options like physical activity and weight loss.

We explored differences between patient subgroups in different settings of treatment descriptively, but we did not find any large differences. One issue that should be taken into account is that patients with replaced joints seem to be overrepresented in our study population, because 47% of patients reported 1 or more joint replacements. However, participants were asked to take the perspective of all patients with OA, and FAQ ranking proved similar in those with and without replacements. This outcome suggests that information is important for every OA patient. Because of our small subgroups, this result should be interpreted with caution. Brembo et al (3) identified informational needs related to the disease continuum of hip OA patients. For instance, a key question during the phase when symptoms significantly decrease quality of life is: "I can't stand the pain, is it time for surgery?" However, the researchers' aim was specifically focused on identifying needs per stages of disease, rather than studying differences. Further research into informational needs at different stages of OA is therefore recommended.

There are several limitations to our study that should be mentioned. First, with a response rate between 18% and 24% there might be selection bias. We invited a large sample of patients from primary care and health care professionals from different settings, but we have no characteristics of the nonresponders and could not compare responders to nonresponders. Responders might have higher informational needs, but this possibility does not necessarily affect prioritization. Second, we included health care professionals from different disciplines, and there might be differences between these subgroups in prioritization. Although no clear guidelines for a minimal sample size for the MaxDiff method are given, in our view our sample size did not allow comparisons between subgroups of health care professionals. Health care professionals were invited from primary care and from 3 different medical centers: 1 university medical center, 1 specialized hospital for rheumatology, orthopedic surgery, and rehabilitation medicine, and 1 local general hospital. This variety of hospital types assures a good representation of The Netherlands. However, because of differences in health care systems, generalizability to other countries should be taken with caution. Last, because participants had to answer 24 subsets in the MaxDiff exercise, they may have given less attention at the end of the questionnaire, but using the RLH as an indicator should have limited this problem.

A strength of our study is that we included new FAQs based on input from a wide range of patients and health care professionals, from different fields in both primary and secondary care, that were not evaluated in previous studies and that are not men-

tioned in national and international guidelines. By investigating informational needs beyond guideline recommendations, we provided specific practical points for information that can be given to patients in daily practice and in future interventions. In addition to providing basic information, health care professionals can spend time on highly ranked topics at the expense of identified topics that are considered less important.

This study provides informational needs of patients with hip or knee OA that go beyond guideline recommendations. Our results provide starting points for optimizing patient education and improving information given in daily clinical practice. A next step should be to formulate answers for the most important FAQs, with health care professionals from different disciplines, to provide patients with consistent information from a common perspective. These answers can be used in educational programs and materials.

ACKNOWLEDGMENTS

The authors thank Gertie Schouten, Wilma Peters, and Heleen Govers (patient representatives) for their contribution to the project group. We thank Tanya Lévy (consultant health innovations, VGZ), Pieter van Haren (manager Innovations, VGZ), Marjolein de Rijk (manager Partnerships, VGZ), Wim Scheurs (orthopedic surgeon, Radboudumc), Marlies van Pelt (physiotherapist, Fysiotherapie Hatert), and John Reijnen (physiotherapist, Medical Training Malden) for their work in the project group. We also thank Tim Pelle (PhD student, Sint Maartenskliniek) for his help with the inventory process of the FAQs.

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 den Ende 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. Claassen, Kremers–van de Hei, van den Hoogen, van der Laan, Rijnen, Koëter, Botman, Busch, Schers, van den Ende.

Acquisition of data. Claassen.

Analysis and interpretation of data. Claassen, Kremers–van de Hei, Koëter, Schers, van den Ende.

REFERENCES

1. Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: translating evidence into action. Health Aff (Millwood) 2001;20:64–78.
2. Novak M, Costantini L, Schneider S, Beanlands H. Approaches to self-management in chronic illness. Semin Dial 2013;26:188–94.
3. Brembo EA, Kapstad H, Eide T, Mansson L, Van DS, Eide H. Patient information and emotional needs across the hip osteoarthritis continuum: a qualitative study. BMC Health Serv Res 2016;16:88.
4. Carpenter DM, DeVellis RF, Fisher EB, DeVellis BM, Hogan SL, Jordan JM. The effect of conflicting medication information and physician support on medication adherence for chronically ill patients. Patient Educ Couns 2010;81:169–76.
5. Carpenter DM, Elstad EA, Blalock SJ, DeVellis RF. Conflicting medication information: prevalence, sources, and relationship to medication adherence. J Health Commun 2014;19:67–81.
6. Hameen-Anttila K, Nordeng H, Kokki E, Jyrkka J, Lupattelli A, Vainio K, et al. Multiple information sources and consequences of conflicting information about medicine use during pregnancy: a multinational Internet-based survey. J Med Internet Res 2014;16:e60.
7. Carpenter DM, Geryk LL, Chen AT, Nagler RH, Dieckmann NF, Han PK. Conflicting health information: a critical research need. Health Expect 2016;19:1173–82.
8. 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.
9. Jordan KM, Arden NK, Doherty M, Bannwarth B, Bijlsma JW, Dieppe P, et al. EULAR recommendations 2003: an evidence based approach to the management of knee osteoarthritis: report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). Ann Rheum Dis 2003;62:1145–55.
10. National Clinical Guideline Centre. Osteoarthritis: care and management in older adults. 2014. URL: https://www.nice.org.uk/guidance/cg177/resources/osteoarthritis-care-and-management-pdf-35109757272517.
11. McHugh GA, Campbell M, Luker KA. Quality of care for individuals with osteoarthritis: a longitudinal study. J Eval Clin Pract 2012;18:534–41.
12. Egerton T, Diamond LE, Buchbinder R, Bennell KL, Slade SC. A systematic review and evidence synthesis of qualitative studies to identify primary care clinicians' barriers and enablers to the management of osteoarthritis. Osteoarthritis Cartilage 2017;25:625–38.
13. French SD, Bennell KL, Nicolson PJ, Hodges PW, Dobson FL, Hinman RS. What do people with knee or hip osteoarthritis need to know? An international consensus list of essential statements for osteoarthritis. Arthritis Care Res (Hoboken) 2015;67:809–16.
14. Montori VM, Brito JP, Murad MH. The optimal practice of evidence-based medicine: incorporating patient preferences in practice guidelines. JAMA 2013;310:2503–4.
15. Porcheret M, Grime J, Main C, Dziedzic K. Developing a model osteoarthritis consultation: a Delphi consensus exercise. BMC Musculoskelet Disord 2013;14:25.
16. Cheung KL, Wijnen BF, Hollin IL, Janssen EM, Bridges JF, Evers SM, et al. Using best-worst scaling to investigate preferences in health care. Pharmacoeconomics 2016;34:1195–209.
17. Muhlbacher AC, Zweifel P, Kaczynski A, Johnson FR. Experimental measurement of preferences in health care using best-worst scaling (BWS): theoretical and statistical issues. Health Econ Rev 2015;6:5.
18. Meyfroidt S, Hulscher M, De Cock D, Van der Elst K, Joly J, Westhovens R, et al. A maximum difference scaling survey of barriers to intensive combination treatment strategies with glucocorticoids in early rheumatoid arthritis. Clin Rheumatol 2015;34:861–9.
19. Van Hulst LT, Kievit W, van Bommel R, van Riel PL, Fraenkel L. Rheumatoid arthritis patients and rheumatologists approach the decision to escalate care differently: results of a maximum difference scaling experiment. Arthritis Care Res (Hoboken) 2011;63:1407–14.
20. Wirth R, Wolfrath A. Using MaxDiff for evaluating very large sets of items. Sawtooth Software Conference; 2012 March 12; Orlando, Florida.
21. Sawtooth Software. Identifying "bad" responders: fit statistic and identifying random responders. 2017. URL: https://www.sawtoothsoftware.com/help/lighthouse-studio/manual/index.html?hid_web_maxdiff_badrespondents.html.

22. Mann C, Gooberman-Hill R. Health care provision for osteoarthritis: concordance between what patients would like and what health professionals think they should have. Arthritis Care Res (Hoboken) 2011;63:963–72.
23. Smith TO, Purdy R, Lister S, Salter C, Fleetcroft R, Conaghan PG. Attitudes of people with osteoarthritis towards their conservative management: a systematic review and meta-ethnography. Rheumatol Int 2014;34:299–313.
24. Chou L, Ellis L, Papandony M, Seneviwickrama KL, Cicuttini FM, Sullivan K, et al. Patients' perceived needs of osteoarthritis health information: a systematic scoping review. PLoS One 2018;13:e0195489.
25. Winter Di Cola J, Juma S, Kennedy D, Dickson P, Denis S, Robarts S, et al. Patients' perception of navigating "the system" for arthritis management: are they able to follow our recommendations? Physiother Can 2014;66:264–271.
26. Moe RH, Grotle M, Kjeken I, Olsen IC, Mowinckel P, Haavardsholm EA, et al. Effectiveness of an integrated multidisciplinary osteoarthritis outpatient program versus outpatient clinic as usual: a randomized controlled trial. J Rheumatol 2016;43:411–8.
27. Selten EM, Vriezekolk JE, Geenen R, van der Laan WH, van der Meulen-Dilling RG, Nijhof MW, et al. Barriers impeding the use of non-pharmacological, non-surgical care in hip and knee osteoarthritis: the views of general practitioners, physical therapists and medical specialists. J Clin Rheumatol 2017;23:405–10.
28. Baumann M, Euller-Ziegler L, Guillemin F. Evaluation of the expectations osteoarthritis patients have concerning healthcare, and their implications for practitioners. Clin Exp Rheumatol 2007;25:404–9.
29. Chen AT. Information seeking over the course of illness: the experience of people with fibromyalgia. Musculoskeletal Care 2012;10:212–20.
30. Van Woerkum CM. The Internet and primary care physicians: coping with different expectations. Am J Clin Nutr 2003;77 Suppl:1016S–8S.
31. Papandony MC, Chou L, Seneviwickrama M, Cicuttini FM, Lasserre K, Teichtahl AJ, et al. Patients' perceived health service needs for osteoarthritis (OA) care: a scoping systematic review. Osteoarthritis Cartilage 2017;25:1010–25.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 893–902
DOI 10.1002/acr.23712

American College of Rheumatology
Empowering Rheumatology Professionals

Antinuclear Antibody–Negative Systemic Lupus Erythematosus in an International Inception Cohort

May Y. Choi,[1] Ann E. Clarke,[1] Yvan St. Pierre,[2] John G. Hanly,[3] Murray B. Urowitz,[4] Juanita Romero-Diaz,[5] Caroline Gordon,[6] Sang-Cheol Bae,[7] Sasha Bernatsky,[2] Daniel J. Wallace,[8] Joan T. Merrill,[9] David A. Isenberg,[10] Anisur Rahman,[10] Ellen M. Ginzler,[11] Michelle Petri,[12] Ian N. Bruce,[13] Mary A. Dooley,[14] Paul R. Fortin,[15] Dafna D. Gladman,[4] Jorge Sanchez-Guerrero,[16] Kristjan Steinsson,[17] Rosalind Ramsey-Goldman,[18] Munther A. Khamashta,[19] Cynthia Aranow,[20] Graciela S. Alarcón,[21] Susan Manzi,[22] Ola Nived,[23] Asad A. Zoma,[24] Ronald F. van Vollenhoven,[25] Manuel Ramos-Casals,[26] Guillermo Ruiz-Irastorza,[27] S. Sam Lim,[28] Kenneth C. Kalunian,[29] Murat Inanc,[30] Diane L. Kamen,[31] Christine A. Peschken,[32] Soren Jacobsen,[33] Anca Askanase,[34] Thomas Stoll,[35] Jill Buyon,[36] Michael Mahler,[37] and Marvin J. Fritzler[1]

Objective. The spectrum of antinuclear antibodies (ANAs) is changing to include both nuclear staining as well as cytoplasmic and mitotic cell patterns (CMPs) and accordingly a change is occurring in terminology to anticellular antibodies. This study examined the prevalence of indirect immunofluorescence (IIF) anticellular antibody staining using the Systemic Lupus International Collaborating Clinics inception cohort.

Methods. Anticellular antibodies were detected by IIF on HEp-2000 substrate using the baseline serum. Three serologic subsets were examined: ANA positive (presence of either nuclear or mixed nuclear/CMP staining), anticellular antibody negative (absence of any intracellular staining), and isolated CMP staining. The odds of being anticellular antibody negative versus ANA or isolated CMP positive was assessed by multivariable analysis.

Results. A total of 1,137 patients were included; 1,049 (92.3%) were ANA positive, 71 (6.2%) were anticellular antibody negative, and 17 (1.5%) had an isolated CMP. The isolated CMP–positive group did not differ from the ANA-positive or anticellular antibody–negative groups in clinical, demographic, or serologic features. Patients who were older (odds ratio [OR] 1.02 [95% confidence interval (95% CI) 1.00, 1.04]), of white race/ethnicity (OR 3.53 [95% CI 1.77, 7.03]), or receiving high-dose glucocorticoids at or prior to enrollment (OR 2.39 [95% CI 1.39, 4.12]) were more likely to be anticellular antibody negative. Patients on immunosuppressants (OR 0.35 [95% CI 0.19, 0.64]) or with anti-SSA/Ro 60 (OR 0.41 [95% CI 0.23, 0.74]) or anti–U1 RNP (OR 0.43 [95% CI 0.20, 0.93]) were less likely to be anticellular antibody negative.

Conclusion. In newly diagnosed systemic lupus erythematosus, 6.2% of patients were anticellular antibody negative, and 1.5% had an isolated CMP. The prevalence of anticellular antibody–negative systemic lupus erythematosus will likely decrease as emerging nomenclature guidelines recommend that non-nuclear patterns should also be reported as a positive ANA.

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

The Montreal General Hospital Lupus Clinic is supported by the Singer Family Fund for Lupus Research. The Hopkins Lupus Cohort is supported by the NIH (grants AR-043727 and AR-069572). Dr. Hanly's work was supported by the Canadian Institutes of Health Research (grant MOP-88526). Dr. Gordon's work was supported by Lupus UK, Sandwell and West Birmingham Hospitals NHS Trust, and the NIHR/Wellcome Trust Clinical Research Facility in Birmingham. Dr. Bae's work was supported by the Korea Healthcare technology R&D project, Ministry for Health and Welfare (A120404). Dr. Isenberg's and Dr. Rahman's work was supported by the NIHR University College London Hospitals Biomedical Research Centre. Dr. Bruce's work was supported by Arthritis Research UK, the NIHR Manchester Biomedical Research Centre, and the NIHR/Wellcome Trust Manchester Clinical Research Facility. Dr. Dooley's work was supported by the NIH (grant RR-00046). Dr. Ramsey-Goldman's work was supported by the NIH (grants 1U54-TR-001353 [formerly 8UL1-TR-000150], UL-1RR-025741, K24-AR-02318, and P60AR064464 [formerly P60-AR-48098]). Dr. Ruiz-Irastorza's work was supported by the Department of Education, Universities, and Research of the Basque Government.

[1]May Y. Choi, MD, Ann E. Clarke, MD, MSc, Marvin J. Fritzler, PhD, MD: University of Calgary, Cumming School of Medicine, Calgary, Alberta, Canada; [2]Yvan St. Pierre, MSc, Sasha Bernatsky, MD, PhD: McGill University Health Centre, Montreal, Quebec, Canada; [3]John G. Hanly, MD: Queen Elizabeth II Health Sciences Centre and Dalhousie University, Halifax, Nova Scotia, Canada; [4]Murray B. Urowitz, MD, Dafna D. Gladman, MD: Toronto Western Hospital, University of Toronto, Toronto, Ontario, Canada; [5]Juanita Romero-Diaz, MD, MS: Instituto Nacional de Ciencias Médicas y Nutrición, Mexico City, Mexico; [6]Caroline Gordon, MD: University of Birmingham, Birmingham, UK; [7]Sang-Cheol Bae, MD, PhD, MPH: Hanyang University Hospital for Rheumatic Diseases, Seoul, Korea; [8]Daniel J. Wallace, MD: Cedars-Sinai/David Geffen School of Medicine at University of California

SIGNIFICANCE & INNOVATIONS

- This is the first study to examine the prevalence of anticellular antibody negativity defined as the absence of any intracellular indirect immunofluorescence staining in a large systemic lupus erythematosus (SLE) cohort at inception.
- In 1,137 patients newly diagnosed with SLE, 6.2% (71) were anticellular antibody negative and 1.5% (17) had an isolated cytoplasmic and mitotic pattern (CMP). Therefore, among these 88 patients, 20% (17) would be misclassified as antinuclear antibody negative under the traditional definition, when in fact they have antibodies directed against a variety of CMP targets.
- Anticellular antibody negativity was more likely in patients who were older, were of white race/ethnicity, or were receiving high-dose glucocorticoids, and it was less likely in those patients using immunosuppressants. Longitudinal data are needed to assess how anticellular antibody status is influenced by the disease course and therapy.

INTRODUCTION

Autoantibodies directed against nuclear autoantigens (antinuclear antibodies [ANAs]) and other intracellular autoantigens are a serologic hallmark of systemic lupus erythematosus (SLE) and other ANA-associated rheumatic diseases (AARD), such as systemic sclerosis, mixed connective tissue disease, and Sjögren's syndrome (1–3). ANAs are widely regarded as an important classification criterion of SLE, as officially recognized by both the American College of Rheumatology (ACR) (4) and the Systemic Lupus International Collaborating Clinics (SLICC) (5). ANA positivity is traditionally defined as the presence of an indirect immunofluorescence (IIF) staining pattern localized to the nucleus, while isolated cytoplasmic and mitotic cell patterns (CMPs), although staining positive by IIF, often are not reported or classified as ANA-positive and are not included in the ANA test reports by some laboratories. The International Consensus on ANA Patterns (ICAP) Committee has debated a suggestion that CMPs should be included in ANA result reports and that there should be a change in terminology to anticellular antibodies, because CMPs are increasingly recognized as clinically relevant (6–8) and have implications for the diagnosis and classification of AARDs (9). For instance, antiribosomal P proteins are highly specific for SLE and are associated with certain clinical and serologic SLE features (10,11), but antiribosomal P antibodies may be reported as ANA IIF negative, because their prototypical staining pattern is localized to the cytoplasm (12). Therefore, ANA IIF exhibits limited sensitivity for the detection of antiribosomal P antibodies (13). After debate, however, the ICAP recognized that current disease classification criteria are predicated on a more traditional definition of ANA and that jurisdictional precedents (i.e., reimbursement fee structures) only allow reporting of classical ANA results, so the ICAP concluded that the reclassification of ANA to include CMPs should be delayed (9).

Inclusion of these additional CMPs in the ANA test results would likely help minimize misclassification of SLE patients, and the prevalence of anticellular antibody–negative SLE (i.e., the complete absence of any intracellular IIF staining patterns) will

Los Angeles; [9]Joan T. Merrill, MD: Oklahoma Medical Research Foundation, Oklahoma City; [10]David A. Isenberg, MD, Anisur Rahman, PhD: University College London, London, UK; [11]Ellen M. Ginzler, MD, MPH: State University of New York Downstate Medical Center, Brooklyn; [12]Michelle Petri, MD, MPH: Johns Hopkins University School of Medicine, Baltimore, Maryland; [13]Ian N. Bruce, MD: Arthritis Research UK, University of Manchester, NIHR Manchester Biomedical Research Centre, Central Manchester University Hospitals NHS Foundation Trust, and Manchester Academic Health Science Centre, Manchester, UK; [14]Mary A. Dooley, MD, MPH: University of North Carolina, Chapel Hill; [15]Paul R. Fortin, MD, MPH: CHU de Québec–Université Laval, Quebec City, Quebec, Canada; [16]Jorge Sanchez-Guerrero, MD, MS: Mount Sinai Hospital and University Health Network, University of Toronto, Toronto, Ontario, Canada; [17]Kristjan Steinsson, MD, PhD: Landspitali University Hospital, Reykjavik, Iceland; [18]Rosalind Ramsey-Goldman, MD, DrPh: Northwestern University and Feinberg School of Medicine, Chicago, Illinois; [19]Munther A. Khamashta, MD, PhD: St Thomas' Hospital and King's College, London School of Medicine, London, UK; [20]Cynthia Aranow, MD: Feinstein Institute for Medical Research, Manhasset, New York; [21]Graciela S. Alarcón, MD, MPH: University of Alabama at Birmingham; [22]Susan Manzi, MD, MPH: Allegheny Health Network, Pittsburgh Pennsylvania; [23]Ola Nived, MD, PhD: University Hospital Lund, Lund, Sweden; [24]Asad A. Zoma, MBChB: Hairmyres Hospital, East Kilbride, Scotland, UK; [25]Ronald F. van Vollenhoven, MD, PhD: University of Amsterdam, Amsterdam, The Netherlands; [26]Manuel Ramos-Casals, MD, PhD: Hospital Clínic, Barcelona, Spain; [27]Guillermo Ruiz-Irastorza, MD, PhD: Hospital Universitario Cruces, University of the Basque Country, Barakaldo, Spain; [28]S. Sam Lim, MD, MPH: Emory University School of Medicine, Atlanta, Georgia; [29]Kenneth C. Kalunian, MD: University of California San Diego School of Medicine, La Jolla; [30]Murat Inanc, MD: Istanbul University, Istanbul, Turkey; [31]Diane L. Kamen, MD, MSCR: Medical University of South Carolina, Charleston; [32]Christine A. Peschken, MD, MSc: University of Manitoba, Winnipeg, Manitoba, Canada; [33]Soren Jacobsen, MD, DMSc: Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; [34]Anca Askanase, MD, MPH: Hospital for Joint Diseases and New York University, New York; [35]Thomas Stoll, MD: Kantousspital, Schaffhausen, Switzerland; [36]Jill Buyon, MD: New York University School of Medicine, New York; [37]Michael Mahler, PhD: Inova Diagnostics Inc., San Diego, California.

Dr. Clarke has received consulting fees from AstraZeneca/MedImmune, Bristol-Myers Squibb, and Exagen Diagnostics (less than 10,000 each). Dr. Gordon has received consulting fees, speaking fees, and/or honoraria from Eli Lilly, UCB, GlaxoSmithKline, Merck Serono, and Bristol-Myers Squibb (less than $10,000 each). Dr. Bruce has received consulting fees, speaking fees, and/or honoraria from Eli Lilly, UCB, Roche, Merck Serono, and MedImmune (less than $10,000 each) and grants from UCB, Genzyme Sanofi, and GlaxoSmithKline. Dr. Gladman has received consulting fees, speaking fees, and/or honoraria from GlaxoSmithKline (less than $10,000). Dr. Kalunian has received grants from UCB, Human Genome Sciences/GlaxoSmithKline, Takeda, Ablynx, Bristol-Myers Squibb, Pfizer, and Kyowa Hakko Kirin, and has received consulting fees from Exagen Diagnostics, Genentech, Eli Lilly, Bristol-Myers Squibb, and Anthera (less than $10,000 each). Dr. Fritzler has received consulting fees from Inova Diagnostics, Werfen International, and Alexion Canada (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Ann E. Clarke, MD, MSc, Division of Rheumatology, Cumming School of Medicine, University of Calgary, 3330 Hospital Drive NW, Calgary, Alberta, Canada T2N 4N1. E-mail: aeclarke@ucalgary.ca.

Submitted for publication November 29, 2017; accepted in revised form July 17, 2018.

accordingly be decreased (12). The exact prevalence of ANA-negative SLE using the traditional definition (i.e., the absence of IIF staining localized only to the nucleus) has been reported to range from 1% to 28% (14–17). A recent systematic review and meta-analysis of 64 studies showed that an ANA of 1:80 was highly sensitive at 97.8% (95% confidence interval [95% CI] 96.8, 98.5), but not specific (74.7% [95% CI 66.7, 81.3]) for SLE (18). Pisetsky et al (14) compared different commercial ANA assays, including the HEp-2000 substrate, in an established SLE cohort and demonstrated significant variation in frequencies of ANA positivity that ranged from 77.7% to 95.1%. In studies to date, there are several factors (laboratory performance, study design, and clinical factors) that could influence the ANA results. Laboratory performance factors could include the ANA kit selected, the definition of an ANA (i.e., whether it includes isolated CMPs), the ANA IIF screening dilution chosen, and technical errors such as variable substrate sensitivity and specificity for the detection of autoantibodies directed against DNA, SSA/Ro 60, Ro 52/tripartite motif 21 (TRIM21), ribosomal P, and other intracellular autoantigens. The prevalence of ANA positivity is also likely impacted by whether it is measured cross-sectionally or longitudinally along the disease course. ANA status is also potentially influenced by the level of disease activity, concurrent treatment with glucocorticoids and other immune-modulating drugs, and persistent proteinuria leading to renal immunoglobulin loss (2,9,15,19,20).

The purpose of this study was to examine the prevalence of anticellular antibody negativity (no intracellular IIF pattern) in a large international SLE inception cohort and to assess demographic, clinical, or other autoantibody characteristics associated with these redefined subgroups of patients with SLE.

MATERIALS AND METHODS

Study design and setting. This study was conducted using data and patient sera collected by SLICC, a network of 53 investigators in 43 academic medical centers in 16 countries (21–23). Between 1999 and 2011, SLICC investigators enrolled patients fulfilling the ACR classification criteria for definite SLE (4) within 15 months of diagnosis. The study was approved by the institutional review board at each participating site and complied with the Helsinki Declaration.

Anticellular antibody by IIF assay. The earliest available serum at enrollment from each patient was analyzed at the Mitogen Advanced Diagnostic Laboratory (University of Calgary). Aliquots of the anonymized SLE sera obtained from the central SLICC biobank were stored at –80° C until required for immunoassays. The IIF immunoassay was initially performed at a screening dilution of 1:160 (24) using HEp-2000 cell substrate (ImmunoConcepts) and fluorescein isothiocyanate conjugated to antihuman IgG (H + L) according to the manufacturer's instructions. IIF results were read by technologists with >10 years of experience at Mitogen Advanced Diagnostics, as previously described (25). The HEp-2000 substrate had been transfected with the SSA/Ro 60 complementary DNA, which was then overexpressed in the cells, as an approach to intentionally increase the detection of anti-SSA/Ro 60 autoantibodies and thereby increasing the sensitivity of this substrate (25,26). The results obtained at a single center (Mitogen) were used for the ANA analysis in this study, because the ANA analyses performed at each regional site had a wide variation in testing parameters (date of test performance, serum screening dilutions, test kits and protocols, microscopes, readers, etc.) and thus were not comparable across sites. For the purposes of this study, patients were divided into 3 groups depending on their anticellular antibody IIF patterns: ANA positive (the presence of nuclear IIF or mixed nuclear and CMP staining), anticellular antibody negative (no intracellular staining detected), and isolated CMP staining.

Detection of anti–double-stranded DNA (anti-dsDNA) and other autoantibodies. All samples were also tested for the presence of anti-dsDNA antibodies by chemiluminescence immunoassay (QUANTA Flash, Inova Diagnostics) as previously described (27) using a cutoff of 70 IU/ml, established in accord with the SLICC classification criterion for anti-dsDNA positivity, which requires that the cutoff for the anti-dsDNA antibody level be above the laboratory reference range (or >2-fold the reference range if tested by enzyme-linked immunosorbent assay) (5).

Antibodies to proliferating cell nuclear antigen, ribosomal P, recombinant Ro 52/TRIM21, native SSA/Ro 60, SSB/La, Sm, and U1 RNP were detected using the extractable nuclear antigen FIDIS Connective Profile, kit 13 addressable laser bead immunoassay (TheraDiag) on a Luminex 200 flow luminometer, according to the manufacturer's instructions, and using MLX-Booster software. Other autoantibodies, such as IgG anticardiolipin, IgG anti–β2-glycoprotein 1, and lupus anticoagulant, were measured in a central laboratory as previously described (28). ANA IIF patterns were classified according to the new ICAP standards (http://www.anapatterns.org/index.php) (9).

Clinically defined samples. Demographic and clinical data were collected at enrollment and included the age at diagnosis, sex, postsecondary education, disease duration, race/ethnicity, smoking status, alcohol use, hypertension, nephritis at enrollment, proteinuria at enrollment (≥3 grams/day), ACR classification criteria fulfilled (total and individual), Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) global score and organ system scores, and medication use (glucocorticoids, high-dose glucocorticoids [any pulse steroid or prednisone ≥40 mg/day], antimalarials, and immunosuppressive agents, including biologics) at or prior to cohort enrollment (see Supplementary Table 1, available on the *Arthritis Care &*

Table 1. Baseline demographic, clinical, and autoantibody profiles of antinuclear antibody (ANA)–positive (presence of any nuclear indirect immunofluorescence [IIF] pattern), anticellular antibody (ACA)–negative (no IIF pattern), and isolated cytoplasmic/mitotic (CMP) groups*

	ANA+ (n = 1,049)	ACA– (n = 71)	Isolated CMP (n = 17)†	ANA+ and ACA–	ANA+ and CMP	ACA– and CMP
Demographics						
Age at diagnosis, years, mean	34.7‡	40.9‡	35.8	−6.2 (−9.4, −2.9)	−1 (−7.5, 5.4)	5.1 (−2.4, 12.7)
Female, %	89.7	90.1	100	−0.4 (−7.6, 6.7)	−10.3 (−24.8, 4.2)	−9.9 (−24.2, 4.5)
Postsecondary education, %	66.7‡	76.1§	31.3†*	−9.5 (−20.1, 1.2)	35.4 (12.5, 58.3)	44.9 (20, 69.8)
Disease duration, years, mean	0.47	0.42	0.35	0.05 (−0.03, 0.14)	0.12 (−0.05, 0.29)	0.07 (−0.12, 0.25)
Race/ethnicity, %						
Asian	23.2‡	4.2‡	11.8	19 (13.7, 24.3)	11.5 (−4.1, 27)	−7.5 (−23.6, 8.5)
African descendant	16.2‡	7.0‡	5.9	9.2 (2.8, 15.5)	10.3 (−1.1, 21.7)	1.2 (−11.5, 13.8)
Hispanic	3.4	2.8	0	0.5 (−3.5, 4.5)	3.4 (−5.2, 11.9)	2.8 (−1, 6.7)
White	52.4†*	84.5‡	76.5§	−32.1 (−41.1, −23.2)	−24.1 (−44.5, −3.7)	8 (−13.8, 29.9)
Other	4.8	1.4	5.9	3.4 (−1.6, 8.4)	−1.1 (−12.3, 10.2)	−4.5 (−16, 7)
Smoking status, %						
Current smoker	15.1	21.9	18.8	−6.8 (−17.1, 3.6)	−3.7 (−22.9, 15.6)	3.1 (−18.5, 24.8)
Former smoker	21.1	26.6	25	−5.5 (−16.6, 5.6)	−3.9 (−25.3, 17.5)	1.6 (−22.3, 25.4)
High alcohol use, %	1.5	1.5	0	0 (−3, 3)	1.5 (−4.6, 7.5)	1.5 (−4.7, 7.6)
Hypertension, %	32.6‡	29.6§	58.8†*	3 (−8, 14)	−26.2 (−49.8, −2.7)	−29.2 (−54.9, −3.6)
Nephritis at enrollment, %	28.7	26.6	50	2.1 (−9, 13.3)	−21.3 (−46, 3.4)	−23.4 (−50.2, 3.3)
Proteinuria at enrollment, %	4.5	3.3	12.5	1.2 (−3.4, 5.9)	−8 (−24.3, 8.3)	−9.2 (−26, 7.6)
No. of ACR criteria, mean	4.8	4.7	4.7	0.1 (−0.1, 0.4)	0.1 (−0.4, 0.6)	0 (−0.5, 0.5)
SLEDAI-2K score, mean	5.4‡	4.1‡	5.4	1.3 (0, 2.6)	0 (−2.7, 2.6)	−1.3 (−3.8, 1.1)
Neurological	0.3	0.3	0	−0.1 (−0.5, 0.3)	0.3 (−0.5, 1)	0.3 (−0.5, 1.2)
Mucocutaneous	1.1	1	1.3	0.1 (−0.4, 0.5)	−0.1 (−1.1, 0.8)	−0.2 (−1, 0.6)
Musculoskeletal	0.8	0.7	1.3	0.1 (−0.3, 0.5)	−0.4 (−1.3, 0.4)	−0.5 (−1.4, 0.4)
Renal	1.4	0.7	1.8	0.7 (−0.1, 1.5)	−0.4 (−2, 1.2)	−1.1 (−2.5, 0.4)
Serositis	0.1	0.1	0	0 (−0.1, 0.1)	0.1 (−0.1, 0.3)	0.1 (−0.2, 0.4)
Constitutional	0	0	0	0 (0, 0.1)	0 (−0.1, 0.1)	0 (0, 0.1)
Immunologic	1.6‡	1.1‡	1.1	0.5 (0.1, 0.9)	0.5 (−0.4, 1.3)	0 (−0.9, 0.8)
Hematologic	0.1	0	0	0.1 (0, 0.1)	0.1 (−0.1, 0.3)	0 (−0.1, 0.1)
Medications, % ever used						
Glucocorticoids	80.6	74.6	82.4	6 (−4.4, 16.4)	−1.7 (−20, 16.6)	−7.7 (−28.5, 13.1)
High-dose glucocorticoids	42.3	46.5	58.8	−4.2 (−16.1, 7.8)	−16.5 (−40.1, 7.1)	−12.3 (−38.5, 13.8)
Antimalarials	74.3	69	52.9	5.2 (−5.8, 16.3)	21.3 (−2.6, 45.2)	16.1 (−10, 42.1)
Immunosuppressants	43.7‡	23.9†*	58.8§	19.7 (9.3, 30.1)	−15.2 (−38.7, 8.4)	−34.9 (−60.3, −9.5)
Autoantibodies, %						
dsDNA	28.4‡	11.3‡	17.7	17.2 (9.3, 25)	10.8 (−7.5, 29.1)	−6.4 (−25.9, 13.2)
PCNA	7.3	1.4	11.8	5.9 (−0.2, 12)	−4.4 (−19.8, 11)	−10.4 (−25.9, 5.2)
Ribosomal P	16.1‡	5.6‡	11.8	10.5 (4.7, 16.3)	4.3 (−11.1, 19.8)	−6.1 (−22.4, 10.1)

(continued)

Table 1. (*Cont'd*)

	ANA+ (n = 1,049)	ACA– (n = 71)	Isolated CMP (n = 17)†	ANA+ and ACA–	ANA+ and CMP	ACA– and CMP
Ro 52/TRIM21	35.9‡	21.1‡	23.5	14.8 (4.9, 24.7)	12.4 (–8, 32.8)	–2.4 (–24.7, 19.9)
SSA/Ro 60	47.3‡	22.5‡	29.4	24.7 (14.6, 34.9)	17.9 (–4, 39.7)	–6.9 (–30.6, 16.9)
SSB/La	15.9‡	5.6‡	11.8	10.3 (4.5, 16.1)	4.2 (–11.3, 19.6)	–6.1 (–22.4, 10.1)
Sm	24.7‡	5.7‡	11.8	19 (12.9, 25)	12.9 (–2.6, 28.5)	–6.1 (–22.3, 10.2)
U1 RNP	32.4‡	11.3‡	11.8	21.1 (13.3, 29)	20.6 (–1.7, 43)	–0.5 (–17.5, 16.5)
Lupus anticoagulant	20.8	20.6	6.7	0.1 (–10.2, 10.5)	14.1 (–6.5, 34.7)	14 (–2.1, 30.1)
Anticardiolipin	12.6	11.1	12.5	1.5 (–6.6, 9.5)	0.1 (–16.3, 16.4)	–1.4 (–19.4, 16.6)
Anti–β2-glycoprotein 1	15	15.9	12.5	0.8 (–10.1, 8.5)	2.5 (–13.8, 18.9)	3.4 (–15.2, 21.9)

* Values are the difference (95% confidence interval) unless indicated otherwise. ACR = American College of Rheumatology; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; dsDNA = double-stranded DNA; PCNA = proliferating cell nuclear antigen; TRIM21 = tripartite motif 21; SSA = Sjögren's syndrome antigen A; SSB = Sjögren's syndrome antigen B; RNP = ribonucleoprotein.
† Some predictors had a small number of missing values. When these occurred, the observations were excluded from the relevant analysis. In particular, for the small group of patients with an isolated CMP, the data included 1 missing value for education, smoking status, nephritis, proteinuria, no. of ACR criteria, anticardiolipin, and anti–β2-glycoprotein 1, and 2 missing values for high alcohol use and lupus anticoagulant.
‡ Values with the same footnote symbol are significantly different from each other.
§ Values with the same footnote symbol are significantly different from each other.
†* Values are significantly different from ‡ and §, but ‡ and § are not different from each other.

Research web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23712/abstract).

Statistical analysis. Statistical analysis was performed using Stata software, version 14.1. A 3-way comparison was performed between patients who were ANA positive versus anticellular antibody negative versus having an isolated CMP. Univariable and multivariable logistic regression analyses were used to examine potential predictors of the odds of being anticellular antibody negative or ANA positive or having an isolated CMP. As a secondary analysis, 3 additional univariable and multivariable logistic regressions were performed: anticellular antibody negative versus ANA positive, isolated CMP positive versus ANA positive, and isolated CMP positive versus anticellular antibody negative.

Potential univariable predictors included the demographic, clinical, and serologic data listed above. For the most informative multivariable model, only statistically significant predictors at the 95% CI were included, after eliminating all other potential predictors individually, starting with the least likely to be associated with the outcome.

RESULTS

Cohort demographic, clinical, and serologic characteristics. The baseline demographic, clinical, and serologic characteristics of the 3 serologic groups (ANA positive, anticellular antibody negative, and isolated CMP positive) are shown in Table 1. Overall, 1,137 patients had sera available; their mean ± SD age at diagnosis was 35.1 ± 13.5 years (median 33 years), 89.9% were female, 66.7% (724 of 1,085) had obtained postsecondary education, the mean ± SD disease duration was 0.46 ± 0.35 years, and 45.2% (511 of 1,130) were not of white race/ethnicity. A total of 312 of 1,084 (29%) of the cohort had lupus nephritis at enrollment, the mean ± SD global SLEDAI-2K score was 5.3 ± 5.3, and 80.3% (913 of 1,137) had a history (either at or prior to enrollment) of glucocorticoid use, 73.6% (837 of 1,137) of antimalarial use, and 42.7% (485 of 1,137) of immunosuppressant use, including 4 patients who had received biologics (rituximab only).

Nuclear and CMP anticellular antibody IIF patterns. The distribution of patients based on IIF staining patterns and specificities is shown in Supplementary Figures 1 and 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr. 23712/abstract. Overall, 1,049 of 1,137 patients (92.3%) were ANA positive, which included 877 isolated nuclear (77.1%) and 172 mixed nuclear patterns and CMPs (15.1%). A total 71 of 1,137 patients (6.2%) were anticellular antibody negative (i.e., with no detectable IIF staining), and 17 of 1,137 patients (1.5%) had an isolated CMP. Therefore, 7.7% of patients were either anticellular antibody negative or had an isolated CMP. Isolated CMPs and their related ICAP designations included 41.2% (7 of 17) cytoplasmic dense fine speckled (ANA pattern AC-19), 23.5% (4 of 17) cytoplasmic fine speckled (AC-20), 5.9% (1 of 17) cytoplasmic discrete dots (AC-18), 5.9% (1 of 17) mitotic chromosomal envelope (AC-28), and 23.5% (4 of 17) mixed CMP (ICAP does not have a pattern designation for mixed patterns at this time).

Table 2. Univariable and multivariable analysis of demographic, clinical, and serologic profiles of anticellular antibody–negative versus antinuclear antibody–positive combined with isolated cytoplasmic/mitotic patterns*

	Univariate	Multivariate
Demographics		
Age at diagnosis	1.03 (1.01, 1.05)†	1.02 (1.00, 1.04)†
Female	1.03 (0.46, 2.31)	–
Postsecondary education	1.63 (0.92, 2.91)	–
Disease duration	0.66 (0.32, 1.34)	–
Race/ethnicity		
Asian	0.15 (0.05, 0.47)†	–
African descendant	0.40 (0.16, 1.00)†	–
Hispanic	0.85 (0.20, 3.60)	–
White	4.88 (2.54, 9.38)†	3.53 (1.77, 7.03)†
Other	0.28 (0.38, 2.07)	–
Smoking status		
Current smoker	1.57 (0.85, 2.90)	–
Former smoker	1.35 (0.76, 2.39)	–
High alcohol use	1.04 (0.14, 8.00)	–
Hypertension	0.85 (0.50, 1.44)	–
Nephritis at enrollment	0.88 (0.50, 1.56)	–
Proteinuria at enrollment	0.70 (0.17, 2.95)	–
No. of ACR criteria	0.89 (0.69, 1.14)	–
SLEDAI-2K score	0.94 (0.89, 1.00)†	–
Neurological	1.03 (0.90, 1.18)	–
Mucocutaneous	0.98 (0.85, 1.12)	–
Musculoskeletal	0.97 (0.84, 1.14)	–
Renal	0.91 (0.81, 1.01)	–
Serositis	1.12 (0.72, 1.75)	–
Constitutional	0.37 (0.05, 2.73)	–
Immunologic	0.82 (0.69, 0.96)†	–
Hematologic	0.41 (0.13, 1.27)	–
Medications, ever using		
Glucocorticoids	0.71 (0.40, 1.23)	–
High-dose glucocorticoids	1.17 (0.72, 1.90)	2.39 (1.39, 4.12)†
Antimalarials	0.79 (0.47, 1.32)	–
Immunosuppressants/biologics	0.40 (0.23, 0.70)†	0.35 (0.19, 0.64)†
Autoantibodies		
dsDNA	0.32 (0.15, 0.68)†	–
PCNA	0.18 (0.02, 1.30)	–
Ribosomal P	0.31 (0.11, 0.87)†	–
Ro52/TRIM21	0.48 (0.27, 0.86)†	–
SSA/Ro 60	0.33 (0.19, 0.58)†	0.41 (0.23, 0.74)†
SSB/La	0.32 (0.11, 0.88)†	–
Sm	0.19 (0.07, 0.52)†	–
U1 RNP	0.27 (0.13, 0.57)†	0.43 (0.20, 0.93)†
Lupus anticoagulant	1.00 (0.54, 1.88)	–
Anticardiolipin	0.87 (0.39, 1.95)	–
Anti–β2-glycoprotein 1	1.07 (0.53, 2.15)	–

* Values are the odds ratio (95% confidence interval). ACR = American College of Rheumatology; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; dsDNA = double-stranded DNA; PCNA = proliferating cell nuclear antigen; TRIM21 = tripartite motif 21; SSA = Sjögren's syndrome antigen A; SSB = Sjögren's syndrome antigen B; RNP = ribonucleoprotein.
† Statistically significant.

Comparison of isolated CMP positive with ANA positive and anticellular antibody negative and comparison of ANA positive and anticellular antibody negative. Patients with isolated CMPs were not clinically or serologically different from ANA-positive or anticellular antibody–negative patients for most variables (Table 1). In contrast, ANA-positive patients were markedly different from anticellular antibody–negative patients in terms of age at diagnosis (34.7 versus 40.9 years), race/ethnicity (a higher proportion of Asians and African descendants, but fewer patients of white race/ethnicity), disease activity (SLEDAI-2K score 5.4 versus 4.1), use of immunosuppressants at or prior to enrollment (43.7% versus 23.9%), and frequency of SLE-related autoantibodies. Interestingly, despite a negative anticellular antibody IIF on HEp-2000 substrate, some SLE-related autoantibodies were still detected, notably anti-dsDNA by chemiluminescence immunoassay (11.3%), and anti–Ro 52/TRIM21 (21.1%), anti-SSA/Ro 60 (22.5%), and anti–U1 RNP by addressable laser bead immunoassay (11.3%).

Multivariable analysis of anticellular antibody–negative patients versus ANA-positive patients combined with isolated CMP–positive patients. Because the isolated CMP–positive group did not differ from the ANA-positive or anticellular antibody–negative groups for most variables, we chose to combine the isolated CMP–positive with the ANA-positive groups for the primary multivariable analysis. In that analysis (Table 2), patients who were older (odds ratio [OR] per year 1.02 [95% CI 1.00, 1.04]), of white race/ethnicity (OR 3.53 [95% CI 1.77, 7.03]), or receiving high doses of glucocorticoids at or prior to enrollment (OR 2.39 [95% CI 1.39, 4.12]) were more likely to be anticellular antibody negative. Patients who were

receiving immunosuppressants at or prior to enrollment (OR 0.35 [95% CI 0.19, 0.64]) or who had anti-SSA/Ro 60 (OR 0.41 [95% CI 0.23, 0.74]) or anti–U1 RNP (OR 0.43 [95% CI 0.20, 0.93]) were less likely to be anticellular antibody negative.

Multivariable analysis of anticellular antibody negative versus ANA positive, isolated CMP positive versus ANA positive, and isolated CMP positive versus anticellular antibody negative. In the secondary multivariable analysis comparing the odds of being anticellular antibody negative versus being ANA positive, the predictors were identical to those in the multivariable analysis of the anticellular antibody–negative patients versus the ANA-positive patients combined with the isolated CMP–positive patients (see Supplementary Table 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23712/abstract).

In secondary multivariable analyses comparing the odds of being isolated CMP positive versus ANA positive or being isolated CMP positive versus anticellular antibody negative, patients who had not attained postsecondary education or who were hypertensive were more likely to be isolated CMP positive (see Supplementary Tables 3 and 4, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23712/abstract).

DISCUSSION

To our knowledge, this is the first study of ANA IIF in a large SLE inception cohort redefining negative ANA as the absence of any intracellular IIF staining, which we referred to as anticellular antibody negative. Traditionally, ANA negative referred only to the absence of any IIF staining localized to the nucleus. This definition is an important consideration, especially for AARDs such as SLE, where the ANA test has a central role in establishing the diagnosis. The need to clarify this issue is exigent, because the topic is currently under international review (9), and the state of nomenclature uncertainty is the source of variability in ANA definitions and related clinical reports by different laboratories. Some laboratories do not report CMP staining, whereas others provide 2 reports: one that specifies nuclear staining patterns and titers and another that indicates whether CMP staining is present. In the broader definition of ANA test results, the inclusive definition of ANA and CMP together is more accurately referred to as anticellular antibody (1,9,29). However, because the ANA rubric is embedded in historical and scientific literature, the anticellular antibody terminology is held in abeyance until wider consensus and clinician education is achieved (3,9,29). The results of the current study provide some insight into the potential diagnostic and clinical implications for patients with SLE as a consequence of changing the definition of ANA to the wider anticellular antibody paradigm.

In our analysis of patients enrolled in the SLICC inception cohort, we demonstrated that the prevalence of ANA-negative SLE by routine IIF on a HEp-2000 substrate at a serum dilution of 1:160 was 7.7% (88 of 1,137). However, if isolated CMPs (17 of 1,137 [1.5%]) were subsequently excluded from the ANA-negative pool of 88 patients, the prevalence of anticellular antibody–negative SLE would decrease to 6.2% (71 of 1,137). Accordingly, among these 88 ANA-negative patients, nearly 1 in 5 is misclassified as ANA-negative, when they in fact have antibodies directed against a variety of CMP targets (8). Therefore, clinicians should be aware of which approach their laboratory employs for routine ANA IIF testing, because some patients with a high pretest probability of an AARD may have a negative ANA test, when in fact the test should be regarded as positive if CMP staining is present.

In our study, SLE patients with an isolated CMP could not be readily differentiated from ANA-positive and anticellular antibody–negative patients based on clinical or conventional serologic features. These results must be interpreted cautiously, however, given the small sample size (n = 17) of patients with an isolated CMP. In contrast, there were many differences between the anticellular antibody–negative and ANA-positive patients, consistent with the current literature indicating that ANA-negative SLE follows a more benign clinical course characterized by photosensitive skin rashes and arthritis (19,30,31). We demonstrated in the SLICC cohort that anticellular antibody–negative patients were older (age 40.9 versus 34.7 years) and that a higher proportion were of white race/ethnicity (84.5% versus 52.4%). Further, anticellular antibody–negative patients compared to ANA-positive patients had a lower global SLEDAI-2K score (4.1 versus 5.4), less frequent use of immunosuppressants at or prior to enrollment (23.9% versus 43.7%), and a decreased likelihood of having multiple SLE-associated autoantibodies, including anti-dsDNA (11.3% versus 28.4%). These observations likely relate to earlier onset of more aggressive, severe disease in nonwhite patients, who tend to be ANA-positive, corroborating previous studies demonstrating higher disease activity in nonwhite patients with SLE (32,33).

When the anticellular antibody–negative patients were compared to the isolated CMP–positive combined with the ANA-positive patients, all the above observations regarding anticellular antibody–negative versus ANA-positive patients persisted in the univariable analysis. However, in the multivariable analysis, slight differences were observed. Older age and white race/ethnicity remained associated with a greater likelihood of being anticellular antibody negative, and high-dose glucocorticoids now became associated with a greater likelihood of being anticellular antibody negative; immunosuppressant medications (at or prior to enrollment) and certain autoantibodies remained associated with a lower likelihood of being anticellular antibody negative. Our finding that high-dose glucocorticoids are associated with a higher likelihood of anticellular antibody negativity may be attributable to glucocorticoids influencing ANA status (34). However, this possibility is merely speculation, because we have no data on

ANA status prior to the baseline assessment. Patients taking other types of immunosuppressants (i.e., methotrexate, azathioprine, mycophenolate mofetil) were less likely to be anticellular antibody negative, perhaps due to a different effect on B cell responses (35,36). Further, immunosuppressants are potentially a proxy for elements of disease activity that are not measured through the other clinical variables included in the regression. Interestingly, in univariable analysis, all 4 patients treated with rituximab (data not shown) were anticellular antibody negative (OR 11.54 [95% CI 2.00, 66.74]). As suggested in a review by Cross et al (15), previous literature on ANA-negative SLE has been poor at documenting concurrent therapies. In that review and commentary, only 5 of 164 patients (3%) had data on medications during ANA testing. This lack of documentation highlights the need to review concurrent medications and consider other known confounders, such as proteinuria, as we have done.

The ANA status of our cohort was tested on the HEp-2000 substrate, which has been engineered to intentionally increase the detection of anti-SSA/Ro 60, thereby lowering the prevalence of ANA-negative SLE (25,26,37,38). Up to two-thirds of patients with mild SLE and persistently negative ANA tested on rodent liver substrate have been serologically linked to SLE due to precipitating autoantibodies to SSA/Ro 60 (31). These findings are particularly relevant to the clinical subset of SLE that has subacute cutaneous SLE and/or features of secondary Sjögren's syndrome (39). However, even with the technical improvements, such as HEp-2000 substrates, our study and others (20) indicate a persistent gap in autoantibody detection by HEp-2 substrates, which in the current study included anti-SSA/Ro 60 and even anti-dsDNA. For example, 22.5% of the anticellular antibody–negative SLE patients in our study still had anti-SSA/Ro 60 antibodies using extractable nuclear antigen testing; 11.3% of our anticellular antibody–negative patients had anti-dsDNA by chemiluminescence immunoassay. Our observations are consistent with a recent study showing that there is significant lack of agreement between positive results using a conventional multiplex array technology and the IIF on HEp-2 cells (40).

Significant variation in the frequencies of positive ANA in well-characterized SLE patients has been reported (15,17,18,41); some of this variation relates to the performance of different HEp-2 assay kits (14). In the current study, we used a serum dilution of 1:160 to maximize specificity of the test at the possible expense of sensitivity (24). When the IIF test was repeated at a serum dilution of 1:80 on 67 of 71 of the available anticellular antibody–negative samples, we observed that 17 of 67 (25.4%) became clearly positive for nuclear and/or CMP staining (detailed data not shown). A cross-sectional study showed that only 76% of unselected SLE sera had a positive ANA, but a relatively high serum dilution of 1:200 was used (16). Taken together, this finding suggests that newer multiplexed autoantigen array technologies might be considered in the future as a replacement for the ANA IIF.

The presence of anti-dsDNA in ANA-negative SLE patients has been reported by others (42,43). These patients were reported to have more severe complications, including nephritis (44), dystrophic calcification (45), or severe autoimmune neutropenia (46). Thus, the detection of anti-dsDNA antibodies even in ANA-negative cases is still important and may aid in risk assessment for clinical complications. Furthermore, the anti-dsDNA repertoire is diverse, such that there is no current anti-dsDNA assay that is able to detect all of the subpopulations of anti-dsDNA autoantibodies (47). Overall, the reports of anti-dsDNA–positive/ANA-negative sera found in the literature provide evidence that not all anti-dsDNA antibodies are detected on conventional HEp-2 substrates and that unique dsDNA epitopes may be missed by HEp-2 IIF screening tests.

Biomarkers such as autoantibodies and a variety of immune-related and inflammation-related molecules can appear years prior to clinical symptoms and/or the diagnosis of SLE and can accrue over time (40,48). Therefore, longitudinal studies are needed to evaluate the serologic status of anticellular antibody–negative and isolated CMP–positive patients over time and to evaluate whether the status varies with disease activity, damage accrual, therapeutic interventions, and/or specific substrate assays. Even among ANA-negative patients with lupus nephritis, the patient can take up to 10 years to seroconvert from ANA negative to positive (17,40). Some patients may only have detectable positive serologic results when there is uncontrolled disease activity due to loss of self-tolerance from chronic autoreactivity of T and B cells (17).

There are some limitations to our study. First, the similarities reported between CMP-positive and ANA-positive patients are likely confounded by the high proportion of ANA-positive patients also expressing a CMP (21.5%). Overall, approximately 17% of patients in the entire cohort expressed a CMP (189 of 1,137), but the majority (172 of 189 [91.0%]) were seen in conjunction with nuclear IIF patterns. As a result, the isolated CMP–positive group size (n = 17) was small, limiting the statistical power of our analysis. We also did not perform statistical correction for multiple comparisons, which is consistent with the exploratory and hypothesis-generating aspect of our study. Additionally, we evaluated ANA status only at disease inception, but we have the capacity with this inception cohort, where data and sera are collected longitudinally, to evaluate ANA status and factors influencing it over the disease course.

ACKNOWLEDGMENTS

The authors thank Ms. Haiyan Hou and Meifeng Zhang (Mitogen Advanced Diagnostics, University of Calgary) for technical 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. Clarke 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. Choi, Clarke, Wallace, Petri, Khamashta, Inanc, Askanase, Fritzler.

Acquisition of data. Choi, Clarke, Hanly, Urowitz, Romero-Diaz, Gordon, Bae, Bernatsky, Wallace, Merrill, Isenberg, Rahman, Ginzler, Petri, Bruce, Dooley, Fortin, Gladman, Sanchez-Guerrero, Steinsson, Ramsey-Goldman, Khamashta, Aranow, Alarcón, Manzi, Nived, van Vollenhoven, Ramos-Casals, Ruiz-Irastorza, Lim, Kalunian, Inanc, Kamen, Peschken, Jacobsen, Askanase, Stoll, Buyon, Fritzler.

Analysis and interpretation of data. Choi, Clarke, St. Pierre, Hanly, Romero-Diaz, Gordon, Bernatsky, Wallace, Petri, Bruce, Dooley, Sanchez-Guerrero, Steinsson, Khamashta, Aranow, Zoma, van Vollenhoven, Lim, Kalunian, Inanc, Kamen, Buyon, Mahler, Fritzler.

ADDITIONAL DISCLOSURE

Michael Mahler is an employee of Inova Diagnostics Inc.

REFERENCES

1. Agmon-Levin N, Damoiseaux J, Kallenberg C, Sack U, Witte T, Herold M, et al. International recommendations for the assessment of autoantibodies to cellular antigens referred to as anti-nuclear antibodies. Ann Rheum Dis 2014;73:17–23.
2. Mahler M, Meroni PL, Bossuyt X, Fritzler MJ. Current concepts and future directions for the assessment of autoantibodies to cellular antigens referred to as anti-nuclear antibodies. J Immunol Res 2014;2014:315179.
3. Pisetsky DS. Antinuclear antibody testing: misunderstood or misbegotten? Nat Rev Rheumatol 2017;13:495–502.
4. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.
5. Petri M, Orbai AM, Alarcón GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of systemic lupus international collaborating clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum 2012;64:2677–86.
6. Rattner JB, Mack G, Fritzler MJ. Autoantibodies to components of the mitotic spindle apparatus. Mol Biol Rep 1998;25:143–55.
7. Mack GJ, Rees J, Sandblom O, Balczon R, Fritzler MJ, Rattner JB. Autoantibodies to a group of centrosomal proteins in human autoimmune sera reactive with the centrosome. Arthritis Rheum 1998;41:551–8.
8. Stinton LM, Eystathioy T, Selak S, Chan EK, Fritzler MJ. Autoantibodies to protein transport and messenger RNA processing pathways: endosomes, lysosomes, Golgi complex, proteasomes, assemblyosomes, exosomes and GW bodies. Clin Immunol 2004;110:30–44.
9. Damoiseaux J, von Muhlen CA, Garcia-de la Torre I, Carballo OG, de Melo CW, Francescantonio PL, et al. International consensus on ANA patterns (ICAP): the bumpy road towards a consensus on reporting ANA results. Auto Immun Highlights 2016;7:1.
10. Mahler M, Kessenbrock K, Szmyrka M, Takasaki Y, Garcia-De La Torre I, Shoenfeld Y, et al. International multicenter evaluation of autoantibodies to ribosomal P proteins. Clin Vaccine Immunol 2006;13:77–83.
11. Shi ZR, Cao CX, Tan GZ, Wang L. The association of serum anti-ribosomal P antibody with clinical and serological disorders in systemic lupus erythematosus: a systematic review and meta-analysis. Lupus 2015;24:588–96.
12. Mahler M, Kessenbrock K, Raats J, Fritzler MJ. Technical and clinical evaluation of anti-ribosomal P protein immunoassays. J Clin Lab Anal 2004;18:215–23.
13. Mahler M, Ngo J, Schulte-Pelkum J, Luettich T, Fritzler MJ. Limited reliability of the indirect immunofluorescence technique for the detection of anti-Rib-P antibodies. Arthritis Res Ther 2008;10:R131.
14. Pisetsky DS, Spencer DM, Lipsky PE, Rovin BH. Assay variation in the detection of antinuclear antibodies in the sera of patients with established SLE. Ann Rheum Dis 2018;77:911–3.
15. Cross LS, Aslam A, Misbah SA. Antinuclear-negative lupus as a distinct diagnostic entity: does it no longer exist? QJM 2004;97:303–8.
16. Sjowall C, Sturm M, Dahle C, Bengtsson AA, Jonsen A, Sturfelt G, et al. Abnormal antinuclear antibody titers are less common than generally assumed in established cases of systemic lupus erythematosus. J Rheumatol 2008;35:1994–2000.
17. Simmons SC, Smith ML, Chang-Miller A, Keddis MT. Antinuclear antibody-negative Lupus nephritis with full house nephropathy: a case report and review of the literature. Am J Nephrol 2015;42:451–9.
18. Leuchten N, Hoyer A, Brinks R, Schoels M, Schneider M, Smolen J, et al. Performance of antinuclear antibodies for classifying systemic lupus erythematosus: a systematic literature review and meta-regression of diagnostic data. Arthritis Care Res (Hoboken) 2018;70:428–38.
19. Reichlin M. ANA negative systemic lupus erythematosus sera revisited serologically. Lupus 2000;9:116–9.
20. Fritzler MJ. Choosing wisely: review and commentary on anti-nuclear antibody (ANA) testing. Autoimmun Rev 2016;15:272–80.
21. Urowitz MB, Gladman DD, Ibanez D, Sanchez-Guerrero J, Romero-Diaz J, Gordon C, et al. American College of Rheumatology criteria at inception, and accrual over 5 years in the SLICC inception cohort. J Rheumatol 2014;41:875–80.
22. Hanly JG, O'Keeffe AG, Su L, Urowitz MB, Romero-Diaz J, Gordon C, et al. The frequency and outcome of lupus nephritis: results from an international inception cohort study. Rheumatology (Oxford) 2016;55:252–62.
23. Isenberg DA, Ramsey-Goldman R, Gladman D, Hanly JG. The Systemic Lupus International Collaborating Clinics (SLICC) group: it was 20 years ago today. Lupus 2011;20:1426–32.
24. Tan EM, Feltkamp TE, Smolen JS, Butcher B, Dawkins R, Fritzler MJ, et al. Range of antinuclear antibodies in "healthy" individuals. Arthritis Rheum 1997;40:1601–11.
25. Fritzler MJ, Hanson C, Miller J, Eystathioy T. Specificity of autoantibodies to SS-A/Ro on a transfected and overexpressed human 60 kDa Ro autoantigen substrate. J Clin Lab Anal 2002;16:103–8.
26. Keech CL, Howarth S, Coates T, Rischmueller M, McCluskey J, Gordon TP. Rapid and sensitive detection of anti-Ro (SS-A) antibodies by indirect immunofluorescence of 60 kDa Ro HEp-2 transfectants. Pathology 1996;28:54–7.
27. Choi MY, Clarke AE, St Pierre Y, Hanly JG, Urowitz MB, Romero-Diaz J, et al. The prevalence and determinants of anti-DFS70 autoantibodies in an international inception cohort of systemic lupus erythematosus patients. Lupus 2017;26:1051–9.
28. Hanly JG, Urowitz MB, Siannis F, Farewell V, Gordon C, Bae SC, et al. Autoantibodies and neuropsychiatric events at the time of systemic lupus erythematosus diagnosis: results from an international inception cohort study. Arthritis Rheum 2008;58:843–53.
29. Chan EK, Damoiseaux J, de Melo CW, Carballo OG, Conrad K, Francescantonio PL, et al. Report on the second International Consensus on ANA Pattern (ICAP) workshop in Dresden 2015. Lupus 2016;25:797–804.
30. Gladman DD, Chalmers A, Urowitz MB. Systemic lupus erythematosus with negative LE cells and antinuclear factor. J Rheumatol 1978;5:142–7.
31. Maddison PJ, Provost TT, Reichlin M. Serological findings in patients with "ANA-negative" systemic lupus erythematosus. Medicine (Baltimore) 1981;60:87–94.

32. Fernández M, Alarcón GS, Calvo-Alén J, Andrade R, McGwin G Jr, Vilá LM, et al. A multiethnic, multicenter cohort of patients with systemic lupus erythematosus (SLE) as a model for the study of ethnic disparities in SLE. Arthritis Rheum 2007;57: 576–84.
33. Feldman CH, Hiraki LT, Liu J, Fischer MA, Solomon DH, Alarcón GS, et al. Epidemiology and sociodemographics of systemic lupus erythematosus and lupus nephritis among US adults with Medicaid coverage, 2000–2004. Arthritis Rheum 2013;65:753–63.
34. Barnett EV, North AF Jr, Condemi JJ, Jacox RF, Vaughan JH. Antinuclear factors in systemic lupus erythematosus and rheumatoid arthritis. Ann Intern Med 1965;63:100–8.
35. Weisbart RH, Colburn K. Effect of corticosteroids on serum antinuclear antibodies in man. Immunopharmacology 1984;8:97–101.
36. Lindholm C, Borjesson-Asp K, Zendjanchi K, Sundqvist AC, Tarkowski A, Bokarewa M. Longterm clinical and immunological effects of anti-CD20 treatment in patients with refractory systemic lupus erythematosus. J Rheumatol 2008;35:826–33.
37. Bossuyt X, Meurs L, Mewis A, Marien G, Blanckaert N. Screening for autoantibodies to SS-A/Ro by indirect immunofluorescence using HEp-2000TM cells. Ann Clin Biochem 2000;37:216–9.
38. Peene I, Van Ael W, Vandenbossche M, Vervaet T, Veys E, De Keyser F. Sensitivity of the HEp-2000 substrate for the detection of anti-SSA/Ro60 antibodies. Clin Rheumatol 2000;19:291–5.
39. Sontheimer RD, Maddison PJ, Reichlin M, Jordon RE, Stastny P, Gilliam JN. Serologic and HLA associations in subacute cutaneous lupus erythematosus, a clinical subset of lupus erythematosus. Ann Intern Med 1982;97:664–71.
40. Perez D, Gilburd B, Cabrera-Marante O, Martinez-Flores JA, Serrano M, Naranjo L, et al. Predictive autoimmunity using autoantibodies: screening for anti-nuclear antibodies. Clin Chem Lab Med 2018;56:1771–7.
41. Worrall JG, Snaith ML, Batchelor JR, Isenberg DA. SLE: a rheumatological view. Analysis of the clinical features, serology and immunogenetics of 100 SLE patients during long-term follow-up. Q J Med 1990;74:319–30.
42. Compagno M, Jacobsen S, Rekvig OP, Truedsson L, Heegaard NH, Nossent J, et al. Low diagnostic and predictive value of anti-dsDNA antibodies in unselected patients with recent onset of rheumatic symptoms: results from a long-term follow-up Scandinavian multicentre study. Scand J Rheumatol 2013;42:311–6.
43. Baronaite R, Engelhart M, Mork HT, Thamsborg G, Slott JH, Stender S, et al. A comparison of anti-nuclear antibody quantification using automated enzyme immunoassays and immunofluorescence assays. Autoimmune Dis 2014;2014:534759.
44. Lindstedt G, Lundberg PA, Westberg G, Kaijser B. S.L.E. nephritis with positive tests for antibodies against native D.N.A. but negative tests for antinuclear antibodies [letter]. Lancet 1977;2:135.
45. Morris CN, Calobrisi SD, Matteson EL. Antinuclear antibody negative lupus associated with dystrophic calcification. J Rheumatol 1998;25:825–6.
46. Zhao M. ANA-negative presentation of SLE in man with severe autoimmune neutropenia. Case Rep Med 2016;2016:6853936.
47. Compagno M, Rekvig OP, Bengtsson AA, Sturfelt G, Heegaard NH, Jonsen A, et al. Clinical phenotype associations with various types of anti-dsDNA antibodies in patients with recent onset of rheumatic symptoms: results from a multicentre observational study. Lupus Sci Med 2014;1:e000007.
48. Choi MY, Barber MR, Barber CE, Clarke AE, Fritzler MJ. Preventing the development of SLE: identifying risk factors and proposing pathways for clinical care. Lupus 2016;25:838–49.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 903–913
DOI 10.1002/acr.23721

American College of Rheumatology
Empowering Rheumatology Professionals

Area-Level Predictors of Medication Nonadherence Among US Medicaid Beneficiaries With Lupus: A Multilevel Study

Candace H. Feldman,[1] Karen H. Costenbader,[1] Daniel H. Solomon,[1] S. V. Subramanian,[2] and Ichiro Kawachi[2]

Objective. Adherence to hydroxychloroquine (HCQ) treatment in patients with systemic lupus erythematous (SLE) is suboptimal. Although individual-level factors, including younger age and non-white race/ethnicity, have been implicated, contextual factors have not been explored. The aim of this study was to investigate the effect of contextual factors, including racial composition, socioeconomic status, and the concentration of health care resources, on adherence to HCQ among SLE patients enrolled in Medicaid.

Methods. We identified SLE patients from 28 states in the US who enrolled in Medicaid (2000–2010) and in whom HCQ treatment was newly initiated (no use for ≥6 months). We required 12 months of continuous enrollment with complete drug dispensing data and measured adherence using the proportion of days covered (PDC). We identified individual-level variables from Medicaid, zip code–level, county-level and state-level sociodemographic variables from the American Community Survey, and health resources from Area Health Resources Files. We used 4-level hierarchical multivariable logistic regression models to examine the odds ratios (ORs) and 95% credible intervals (95% CrIs) of adherence (PDC ≥80%) versus nonadherence.

Results. Among 10,268 patients with SLE in whom HCQ treatment was initiated, 15% were adherent to treatment. After we adjusted for individual-level characteristics, we observed lower odds of adherence among patients living in zip code areas with a higher percentage of black individuals (highest tertile OR 0.81 [95% CrI 0.69–0.96] versus lowest tertile). This association persisted after controlling for area-level educational attainment, percent below federal poverty level (FPL), urbanicity, and health care resources. We did not observe statistically significant associations with zip code–level percent Hispanic, percent white, education, or percent below FPL. The odds of adherence were higher in counties with more hospitals (OR 1.30 [95% CrI 1.07–1.58]).

Conclusion. Among Medicaid beneficiaries with SLE, we observed significant effects of racial composition and hospital concentration on HCQ adherence. Interventions that acknowledge and address contextual factors should be considered in order to reduce high rates of nonadherence in vulnerable populations.

INTRODUCTION

SLE is a multisystem autoimmune disease that disproportionately affects racial/ethnic minorities and individuals living in areas of lower socioeconomic status (SES) (1,2). Medication nonadherence is common among patients with SLE and is more pronounced among racial/ethnic minorities and individuals with lower SES, who also are affected by the highest burden of adverse outcomes (3–5). A previous study demonstrated an effect of neighborhood poverty on accumulation of SLE-specific damage; however, the role of medication adherence was not investigated (6). The majority of adherence-related studies focus on the isolated contribution of patient-related and disease-related characteristics. Despite the increased awareness of the importance of social determinants on health behaviors and outcomes, few studies have examined the association between these factors and adherence (7–9).

Therefore, the aim of this study was to investigate the effect of contextual factors, including racial composition, SES, and health care resource concentration, on adherence among SLE patients receiving Medicaid, the largest federal/state public health insurance program for low-income Americans. We focused on adherence to hydroxychloroquine (HCQ) treatment,

Dr. Feldman's work was supported by a Rheumatology Research Foundation Investigator Award and the NIH (grant K23-AR-071500 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases). Dr. Costenbader's work was supported by NIH grants R01-057327 and K24-AR-066109. Dr. Solomon's work was supported by NIH grant K24-AR-055989.

[1]Candace H. Feldman, MD, ScD, Karen H. Costenbader, MD, MPH, Daniel H. Solomon, MD, MPH: Brigham and Women's Hospital, Boston, Massachusetts; [2]S. V. Subramanian, PhD, Ichiro Kawachi, MB, ChB, PhD: Harvard T. H. Chan School of Public Health, Boston, Massachusetts.

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

Address correspondence to Candace H. Feldman, MD, ScD, 60 Fenwood Road, Office 6016P, Boston, MA 02115. E-mail: cfeldman@bwh.harvard.edu.

Submitted for publication September 21, 2017; accepted in revised form July 24, 2018.

SIGNIFICANCE & INNOVATIONS

- Despite the increasingly recognized importance of social determinants regarding health and health behaviors, the influence of contextual factors including area-level sociodemographics and health care resource concentration on medication adherence has not been well studied.
- The current study showed nearly 20% reduced odds of adherence to hydroxychloroquine (HCQ) treatment among patients with systemic lupus erythematosus living in zip code–level areas with higher percentages of black individuals, after adjusting for individual demographics and health-related factors as well as zip code–level poverty. This suggests a possible effect of racial residential segregation on HCQ adherence.
- Although a higher concentration of hospitals was associated with increased odds of adherence, other health care resources including numbers of physicians and pharmacists per capita were not associated with improved adherence, suggesting that other factors, such as the quality of health care resources in underserved areas, may play a stronger role.
- The study findings suggest that contextual factors play a role in HCQ adherence behavior among SLE patients and should be considered in the design of interventions that aim to reduce the racial/ethnic disparities observed.

because it is the standard of care for both active treatment and prevention of complications for nearly all patients with SLE (10).

We utilized the social ecological model (SEM), which emphasizes the importance of multiple reciprocal levels of influence on behavior, to consider potential individual-level and area-level contributors to treatment adherence (Figure 1) (11). The theoretical framework proposed by August and Billimek linking disadvantaged neighborhoods to an increased likelihood of medication nonadherence provided potential mechanisms to explain the interplay between SEM levels and adherence (12). Those investigators hypothesized that living in a disadvantaged neighborhood results in increased exposure to environmental and social stressors and reduces self-regulatory resources to engage in healthy behaviors. Individuals often place disproportionate weight on current costs and benefits relative to future costs and benefits, and in environments with a high level of social stressors, this is more pronounced (13,14). Adherence to medications for chronic disease for potential future benefit may therefore be outweighed by more immediate concerns. Social factors may also contribute to norms that devalue adherence (15,16). Taken together, these aspects may result in unfavorable beliefs about medication adherence, leading to nonadherence (12). Based on this theoretical framework and the SEM, we hypothesized that residing in racially segregated, lower SES areas with fewer health care–related resources would be associated with increased odds of nonadherence to HCQ treatment.

PATIENTS AND METHODS

Patient population. We used Medicaid Analytic eXtract (MAX) files from the 29 most populous US states (~86% of all Medicaid beneficiaries) from 2000 to 2010 with drug dispensing data, billing claims, and health care utilization data. Although Ohio was part of our data set, it was excluded from the analyses due to incomplete dispensing data (n = 388 patients). We identified patients ages 18–65 years with prevalent SLE, defined as ≥2 International Classification of Diseases, Ninth Revision (ICD-9) codes for SLE (710.0) from hospital discharge diagnoses or physician visit claims ≥30 days apart, and HCQ dispensing within 365 days of an SLE ICD-9 code. We restricted our cohort to new users of HCQ, defined as 183 days of continuous enrollment prior to the first dispensing of HCQ (index date) with no use of HCQ during this time. Because >90% of Medicaid beneficiaries received 1-month supplies of HCQ, a 6-month period was determined to be sufficient to define initiation. We were unable to validate our algorithm due to federal restrictions on the use of identifiable Medicaid claims data for research purposes. However, similar claims-based algorithms performed well (sensitivity 98.2%, specificity 72.5%), and adding HCQ use likely increased specificity (17). We required ≥365 days of continuous enrollment following the index date in order to assess adherence. We excluded patients if their dispensing data were missing (n = 333 patients), if they were hospitalized for the entire follow-up period (n = 18 patients), or if their zip codes were not reported (n = 253 patients) or were discordant with their state of residence (n = 46 patients).

Outcome for adherence. Adherence was determined using drug dispensing data: the claims processed by Medicaid when a prescription is filled at a pharmacy and either picked up by or is mailed to a patient, including medication name/National Coverage Determination (NDC) code, dose, quantity, amount prescribed, and number of refills. We defined adherence using the proportion of days covered (PDC) for the 365-day period beginning at the index date. The PDC was defined as the number of days covered by dispensed HCQ prescriptions divided by 365 days, multiplied by 100. We subtracted hospitalized days from the numerator and denominator. The PDC is a widely used, validated measure to assess adherence in claims data, including Medicaid (18,19). We used a PDC threshold of ≥80% to define adherence, consistent with the literature on adherence to other chronic disease medications (20). In sensitivity analyses, we examined adherence thresholds of ≥90% and ≥70%.

Individual-level covariates. We used MAX data to identify individual-level covariates in the 183-day period prior to and includ-

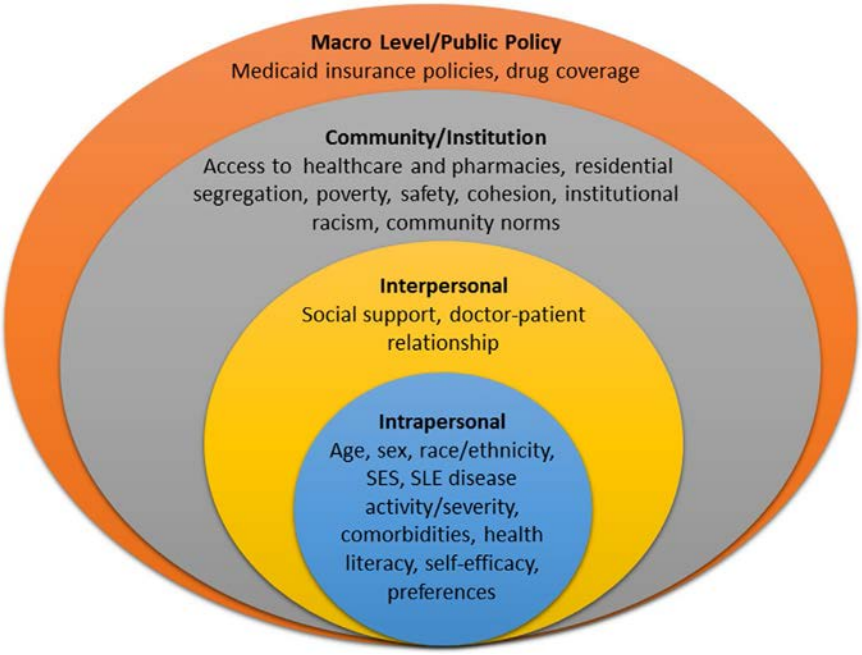


Figure 1. Social ecological model adapted to demonstrate potential multilevel contributors to adherence behavior among patients with systemic lupus erythematosus (SLE). SES = socioeconomic status.

ing the index date. Our choice of variables was informed by prior literature regarding adherence and our hypothesis that they may be related to both initial HCQ prescribing and adherence. We included age, sex, and race/ethnicity. Medicaid uses the terminology "Black or African American," and herein we will use the term "black" to represent this group (21). We used the SLE-specific risk adjustment index, which has been shown to be a better predictor of mortality among SLE patients compared with the Charlson comorbidity index, as a marker of disease severity (22). We separately assessed lupus nephritis, using ≥2 ICD-9 discharge diagnosis codes or physician billing claims for nephritis, proteinuria, and/or renal failure occurring on or after 1 SLE diagnosis code (23) to capture additional cases of lupus nephritis beyond those included in the adjustment index algorithm, because we hypothesized that this would be an important confounder. As proxies for SLE disease activity, the numbers of SLE-related laboratory tests were assessed.

We measured health care utilization (emergency department visits, inpatient visits, and outpatient visits) as markers of care fragmentation and disease activity/severity. We included the number of medications to assess for polypharmacy and glucocorticoids and other immunosuppressive medications to represent disease activity/severity. We included diabetes mellitus (by ICD-9 code or related medication prescription) (24) and antidepressant medication treatment as a marker of depression, given the low positive predictive value of depression-related claims (25). To account for health status and healthy/unhealthy behaviors potentially associated with treatment adherence, we included Current Procedural Terminology (CPT) codes for vaccinations, medication (NDC) codes for pneumocystis prophylaxis, obesity (≥1 ICD-9 code), and smoking (≥1 ICD-9 code), CPT code for smoking cessation counseling or dispensing of related medications (26).

Contextual covariates. From de-identified Medicaid data, we obtained data for zip code (the smallest geographic area available), county, and state of residence. Using American Community Survey (ACS) data (2006–2010), we extracted the percent of the population in each zip code with incomes below the federal poverty level (FPL), zip code–level educational attainment (high school or less, some college, and college graduate and beyond), as well as zip code–level percent black, white, and Hispanic (27–29). We examined a composite score of zip code–level percent black and percent below FPL, divided into tertiles, as well as a variable that included zip codes with the lowest tertile of percent black and of percent FPL in group 1 (n = 760 patients) and with the highest tertile of both in group 3 (n = 934 patients), with others in group 2 (n = 3,236 patients). We used the ACS to determine the state-level Gini coefficient, a marker of income inequality (0–1 scale, where 1 represents the greatest inequality) (29). We used the mean of US Census data from 2000 and 2010 to determine the urbanicity of each county.

To determine county-level concentrations of health care services and health professionals, we used the Area Health Resources Files from 2000 and 2010 (30,31). We obtained the mean numbers of physicians, medicine subspecialists, pharmacists, and hospitals from 2000 and 2010 (2009 for pharmacists) and divided these by each county's population size, per 1,000 individuals. There were 52 counties (4%) without reported pharmacist data. We also determined the county

Health Professional Shortage Area (HPSA) designation (none, partial, whole). HPSAs are characterized by shortages of primary care providers, which may be in a given geographic area, for a specific population, or within a type of facility in a given area. We obtained data for the number of rheumatologists per state from the American College of Rheumatology (2000) and divided this number by population size according to the US Census (2000), per 10,000 individuals. We examined our contextual variables in tertiles (reference, lowest) to facilitate interpretability and to be consistent across levels.

Statistical analyses for multilevel models. In order to examine associations across contextual variables with individual-level adherence and to understand the potential heterogeneities underlying these associations, we constructed hierarchical 4-level multivariable random intercepts logistic regression models with individuals (level 1) nested in zip codes (level 2), in counties (level 3), and in states (level 4) (32). All analyses used Markov chain Monte Carlo (MCMC) procedures using a Metropolis–Hasting algorithm (33). This is a Bayesian approach that relies on sequential learning, whereby prior information is accounted for in the estimates and the distributional assumptions of maximum likelihood methods are not required (33,34). For our fixed-effects estimates, we present the MCMC estimates of odds ratios (ORs) with 95% credible intervals (CrIs). Unlike confidence intervals, CrIs do not have to be normally distributed and provide the potential range of values following the MCMC simulation of many model runs. For random-effects parameters, we present level-specific residual variance estimates and the percent of variance partitioned at each level.

We examined fixed effects of individual-level and contextual variables and accounted for random effects at each geographic level. We conducted sensitivity analyses adjusting for fewer individual-level covariates. We examined variance partitioning at each level in order to understand the level at which most between-area variability was occurring and the degree to which our fixed effects could account for this. For models examining health care resources, we began with the model with the smallest deviance information criterion (DIC) statistic, which is used to compare hierarchical models obtained by MCMC procedures. We used SAS version 9.4 software to organize variables and MLwiN version 2.36 to conduct multilevel analyses (35). MAX data were obtained through a data use agreement from Centers for Medicare & Medicaid Services (CMS), and findings are reported in accordance with their specifications. In this study, we used only de-identified, aggregated data, and the study was approved by the Partners Human Research Committee.

RESULTS

Individual-level and area-level characteristics. We identified 10,268 Medicaid beneficiaries residing within 4,930 zip code areas, in 1,414 counties, in 28 states, all of whom had prevalent SLE and were initiating treatment with HCQ. On average, there was a mean ± SD of 2.1 ± 2.8 (range 1–29) individuals per zip code, 7.3 ± 28.7 (range 1–739) individuals per county, and 366.7 ± 416.9 (range 84–1,960) individuals per state. There were, on average, 3.5 ±7.5 (range 1–190) zip codes per county and 50.5 ± 24.9 (range 13–120) counties per state. The mean ± SD age of the patients was 37.7 ± 11.8 years, and 94% of the patients were female (Table 1). The mean ± SD PDC was 42.3 ± 28.7%, and 1,567 individuals (15.3%) were adherent to HCQ treatment (PDC ≥80%). Area-level sociodemographic and health care resource characteristics are shown in Table 2.

Individual-level fixed effects. In our 4-level random intercepts logistic regression model, we observed lower odds of adherence (versus nonadherence) associated with younger age (versus older age), with black race and Hispanic ethnicity (versus white), with diabetes mellitus, antidepressant medication use, and with more emergency department visits (Table 3). We observed higher odds of adherence associated with Asian race, higher SLE risk adjustment index, and more laboratory tests and medications.

Area-level fixed effects. *Sociodemographics*. After accounting for individual-level fixed effects and random effects at each level, we observed lower odds of adherence (versus nonadherence) in zip codes with higher percentages of black individuals. Comparing tertile 2 to tertile 1 (lowest percentage of black individuals), the OR was 0.85 (95% CrI 0.74–0.98), and comparing tertile 3 to tertile 1, the OR was 0.81 (95% CrI 0.69–0.96) (Table 4). We did not find statistically significant associations with zip code percent white or Hispanic or with percent below FPL or educational attainment. When we adjusted our model by percent below FPL and education, the effect of percent black remained significant (OR 0.80 [95% CrI 0.67–0.96] for tertile 3 versus tertile 1, and OR 0.86 [95% CrI 0.74–0.99] for tertile 2 versus tertile 1). We did not find statistically significant associations between our combined percent black/percent below FPL measures and adherence or with the Gini coefficient.

We examined cross-level interactions between individual race (categorized as black, white, Hispanic, and other) and zip code percent black and with zip code percent below FPL to determine whether the relationship between race and adherence differs by the racial composition or poverty status of the area in which a person lives. The interaction term was significant for black race*percent black tertile 3 (reference, white and tertile 1). The OR for adherence versus nonadherence for individuals of black versus white race living in higher versus lower percent black areas (tertile 3 versus tertile 1) was 0.49 (compared to OR 0.56 in the model without the interaction term). The interaction terms were not significant when black race was used as the reference or for race*percent below FPL.

Table 1. Individual-level baseline characteristics of 10,268 lupus patients enrolled in US Medicaid 2000–2010 in whom HCQ was newly initiated*

Age, mean ± SD years	37.7 ± 11.8
Age group, years	
18–34	4,542 (44.2)
35–50	3,902 (38.0)
51–65	1,824 (17.8)
Sex	
Female	9,669 (94.2)
Male	599 (5.8)
Race/ethnicity	
Black	4,289 (41.8)
White	3,194 (31.1)
Hispanic	2,036 (19.8)
Asian	397 (3.9)
American Indian/Alaska Native	121 (1.2)
Other	231 (2.3)
SLE risk adjustment index, mean ± SD	1.0 ± 1.9
Lupus nephritis	1,049 (10.2)
Diabetes mellitus	964 (9.4)
Smoking	616 (6.0)
Obesity	240 (2.3)
Antidepressant use	2,976 (29)
Preventive care	
Influenza vaccine	178 (1.7)
Pneumococcal vaccine	58 (0.6)
Pneumocystis pneumonia prophylaxis	903 (8.8)
Immunosuppressive medication use†	1,127 (11)
Number of laboratory tests, mean ± SD‡	1.6 ± 2.8
Ever use of glucocorticoids	6,112 (59.5)
Mean no. medications, mean ± SD	4.2 ± 3.4
Health care utilization, median (25th, 75th percentiles)	
Emergency department visits	0 (0, 1)
Hospitalizations	0 (0, 1)
Outpatient visits	2 (0, 6)
No. of hospitalized days, mean ± SD (median [25th, 75th percentiles])	3.9 ± 10.6 (0 [0, 4])

* Baseline characteristics were determined from the 183 days prior to and including the first dispensing of hydroxychloroquine (index date). Data for the following 28 US states were included: Alabama, Arkansas, California, Colorado, Florida, Georgia, Illinois, Indiana, Kentucky, Louisiana, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, New Jersey, New Mexico, New York, North Carolina, Oklahoma, Oregon, Pennsylvania, South Carolina, Texas, Washington, and Wisconsin. Except where indicated otherwise, values are the number (%). SLE = systemic lupus erythematosus.
† Includes azathioprine, cyclophosphamide, leflunomide, methotrexate, mycophenolate mofetil, sulfasalazine, and tacrolimus.
‡ Includes blood urea nitrogen, creatinine, urinalysis, complement (C3, C4), erythrocyte sedimentation rate, C-reactive protein, and anti–double-stranded DNA.

Health care resources. We built our subsequent models including individual-level factors (Table 3) as well as zip code–level percent black both because we aimed to test whether health care resource concentration and urbanicity accounted for the effect of zip code percent black on adherence, and because it had a lower DIC compared to the individual variable–only model. We observed higher odds of adherence versus nonadherence comparing the county-level highest number of hospitals per capita compared to the lowest (OR 1.30 [95% CrI 1.07–1.58]) (Table 5). We found a trend toward lower odds of adherence versus nonadherence associated with residing in a whole health care professional shortage area (OR 0.88 [95% CrI 0.77–1.01]). We did not find statistically significant associations with number of physicians, medicine subspecialists, pharmacists per capita, or rheumatologists. We also explored urbanicity and found a borderline significant relationship with higher percent urban and reduced odds of adherence (OR 0.82 [95% CrI 0.66–1.02], tertile 3 versus tertile 1). The magnitude of the effect of zip code percent black and adherence remained consistent and either significant or nearly significant in all models (Table 5).

Random effects. We observed modest between-state–level, county-level, and zip code–level variation in adherence. Most variance detected was attributable to between-state variation (1.4% in our null model), with <1% attributable to between-county or between–zip code variation. We observed minor differences in random effects in our adjusted models (between-state variance 1–1.5%, between-county variance 0.1–0.3%, and between–zip code variances 0.03–0.2%) (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23721/abstract).

Sensitivity analyses. We examined the relationship between zip code percent black and adherence at 2 different adherence thresholds. For PDC ≥70%, 22.2% of patients in our cohort were now categorized as adherent. At this threshold, we found similar associations as those observed in our primary analyses for tertile 2 versus tertile 1 (OR 0.83 [95% CrI 0.72–0.94]) and for tertile 3 versus tertile 1 (OR 0.81 [95% CrI 0.69–0.94]) as well as for all individual-level variables. For PDC ≥90%, 7.7% of patients in our cohort were considered adherent, and while the magnitude and direction of associations were similar, the 95% CrIs were wider (for tertile 2 versus tertile 1, OR 0.83 [95% CrI 0.69–1.01]; for tertile 3 versus tertile 1, OR 0.89 [95% CrI 0.71–1.11]). In addition, we examined more parsimonious models that included only sex, age, race/ethnicity, SLE risk adjustment index, number of medications, and calendar year at the individual level and found estimates similar in magnitude and significance for fixed and random effects at all levels (see Supplementary Tables 2a and b and Table 3, available

Table 2. Zip code–level, county-level, and state-level demographic and health services characteristics of lupus patients enrolled in US Medicaid 2000–2010 in whom HCQ was newly initiated*

Area-level characteristics	Zip code–level (n = 4,930)	County-level (n = 1,414)	State-level (n = 28)
Percent with income below the FPL, mean ± SD†	14 ± 9	–	–
Tertile 1	5.4 ± 2.2 (0–8.9)	–	–
Tertile 2	12.4 ± 2.1 (8.9–16.3)	–	–
Tertile 3	34.5 ± 7.3 (16.3–70.7)	–	–
Percent black, mean ± SD†	18 ± 23	–	–
Tertile 1	1.1 ± 1.0 (0–3.2)	–	–
Tertile 2	8.9 ± 4.2 (3.2–17.8)	–	–
Tertile 3	45.1 ± 22.1 (17.8–100)	–	–
Percent white, mean ± SD	67 ± 25	–	–
Tertile 1	37.1 ± 16 (0–58.4)	–	–
Tertile 2	71.4 ± 6.9 (58.4–82.5)	–	–
Tertile 3	91.5 ± 4.9 (82.5–100)	–	–
Percent Hispanic, mean ± SD	18 ± 22	–	–
Tertile 1	2 ± 1 (0–3.9)	–	–
Tertile 2	8.3 ± 3.3 (3.9–15.9)	–	–
Tertile 3	42.8 ± 22.6 (15.9–99.1)	–	–
Educational attainment, no. (%)†		–	–
High school or less	3,904 (79.2)	–	–
Some college	288 (5.8)	–	–
College graduate and beyond	738 (15)	–	–
Total number of total MDs, mean ± SD‡	–	1.53 ± 1.75	–
Tertile 1	–	0.43 ± 0.18 (0–0.72)	–
Tertile 2	–	1.09 ± 0.25 (0.72–1.56)	–
Tertile 3	–	3.06 ± 2.3 (1.56–25.3)	–
Total number of hospitals, mean ± SD‡	–	0.03 ± 0.03	
Tertile 1	–	0.01 ± 0.01 (0–0.02)	–
Tertile 2	–	0.03 ± 0.01 (0.02–0.04)	–
Tertile 3	–	0.07 ± 0.04 (0.04–0.59)	–
Total no. of medical subspecialists, mean ± SD‡	–	0.46 ± 0.64	–
Tertile 1	–	0.08 ± 0.05 (0–0.16)	–
Tertile 2	–	0.30 ± 0.09 (0.16–0.47)	–
Tertile 3	–	1 ± 0.87 (0.47–0.5)	–
Total no. of pharmacists, mean ± SD‡	–	0.54 ± 0.33	–
Tertile 1	–	0.23 ± 0.07 (0–0.34)	–
Tertile 2	–	0.45 ± 0.07 (0.34–0.60)	–
Tertile 3	–	0.90 ± 0.30 (0.60–2.73)	–
Health professional shortage area, no. (%)§			
None	–	236 (16.7)	–
Partial	–	615 (43.5)	–
Whole	–	563 (39.8)	–
Percent urban, mean ± SD¶	–	51.2 ± 29.8	–
Tertile 1	–	17.0 ± 12.8 (0–36.8)	–
Tertile 2	–	51.5 ± 8.6 (36.8–67.3)	–
Tertile 3	–	85.2 ± 10.2 (67.3–100)	–

(continued)

Table 2. (*Cont'd*)

Area-level characteristics	Zip code–level (n = 4,930)	County-level (n = 1,414)	State-level (n = 28)
No. of rheumatologists, mean ± SD#	–	–	0.13 ± 0.04
Tertile 1	–	–	0.09 ± 0.01 (0.07–0.11)
Tertile 2	–	–	0.12 ± 0.01 (0.11–0.14)
Tertile 3	–	–	0.17 ± 0.04 (0.14–0.26)
Gini coefficient, mean ± SD†	–	–	0.46 ± 0.02
Tertile 1	–	–	0.44 ± 0.01 (0.43–0.45)
Tertile 2	–	–	0.46 ± 0.003 (0.45–0.47)
Tertile 3	–	–	0.47 ± 0.01 (0.47–0.50)

* Values for tertiles are the mean ± SD (range). HCQ = hydroxychloroquine; FPL = federal poverty level.
† Derived from the American Community Survey (2000–2006).
‡ Per capita, per 1,000 individuals in the county, mean from the Area Health Resources Files 2000 and 2010. MD data were from the American Medical Association hospital data from the Health Resources and Services Administration (HRSA). For pharmacists (n = 1,362), 52 counties (4%) do not report pharmacist data.
§ Defined and reported by the HRSA, 2010.
¶ From the Area Health Resources files, mean estimates from 2000 and 2010 data.
# From the American College of Rheumatology (2000), per 10,000 state population in 2000 (US Census).

on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23721/abstract).

DISCUSSION

In this vulnerable population of patients with SLE among whom only 15% were adherent to HCQ treatment over the first year of use, we observed significant relationships between individual and contextual factors and adherence. Although it is challenging to compare across studies given differences in populations (e.g., academic-based cohorts versus public insurance beneficiaries) and in adherence measures (e.g., self-reported surveys versus claims-based prescription refill data), our findings were similar to estimates of adherence to other SLE-related medications in the Medicaid population (36). In line with prior studies, younger age, black race, Hispanic ethnicity, and antidepressant use were associated with poorer adherence (5,24,37–39). Nonadherence, particularly among black patients with lupus, is associated with poorer outcomes overall (3,5,37,38). Because HCQ is the backbone of SLE care and has been associated with fewer disease flares and improved outcomes, it is plausible that differences in adherence by race/ethnicity may contribute to disparities in outcomes (3,40–42).

Few studies have explored the contribution of contextual factors to these disparities (6). We found a significant dose-response relationship between zip code–level percent black and poorer adherence to HCQ, which remained after adjusting for individual race, and for aggregate area-level poverty and educational attainment. We did not, however, observe associations with zip code–level percent white or percent Hispanic. While zip code is an imperfect proxy for neighborhood (43), these observations suggest a potential effect of racial residential segregation on HCQ nonadherence. Our lack of addresses did not allow us to geocode our data to examine the differential distribution by race across smaller residential units (e.g., census blocks) within larger geographic areas or to look at spatial relationships (44). We hypothesize that in these areas, there may be reduced access to high-quality hospitals, outpatient providers, and pharmacy services and to transportation (44,45). There may be increased stress and depression because of racial discrimination and crime, which may contribute to nonadherence (45,46). There may also be reverse causation, because individuals who are less likely to engage in healthy behaviors may also be less equipped to move out of racially segregated areas. This may be related to the known differences in the opportunities afforded to individuals who live in segregated neighborhoods (45).

It is also possible that social and cultural norms within neighborhoods influence adherence behavior. Racial bias experienced in health care is one plausible mechanism (45). Prior studies suggest differences between black and white patients in their willingness to accept certain SLE-related medications, in beliefs about medication effectiveness, and in health provider trust (47). In one study, >50% of black patients with SLE experienced racial discrimination in health care, and this was linked to increased depression (46). Similar patterns have been described among black patients with hypertension, and racial discrimination has been shown to be associated with medication nonadherence (48). Living near individuals with negative health care experiences because of racism may contribute to nonadherence norms.

Table 3. Odds of HCQ adherence versus nonadherence among lupus patients enrolled in US Medicaid 2000–2010 in whom HCQ was newly initiated*

Individual-level fixed effects	OR (95% CrI)
Male sex (ref. female sex)	1.21 (0.97–1.50)
Age, years (ref. 51–65 years)	
18–34	0.60 (0.51–0.71)†
35–50	0.71 (0.61–0.82)†
Race/ethnicity (ref. white)	
Black	0.61 (0.54–0.71)†
Hispanic	0.71 (0.60–0.83)†
Asian	1.49 (1.15–1.92)†
American Indian/Alaska native	0.78 (0.45–1.92)
Other	0.85 (0.59–1.22)
SLE risk adjustment index	1.04 (1.01–1.07)
Lupus nephritis	1.13 (0.92–1.38)
Diabetes mellitus	0.82 (0.68–1.00)†
Antidepressant use	0.86 (0.76–0.98)†
Glucocorticoid use	1.04 (0.92–1.16)
No. of laboratory tests	1.03 (1.00–1.05)†
No. of medications	1.10 (1.08–1.12)†
Immunosuppressive use	0.91 (0.77–1.07)
Health care utilization	
No. of emergency department visits	0.96 (0.93–1.00)†
No. of outpatient visits	1.00 (0.99–1.01)
No. of hospitalizations	0.95 (0.90–1.00)‡

* Adherence is defined as proportion of days covered (PDC) ≥80%. Nonadherence is defined as PDC <80%. The 4-level random intercepts multivariable regression model additionally adjusted for calendar year of index date and healthy adherer effect variables (obesity, smoking, and preventive care [influenza or pneumonia vaccines, pneumocystis pneumonia prophylaxis]) as well as random effects at each level. HCQ = hydroxychloroquine; OR = odds ratio; 95% CrI = 95% credible interval; SLE = systemic lupus erythematosus.
† 95% CrIs do not cross 1 and Bayesian *P* value <0.05.
‡ Bayesian *P* value <0.05 with 95% CrI that just crosses 1.

Table 4. Odds of HCQ adherence versus nonadherence according to contextual sociodemographic factors*

Area-level sociodemographic factors	OR (95% CrI)
Zip code–level percent black (ref. lowest tertile)	
Tertile 2	0.85 (0.74–0.98)†
Tertile 3	0.81 (0.69–0.96)†
Zip code–level percent white (ref. lowest tertile)	
Tertile 2	0.93 (0.81–1.08)
Tertile 3	1.13 (0.94–1.34)
Zip code–level percent Hispanic (ref. lowest tertile)	
Tertile 2	0.92 (0.77–1.07)
Tertile 3	0.90 (0.74–1.06)
Zip code–level percent below FPL (ref. lowest tertile)	
Tertile 2	0.98 (0.83–1.16)
Tertile 3	1.02 (0.88–1.19)
Zip code–level educational attainment (ref. high school or less)	
Some college	1.10 (0.83–1.46)
College graduate	1.00 (0.83–1.19)
Zip code–level composite percent black/percent below FPL (ref. lowest tertile)‡	
Tertile 2	0.97 (0.83–1.14)
Tertile 3	0.91 (0.76–1.08)
Zip code–level overlapping percent black/percent below FPL (ref. group 1)§	
Group 2	0.88 (0.73–1.04)
Group 3	0.87 (0.70–1.07)
State-level Gini coefficient (ref. least income inequality)	
Tertile 2	1.09 (0.84–1.40)
Tertile 3	1.12 (0.86–1.42)

* Adherence is defined as proportion of days covered (PDC) ≥80%. Nonadherence is defined as PDC <80%. Each variable was examined in a separate model adjusted for individual-level factors (see Table 2), calendar year at index date, and random effects at each level. HCQ = hydroxychloroquine; OR = odds ratio; 95% CrI = 95% credible interval; FPL = federal poverty level.
† 95% CrIs do not cross 1 and Bayesian *P* value <0.05.
‡ Composite score of percent black and percent below the FPL divided into tertiles.
§ Group 1 = zip codes with the lowest tertiles of percent black and percent below FPL. Group 3 = zip codes with the highest tertiles of percent black and percent below FPL.

We did not observe effects from concentrated zip code–level poverty or from income inequality. There was a moderate correlation between zip code percent below FPL and percent black ($r^2 = 0.46$, $P < 0.0001$). While our composite measures of individual race and area poverty were in the expected direction, the estimates were not statistically significant. The effect of zip code percent black on adherence persisted after adjusting for percent below FPL, suggesting that other mechanisms beyond socioeconomic deprivation contribute. We hypothesize that we may not see an independent area-level effect of poverty because of the dominant individual-level effect in this low-income population.

We also observed that adjusting for health care resource concentration did not significantly attenuate the relationship between zip code percent black and adherence. However, we could not assess the quality of services provided. Although we did find an association with greater numbers of hospitals and increased odds of adherence, we did not find a parallel asso-

Table 5. Odds of HCQ adherence versus nonadherence according to contextual health resources and urbanicity, accounting for zip code percent black*

Area-level health resources	OR (95% CrI)	OR (95% CrI) for % black
Total MDs per capita (ref. lowest tertile)		
Tertile 2	0.88 (0.70–1.08)	0.86 (0.73–1.01)†
Tertile 3	0.86 (0.71–1.06)	0.82 (0.70–0.98)
Total subspecialists per capita (ref. lowest tertile)		
Tertile 2	0.92 (0.71–1.16)	0.87 (0.75–1.01)†
Tertile 3	0.84 (0.67–1.04)	0.84 (0.71–1.01)†
No. of hospitals per capita (ref. lowest tertile)		
Tertile 2	1.13 (0.97–1.31)	0.87 (0.76–1.00)†
Tertile 3	1.30 (1.07–1.58)†	0.83 (0.70–0.97)†
Health professional shortage area (primary care) (ref. none)		
Partial	1.10 (0.86–1.43)	0.86 (0.74–1.00)†
Whole	0.88 (0.77–1.01)‡	0.83 (0.70–0.99)†
No. of pharmacists per capita (ref. lowest tertile)		
Tertile 2	0.91 (0.75–1.09)	0.85 (0.74–0.99)†
Tertile 3	0.89 (0.75–1.05)	0.81 (0.68–0.95)†
No. of rheumatologists per capita (ref. lowest tertile)		
Tertile 2	1.22 (0.95–1.55)	0.85 (0.73–0.99)†
Tertile 3	1.16 (0.89–1.47)	0.82 (0.69–0.96)†
Percent urban (ref. lowest tertile)		
Tertile 2	0.92 (0.74–1.17)	0.87 (0.75–1.01)‡
Tertile 3	0.82 (0.66–1.02)‡	0.83 (0.70–1.00)†

* Adherence is defined as proportion of days covered (PDC) ≥80%. Nonadherence is defined as PDC <80%. Each variable was examined in a separate model adjusted for individual-level factors (see Table 2), calendar year at index date, zip code–level percent black, and random effects at each level. All variables are at the county level except for number of rheumatologists, which is at the state level. HCQ = hydroxychloroquine; OR = odds ratio; 95% CrI = 95% credible interval.
† Bayesian *P* value <0.05 with 95% CrI that just crosses 1.
‡ 95% CrIs do not cross 1 and Bayesian *P* value <0.05.

ciation with more health care providers. Hospital-based clinics provide care to medically complex patients, whereas private practice physicians may not accept Medicaid or sicker patients. Adherence may be more related to the quality of care received and patient–provider interactions than to the concentration of services available. Alternatively, there might be more physicians who practice in areas with high concentrations of medically complex patients, and therefore adherence may be poorer overall in these areas to begin with.

Notably, we found only small between-area variation in adherence despite our expectation that this would be greater due to variation in Medicaid eligibility criteria by state. However, variations in poverty-level eligibility among overall poor individuals may not be that significant. It is also possible that we did not see significant differences because HCQ co-payments and quantity limit between the states were similar. It is possible that at the zip code and county levels, the lack of between-area variation observed was a result of a significant proportion of areas with few SLE patients and the high prevalence of overall nonadherence in this population.

This study had a number of strengths. We used multilevel models to examine the relationship across contextual characteristics with adherence while accounting for individual-level factors as well as for potential clustering and heterogeneity by geographic area. To our knowledge, this is the first study to examine the role of contextual factors on adherence in a high-risk SLE population.

Researchers have highlighted that among patients with SLE, we need a better understanding of the contribution of social determinants to disparities in health outcomes; however, few studies have examined this (49). We propose that our findings are hypothesis-generating and provide impetus to explore the mechanisms behind the associations uncovered.

Our study also has limitations. We excluded individuals older than age 65 years because of possible dual eligibility in Medicare. We did not exclude younger dual-eligible individuals because we did not have complete data available. Although we likely captured most prescription claims for this population, adherence may be underestimated for those who changed their drug coverage after the index date. Although low income is part of the criteria for Medicaid eligibility, we did not have other individual-level measures of SES (e.g., occupation and education). In addition, we lacked more granular geographic area data and therefore were unable to assess direct neighborhood effects or geocode our data. Our use of claims data did not enable us to directly measure SLE disease activity; however, SLE-related laboratory tests and the SLE risk adjustment index were used to approximate this.

Our study may also have excluded patients with more severe SLE and lupus nephritis, because we required new use of HCQ but not a new diagnosis of SLE. Although this is not a cohort of patients with incident SLE, we suspect that many of the patients included are in the early stages of their disease course, which is reflected by the low mean SLE risk adjustment index and the low prevalence of lupus nephritis and immunosuppressive medication use despite the vulnerable nature of this population. Due to the nature of claims, there may also be misclassification of patients with SLE and of comorbidities. Certain variables such as obesity, vaccinations, and smoking are also underestimated due to underreporting. It is also possible that antidepressant use may capture both SLE patients with chronic pain and those with depression, because some antidepressants are used to treat both, although these are often comorbid conditions. We measured adherence using the PDC, which reflects drug dispensing and refills; however, filling a medication prescription may not always translate to taking the medication in the manner in which it was prescribed. In addition, the PDC is an aggregate measure over the course of a year and may not capture the dynamic nature of adherence behavior.

In this study of Medicaid beneficiaries with SLE, adherence to HCQ treatment was poor. In addition to reaffirming the role of certain individual-level sociodemographic and disease-related factors on adherence, we propose that contextual influences contribute as well. Our findings should pave the way for further work examining the importance of social determinants, including racial residential segregation and racial discrimination as well as neighborhood-specific health care quality, on health behaviors and outcomes in vulnerable populations.

ACKNOWLEDGMENTS

The authors would like to thank Yongjoo Kim, MD, ScD and Lulu Pan for their assistance with programming.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Feldman, Costenbader, Solomon, Subramanian, Kawachi.

Acquisition of data. Feldman, Costenbader.

Analysis and interpretation of data. Feldman, Costenbader, Solomon, Subramanian, Kawachi.

REFERENCES

1. Feldman CH, Hiraki LT, Liu J, Fischer MA, Solomon DH, Alarcon GS, et al. Epidemiology and sociodemographics of systemic lupus erythematosus and lupus nephritis among US adults with Medicaid coverage, 2000–2004. Arthritis Rheum 2013;65:753–63.
2. Lim SS, Bayakly AR, Helmick CG, Gordon C, Easley KA, Drenkard C. The incidence and prevalence of systemic lupus erythematosus, 2002–2004: the Georgia Lupus Registry. Arthritis Rheumatol 2014;66:357–68.
3. Petri M, Perez-Gutthann S, Longenecker JC, Hochberg M. Morbidity of systemic lupus erythematosus: role of race and socioeconomic status. Am J Med 1991;91:345–53.
4. Gomez-Puerta JA, Barbhaiya M, Guan H, Feldman CH, Alarcon GS, Costenbader KH. Racial/ethnic variation in all-cause mortality among United States Medicaid recipients with systemic lupus erythematosus: a Hispanic and Asian paradox. Arthritis Rheumatol 2015;67:752–60.
5. 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.
6. Yelin E, Trupin L, Yazdany J. A prospective study of the impact of current poverty, history of poverty, and exiting poverty on accumulation of disease damage in systemic lupus erythematosus. Arthritis Rheumatol 2017;69:1612–22.
7. Walker RJ, Strom Williams J, Egede LE. Influence of race, ethnicity and social determinants of health on diabetes outcomes. Am J Med Sci 2016;351:366–73.
8. Williams LK, Joseph CL, Peterson EL, Moon C, Xi H, Krajenta R, et al. Race-ethnicity, crime, and other factors associated with adherence to inhaled corticosteroids. J Allergy Clin Immunol 2007;119:168–75.
9. Rolnick SJ, Pawloski PA, Hedblom BD, Asche SE, Bruzek RJ. Patient characteristics associated with medication adherence. Clin Med Res 2013;11:54–65.
10. Tang C, Godfrey T, Stawell R, Nikpour M. Hydroxychloroquine in lupus: emerging evidence supporting multiple beneficial effects. Intern Med J 2012;42:968–78.
11. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15:351–77.
12. August KJ, Billimek J. A theoretical model of how neighborhood factors contribute to medication nonadherence among disadvantaged chronically ill adults. J Health Psychol 2016;21:2923–33.

13. Loewenstein G, Brennan T, Volpp KG. Asymmetric paternalism to improve health behaviors. JAMA 2007;298:2415–7.
14. O'Donoghue T, Rabin M. Doing it now or later. Am Econ Rev 1999;89:103–24.
15. Brown AF, Ang A, Pebley AR. The relationship between neighborhood characteristics and self-rated health for adults with chronic conditions. Am J Public Health 2007;97:926–32.
16. Ludwig J, Duncan GJ, Gennetian LA, Katz LF, Kessler RC, Kling JR, et al. Neighborhood effects on the long-term well-being of low-income adults. Science 2012;337:1505–10.
17. Bernatsky S, Linehan T, Hanly JG. The accuracy of administrative data diagnoses of systemic autoimmune rheumatic diseases. J Rheumatol 2011;38:1612–6.
18. Steiner JF, Prochazka AV. The assessment of refill compliance using pharmacy records: methods, validity, and applications. J Clin Epidemiol 1997;50:105–16.
19. Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. Prospective validation of eight different adherence measures for use with administrative claims data among patients with schizophrenia. Value Health 2009;12:989–95.
20. Yeaw J, Benner JS, Walt JG, Sian S, Smith DB. Comparing adherence and persistence across 6 chronic medication classes. J Manag Care Pharm 2009;15:728–40.
21. Williams DR, Jackson JS. Race/ethnicity and the 2000 census: recommendations for African American and other black populations in the United States. Am J Public Health 2000;90:1728–30.
22. Ward MM. Development and testing of a systemic lupus-specific risk adjustment index for in-hospital mortality. J Rheumatol 2000;27:1408–13.
23. Chibnik LB, Massarotti EM, Costenbader KH. Identification and validation of lupus nephritis cases using administrative data. Lupus 2010;19:741–3.
24. 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.
25. Noyes K, Liu H, Lyness JM, Friedman B. Medicare beneficiaries with depression: comparing diagnoses in claims data with the results of screening. Psychiatr Serv 2011;62:1159–66.
26. Patrick AR, Shrank WH, Glynn RJ, Solomon DH, Dormuth CR, Avorn J, et al. The association between statin use and outcomes potentially attributable to an unhealthy lifestyle in older adults. Value Health 2011;14:513–20.
27. Minnesota Population Center. National Historical Geographic Information System: Version 11.0 [Database]. University of Minnesota, Minneapolis. URL: https://doi.org/10.18128/d050.v11.0.
28. The U.S. Census Bureau. 2006-2010 American Community Survey. U.S. Census Bureau's American Community Survey Office. 2010. URL: https://www.census.gov/programs-surveys/acs/.
29. Manson S, Schroeder J, Van Riper D, Ruggles S. IPUMS National Historical Geographic Information System, Version 12.0 [Database]. Minneapolis: University of Minnesota; 2017. URL: http//doi.org/10.18128.D050.V12.0.
30. U.S. Health Resources and Services Administration Data Warehouse. Area Health Resources Files. U.S. Department of Health and Human Services; 2017. URL: https://data.hrsa.gov.
31. U.S. Health Resources and Services Administration Data Warehouse. Bureau of Health Professions. 2002. URL: https://ciser.cornell.edu/ASPs/search_athena.asp?IDTITLE=2044.
32. Subramanian SV, O'Malley AJ. Modeling neighborhood effects: the futility of comparing mixed and marginal approaches. Epidemiology 2010;21:475–8.
33. Browne WJ, Draper D. A comparison of Bayseian and likelihood-based methods for fitting multilevel models. Bayesian Analysis 2006;1:472–514.
34. Browne WJ. MCMC estimation in MLwiN. Center for Multilevel Modeling, University of Bristol, UK. 2015. URL: http://www.bris.ac.uk/cmm/media/software/mlwin/downloads/manuals/2-32/mcmc-web.pdf.
35. Rasbash J, Charlton C, Browne WJ, Healy M, Cameron B. MLwiN Version 2.36. (2015) Centre for Multilevel Modelling: University of Bristol.
36. Feldman CH, Collins J, Zhang Z, Xu C, Subramanian SV, Kawachi I, et al. Azathioprine and mycophenolate mofetil adherence patterns and predictors among Medicaid beneficiaries with systemic lupus erythematosus. Arthritis Care Res (Hoboken). 2018 Oct 24. doi: 10.1002/acr.23792. [Epub ahead of print]
37. Julian LJ, Yelin E, Yazdany J, Panopalis P, Trupin L, Criswell LA, et al. Depression, medication adherence, and service utilization in systemic lupus erythematosus. Arthritis Rheum 2009;61:240–6.
38. Mosley-Williams A, Lumley MA, Gillis M, Leisen J, Guice D. Barriers to treatment adherence among African American and white women with systemic lupus erythematosus. Arthritis Rheum 2002;47:630–8.
39. Garcia-Gonzalez A, Richardson M, Garcia Popa-Lisseanu M, Cox V, Kallen MA, Janssen N, et al. Treatment adherence in patients with rheumatoid arthritis and systemic lupus erythematosus. Clin Rheumatol 2008;27:883–9.
40. 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.
41. Fessler BJ, Alarcon GS, McGwin G Jr, Roseman J, Bastian HM, Friedman AW, et al. Systemic lupus erythematosus in three ethnic groups: XVI. Association of hydroxychloroquine use with reduced risk of damage accrual. Arthritis Rheum 2005;52:1473–80.
42. Pons-Estel GJ, Alarcon GS, Gonzalez LA, Zhang J, Vila LM, Reveille JD, et al. Possible protective effect of hydroxychloroquine on delaying the occurrence of integument damage in lupus: LXXI, data from a multiethnic cohort. Arthritis Care Res (Hoboken) 2010;62:393–400.
43. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter? The Public Health Disparities Geocoding Project. Am J Epidemiol 2002;156:471–82.
44. Acevedo-Garcia D, Lochner KA. Residential segregation and health. In: Kawachi I, Berkman L, editors. Neighborhoods and health. New York: Oxford University Press; 2003.
45. Williams DR. Race, socioeconomic status, and health. The added effects of racism and discrimination. Ann N Y Acad Sci 1999;896:173–88.
46. Vina ER, Hausmann LR, Utset TO, Masi CM, Liang KP, Kwoh CK. Perceptions of racism in healthcare among patients with systemic lupus erythematosus: a cross-sectional study. Lupus Sci Med 2015;2:e000110.
47. Vina ER, Masi CM, Green SL, Utset TO. A study of racial/ethnic differences in treatment preferences among lupus patients. Rheumatology (Oxford) 2012;51:1697–706.
48. Cuffee YL, Hargraves JL, Rosal M, Briesacher BA, Schoenthaler A, Person S, et al. Reported racial discrimination, trust in physicians, and medication adherence among inner-city African Americans with hypertension. Am J Public Health 2013;103:e55–62.
49. Williams EM, Ortiz K, Browne T. Social determinants of health, the chronic care model, and systemic lupus erythematosus. Int J Chronic Dis 2014;2014:361792.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 914–924
DOI 10.1002/acr.23837

Impact and Timing of Smoking Cessation on Reducing Risk of Rheumatoid Arthritis Among Women in the Nurses' Health Studies

Xinyi Liu,[1] Sara K. Tedeschi,[2] Medha Barbhaiya,[3] Cianna L. Leatherwood,[2] Cameron B. Speyer,[1] Bing Lu,[2] Karen H. Costenbader,[2] Elizabeth W. Karlson,[2] and Jeffrey A. Sparks[2]

Objective. To investigate the impact and timing of smoking cessation on developing rheumatoid arthritis (RA) and serologic phenotypes.

Methods. We investigated smoking cessation and RA risk in the Nurses' Health Study (NHS) (1976–2014) and the NHS II (1989–2015). Smoking exposures and covariates were obtained by biennial questionnaires. Self-reported RA was confirmed by medical record review for American College of Rheumatology/European League Against Rheumatism criteria. Cox regression was used to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs) for RA serologic phenotypes (all, seropositive, seronegative) according to smoking status, intensity, pack-years, and years since cessation.

Results. Among 230,732 women, we identified 1,528 incident cases of RA (63.4% of which were seropositive) during 6,037,151 person-years of follow-up. Compared with never smoking, current smoking increased the risk of all RA (multivariable HR 1.47, 95% CI 1.27–1.72) and seropositive RA (HR 1.67, 95% CI 1.38–2.01) but not seronegative RA (HR 1.20, 95% CI 0.93–1.55). An increasing number of smoking pack-years was associated with an increased trend for the risk of all RA ($P < 0.0001$) and seropositive RA ($P < 0.0001$). With increasing duration of smoking cessation, a decreased trend for the risk of all RA was observed ($P = 0.009$) and seropositive RA ($P = 0.002$). Compared to recent quitters (<5 years), those who quit ≥30 years ago had an HR of 0.63 (95% CI 0.44–0.90) for seropositive RA. However, a modestly increased risk of RA was still detectable 30 years after quitting smoking (for all RA, HR 1.25 [95% CI 1.02–1.53]; for seropositive RA, HR 1.30 [95% CI 1.01–1.68]; reference, never smoking).

Conclusion. These results confirm that smoking is a strong risk factor for developing seropositive RA and demonstrate for the first time that a behavior change of sustained smoking cessation could delay or even prevent seropositive RA.

INTRODUCTION

Although the etiology of rheumatoid arthritis (RA) remains obscure, previous studies have implicated smoking as an important and potentially modifiable risk factor for the development of RA (1–6). Previous epidemiologic studies have identified cigarette smoking as one of the most important lifestyle risk factors for the development of RA and particularly seropositive RA, which is defined as the presence of rheumatoid factor (RF) or anti–cyclic citrullinated peptide (anti-CCP) (5,7). Smoking may affect the risk of seropositive RA by inducing local tissue inflammation, promoting citrullination, and forming neoepitopes, resulting in autoimmunity (8). Smoking also induces immune cells to secrete proinflammatory cytokines, resulting in a systemic inflammatory state (8–10). Even though smoking cessation may decrease the level of systemic inflammation, other components of the immune system may be permanently altered after autoimmunity is established once a threshold of smoking is reached (11). Although there is strong evidence that ever-smokers (current or past) have a higher risk of seropositive RA compared to never-smokers, it is unclear whether

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of Harvard University, its affiliated academic health care centers, or the National Institutes of Health.

Supported by NIH grants K23-AR-069688, L30-AR-066953, R01-AR-049880, K24-AR-052403, UM1-CA-186107, UM1-CA-176726, P30-AR-070253, and P30-AR-072577.

[1]Xinyi Liu, MS, Cameron B. Speyer, BA: Brigham and Women's Hospital, Boston, Massachusetts; [2]Sara K. Tedeschi, MD, MPH, Cianna L. Leatherwood, MD, MPH, Bing Lu, MD, DrPH, Karen H. Costenbader, MD, MPH, Elizabeth W. Karlson, MD, MS, Jeffrey A. Sparks, MD, MMSc: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [3]Medha Barbhaiya, MD, MPH: Hospital for Special Surgery, New York, New York.

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

Address correspondence to Jeffrey A. Sparks, MD, MMSc, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, 60 Fenwood Road, #6016U, Boston, MA 02115. E-mail: jsparks@bwh.harvard.edu.

Submitted for publication December 3, 2018; accepted in revised form January 22, 2019.

SIGNIFICANCE & INNOVATIONS

- We used data from 2 large prospective cohorts comprised of up to 38 years of follow-up to investigate smoking cessation and RA risk. We identified 1,528 incident cases of RA. We used data from 2 large prospective cohort comprised of up to 38 years of follow-up to investigate smoking cessation and RA risk. We identified 1,528 incident RA cases by medical record review during over 6 million person-years of follow-up.
- Compared to never-smokers, a slightly elevated risk for seropositive RA was still detectable among past-smokers even 30 years after smoking cessation.
- Among past-smokers, the risk of seropositive RA was significantly reduced by 37% (hazard ratio 0.63, 95% confidence interval 0.44–0.90) in women who quit smoking for ≥30 years compared to women who quit smoking for 0 to <5 years.
- Our findings demonstrate that a behavior change of sustained smoking cessation may reduce the risk of seropositive RA.

smoking cessation actually reduces the risk for past-smokers, perhaps to the level of a never-smoker, after long-term cessation.

Previous studies have investigated the association between smoking status, intensity, duration, and pack-years and the risk of RA (2,3,5,7,12–14). However, only a few studies have investigated whether smoking cessation might reduce the risk of developing RA (5,15–17). A previous investigation in the Nurses' Health Study (NHS) suggested that the risk of RA in past-smokers was similar to the risk in never-smokers after 20 years smoking cessation (5). However, a true increased risk of RA after long-term sustained smoking cessation may have been difficult to detect given sample sizes and follow-up limitations. Another study in Sweden suggested a reduced risk of RA with increased duration of smoking compared to current smoking, but these results were not statistically significant (15).

To investigate the association between smoking cessation and RA risk, repeated measures of smoking in a large sample with lengthy follow-up prior to RA diagnosis are required. Therefore, we used 2 longitudinal cohorts of female nurses, the NHS and the NHS II, to investigate smoking cessation and RA risk, using up to 38 years of prospective follow-up data. The aims of this study were to investigate the association between smoking cessation and the risk of RA overall and according to serologic phenotype. We also sought to determine when and to what extent smoking cessation may reduce the increased risk of developing RA overall and according to serologic phenotype. We hypothesized that smoking cessation would reduce the risk of seropositive RA, but that residual RA risk would remain elevated compared to that in never-smokers for many years following smoking cessation.

PARTICIPANTS AND METHODS

Study population. The NHS and NHS II are prospective cohort studies of US female registered nurses. Participants completed baseline and biennial questionnaires regarding lifestyle, health behaviors, medications, and diseases. The NHS began in 1976 and enrolled 121,700 nurses ages 30–55 years; the NHS II was established in 1989 and enrolled 116,430 nurses ages 25–42 years. Both cohorts have >90% follow-up response rates, and only 5% of person-time has been lost to follow-up (18).

For this analysis, we excluded women who reported RA and other connective tissue diseases (CTDs) at baseline, were missing baseline smoking information, or answered only the baseline questionnaire. After exclusions, 117,182 NHS participants who were followed from 1976 to 2014 and 113,550 NHS II participants followed from 1989 to 2015 were analyzed. All participants provided informed consent, and this study was approved by the Partners HealthCare institutional review board.

Smoking exposures. On the baseline questionnaire, the women reported smoking status (never/past/current) and the age at which they began to smoke. Current-smokers were asked the number of cigarettes typically smoked per day, and past-smokers provided the age at which they stopped smoking and the number of cigarettes smoked per day before quitting. On subsequent questionnaires, the women reported smoking status and intensity (1–14, 15–24, ≥25 cigarettes/day). Data for smoking pack-years were derived by multiplying the number of packs of cigarettes smoked per day (20 cigarettes/pack) by the number of years of smoking. Because smokers often stop and restart smoking, all smoking exposure variables were time-updated.

Identification of incident RA. RA cases were identified by a 2-stage procedure. Participants who self-reported a new diagnosis of RA were contacted by mailing the connective tissue disease screening questionnaire (CSQ) (19). The medical records of participants with positive scoring on the CSQ were requested and reviewed independently by 2 rheumatologists to identify RA cases meeting the 1987 American College of Rheumatology (ACR) or 2010 ACR/European League Against Rheumatism classification criteria (20,21). In addition to components of these classification criteria, dates of symptom onset/diagnosis and clinical laboratory results for RF and anti-CCP tests were collected. Therefore, cases had confirmed incident RA, with documented serologic phenotype from medical records. For women who were diagnosed with RA prior to the clinical use of anti-CCP assays in the early 2000s, serologic phenotype was determined solely by RF status based on medical record review. A subset of cases with blood banked prior to or after the date of RA onset had anti-CCP tested for research purposes; therefore, we reclassified a few women as seropositive (n = 23). Thus, partici-

pants were classified as seropositive if RF or anti-CCP assays were positive and seronegative if both RF and anti-CCP assays (if available) were negative.

Covariates. We selected covariates as potential confounders associated with cigarette smoking and RA based on previous studies (2,7,12). Time-updated sociodemographic covariates included age, race, and household income (categorized by quartile of US Census tract–based median household income at the zip code level). Potential time-updated reproduction confounders were oral contraceptive use (categorized as ever-users or never-users), parity/total breastfeeding duration, menopausal status, and postmenopausal hormone (PMH) use. We used a combined "parity/total breastfeeding duration" variable categorized as nulliparous, parous/0 to <1 month, parous/1–11 months, or parous/≥12 months, and a combined variable for menopausal status and PMH use categorized as premenopausal, postmenopausal/never, or postmenopausal/ever. We defined time-updated sedentary physical activity as <3 metabolic equivalent of task (MET) hours/week (22) and categorized time-updated body mass index (BMI) as <25.0, 25.0 to <30.0, or ≥30.0 kg/m^2. Alcohol consumption (also time-updated) was assessed by a semiquantitative food frequency questionnaire (23). We calculated cumulative average alcohol as a long-term measure of intake and categorized as

Table 1. Age-standardized characteristics of participants in the Nurses' Health Study (NHS) in 1988 and NHS II in 1989 categorized by smoking status*

	NHS (n = 98,497)†			NHS II (n = 113,550)		
	Never-smoker	Past-smoker	Current-smoker	Never-smoker	Past-smoker	Current-smoker
Participants, no. (%)	44,776 (45.5)	35,238 (35.8)	18,483 (18.8)	74,181 (65.3)	24,155 (21.3)	15,214 (13.4)
Age, mean ± SD years‡	54.2 ± 7.3	54.5 ± 7.1	53.9 ± 6.9	34.0 ± 4.7	35.2 ± 4.5	34.8 ± 4.6
White race	92.7	94.6	94.7	91.7	94.9	93.2
Median household income quartile						
Quartile 1 (lowest)	29.0	24.4	26.0	24.3	21.2	28.4
Quartile 2	24.9	23.2	24.5	25.5	23.0	25.4
Quartile 3	23.3	25.2	25.6	25.1	26.0	24.4
Quartile 4 (highest)	22.7	27.2	23.9	25.1	29.9	21.7
Body mass index category						
<25.0 kg/m^2	54.0	54.6	63.3	70.3	70.4	69.5
25.0 to <30.0 kg/m^2	29.6	29.5	26.3	18.3	18.6	19.2
≥30.0 kg/m^2	16.4	15.9	10.3	11.4	11.0	11.3
Sedentary physical activity (<3 METs/week)	19.6	19.0	27.3	14.7	14.0	18.0
Parity/breastfeeding duration, months						
Nulliparous	5.2	5.2	5.6	30.1	28.0	33.9
Parous/none to <1	25.8	30.2	35.4	11.8	12.0	16.1
Parous/1–11	27.2	27.8	26.5	22.1	24.2	22.6
Parous/≥12	17.6	14.1	11.0	26.2	25.3	14.0
Menopausal status/PMH use						
Premenopausal	38.9	36.6	35.4	94.3	94.3	91.6
Postmenopausal/never	28.0	27.3	32.6	2.8	2.8	4.0
Postmenopausal/ever	33.1	36.0	31.9	2.9	2.9	4.3
Cumulative average alcohol intake, gm/day						
None to <5	71.4	53.9	52.8	84.2	71.0	67.3
5 to <10	9.8	15.3	12.5	8.7	14.5	13.8
≥10	10.7	23.5	28.4	6.0	13.7	18.0

* Missing values are not shown. METs = metabolic equivalents of task; PMH = postmenopausal hormone. Except where indicated otherwise, values are the percent.
† A total of 117,182 women were in the NHS at baseline in 1976.
‡ Not age-standardized.

none to <5, 5 to <10, and ≥10 gm/day. Missing data for physical activity, BMI, and alcohol were carried forward 1 cycle, and then a missing indicator variable was created for data missing beyond 1 cycle.

Statistical analysis. We pooled data from the NHS and NHS II into a single analysis for statistical efficiency, given exposures with many categories, planned subgroup analyses, and analyses for RA serologic phenotypes. We reported age-adjusted descriptive statistics for covariates across smoking status categories (never/past/current) for the NHS in 1988 and the NHS II in 1989, because these were the first cycles at similar calendar periods.

Person-years of follow-up for each participant accrued from the date of return of the baseline questionnaire to the date of censoring, whichever came first: RA diagnosis, reported other CTD not confirmed as RA, date of death, or end of follow-up for this analysis (June 1, 2014 for the NHS and June 1, 2015 for the NHS II). If participants were missing smoking data during a questionnaire cycle, we did not include these person-years in the analysis.

We used Cox proportional hazards models to test for the association between time-updated smoking intensity according to smoking status, pack-years, and smoking cessation and RA risk. We analyzed smoking status as never (reference), past, or current as well as a 5-level smoking intensity variable of never (reference), past, current 1–14, current 15–24, and current ≥25 cigarettes/day. Pack-years of smoking were categorized as never (reference), >0–10, >10–20, >20–30, >30–40, or >40 pack-years. Grouped by smoking status and years since quitting, smoking cessation was analyzed as never (reference), past (in ordinal categories of years since cessation [≥30, 20 to <30, 10 to <20, 5 to <10, or 0 to <5 years ago]), and current. Base models were adjusted for age, cohort, and questionnaire cycle (each cohort was pooled by similar calendar times; e.g., the 1988 cycle in the NHS was pooled with the 1989 cycle in the NHS II). After we examined the associations of each possible covariate with smoking status and the all-RA outcome separately, the multivariable model was additionally adjusted for oral contraceptive use, parity/breastfeeding, menopausal status/PMH use, BMI, sedentary physical activity, median household income, and alcohol intake. We performed similar analyses for RA serologic phenotypes. We also analyzed additional subgroups categorized as ever-smokers (reference, current smoking) and past-smokers (reference, 0 to <5 years since quitting) to further investigate smoking cessation and RA risk.

We used restricted cubic splines with 3 knots to visualize the risk for RA serologic phenotypes by pack-years (among the entire study sample; reference = never [0] pack-years) and years since quitting smoking (among only past-smokers; reference ≤2 years since cessation) adjusted for covariates (24). Tests for nonlinearity used the likelihood ratio test, comparing the model with only the linear term to the model with the linear and the cubic spline terms. Use of higher numbers of knots in the cubic spline curves showed similar results.

We tested the proportional hazards assumption by including interaction terms between smoking exposures and follow-up time, using likelihood ratio tests to compare nested models with and without interaction terms. The proportional hazards assumption was met in all analyses. Two-sided *P* values less than 0.05 were considered significant in all analyses. All analyses were performed using SAS v.9.4.

Table 2. Hazard ratios for rheumatoid arthritis (RA) serologic phenotypes according to smoking status and intensity*

			Current-smoker		
	Never-smoker	Past-smoker	1–14 cigarettes/day	15–24 cigarettes/day	≥25 cigarettes/day
All RA					
Cases/person-years	675/3,262,927	597/1,946,910	80/316,907	111/330,369	65/180,038
Age-adjusted model	1.00 (Ref.)	1.35 (1.20–1.51)	1.19 (0.94–1.50)	1.58 (1.28–1.94)	1.69 (1.30–2.19)
Multivariable model	1.00 (Ref.)	1.36 (1.21–1.53)	1.23 (0.97–1.55)	1.60 (1.30–1.97)	1.69 (1.30–2.20)
Seropositive RA					
Cases/person-years	415/3,254,327	386/1,940,936	52/315,778	73/329,167	43/179,214
Age-adjusted model	1.00 (Ref.)	1.44 (1.25–1.66)	1.29 (0.97–1.73)	1.75 (1.35–2.25)	1.86 (1.35–2.58)
Multivariable model	1.00 (Ref.)	1.48 (1.28–1.71)	1.36 (1.01–1.82)	1.80 (1.39–2.34)	1.92 (1.39–2.66)
Seronegative RA					
Cases/person-years	260/3,254,901	211/1,940,814	28/315,756	38/328,645	22/178,974
Age-adjusted model	1.00 (Ref.)	1.20 (1.00–1.44)	1.03 (0.70–1.53)	1.32 (0.94–1.88)	1.42 (0.91–2.21)
Multivariable model	1.00 (Ref.)	1.18 (0.98–1.43)	1.03 (0.69–1.53)	1.31 (0.92–1.86)	1.36 (0.86–2.12)

* Multivariable models were adjusted for age, questionnaire period, cohort, oral contraceptive use (ever, never), parity/breastfeeding in months (nulliparous, parous/<1 month, parous/1–11 months, parous/≥12 months), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/never, postmenopausal/ever), body mass index category (underweight/normal, overweight, obese), sedentary physical activity, median household income (quartiles), alcohol intake (none to <5 gm/day, 5 to <10 gm/day, ≥10 gm/day). Except where indicated otherwise, values are the hazard ratio (95% confidence interval).

RESULTS

Characteristics of the participants. Among 117,182 women in the NHS (1976–2014) and 113,550 women in the NHS II (1989–2015), we identified a total of 1,528 incident cases of RA (1,002 in the NHS and 526 in the NHS II). There were 969 (63.4%) seropositive RA cases and 559 (36.6%) seronegative RA cases during 6,037,151 person-years of follow-up. Table 1 shows age-adjusted characteristics of the NHS and the NHS II study participants categorized according to smoking status and at a similar calendar time (1988 and 1989, respectively). Women in the NHS were older in 1988 (mean ± SD age 54.3 ± 7.2 years) compared to women in the NHS II in 1989 (mean ± SD age 34.4 ± SD 4.7 years). There were more smokers in the NHS (18.8% current-smokers and 35.8% past-smokers) than in the NHS II (13.4% current-smokers and 21.3% past-smokers). Within both cohorts, sedentary physical activity and alcohol consumption were higher among smokers than non-smokers, particularly for current-smokers.

Smoking status/intensity and RA risk. Compared to women who never smoked, the multivariable-adjusted hazard ratio (HR) for developing RA was 1.36 (95% CI 1.22–1.53) among past-smokers and 1.46 (95% CI 1.26–1.70) among current-smokers (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23837/abstract), after adjustment for age, questionnaire period, cohort, oral contraceptive use, parity/breastfeeding, menopausal status/PMH use, BMI, sedentary physical activity, median household income, and alcohol intake. When we further examined the intensity of smoking, current-smokers who smoked ≥25 cigarettes per day had a 92% increased risk of seropositive RA compared to never-smokers (multivariable HR 1.92, 95% CI 1.39–2.66) (Table 2). There was no significant association between smoking intensity and seronegative RA.

Pack-years and RA risk. To further investigate the association between smoking and RA risk, we investigated the relationship between pack-years of smoking and RA risk, using restricted cubic splines models (Figure 1). Compared to never-smokers (0 pack-years), there was a statistically significant increasing trend for developing RA with increasing number of pack-years smoked up to 35 pack-years, with the HR plateauing at approximately 1.8 for all RA ($P < 0.0001$) and an HR of 2.3 for seropositive RA ($P < 0.0001$).

Supplementary Table 2 (available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23837/abstract) shows the categories of pack-years and RA risk. Compared to the risk for never-smokers, there was no

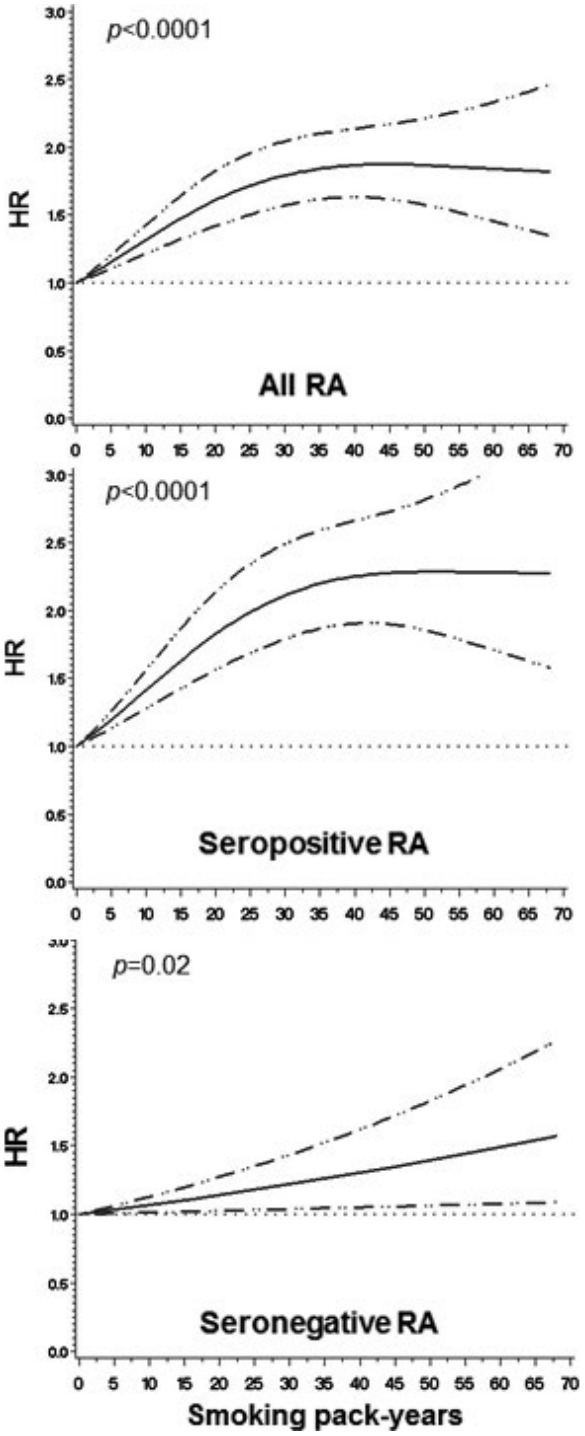


Figure 1. Restricted cubic spline curves showing hazard ratios (HRs) (solid lines) and 95% confidence intervals (95% CIs) (broken lines) for rheumatoid arthritis (RA) serologic phenotypes among all women with RA, according to pack-years of smoking (reference group, 0 pack-years). Curves were adjusted for the covariates listed in Table 2. Supplementary Table 2 (available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23837/abstract) shows data analyzed by categories of smoking pack-years. *P* values are for trend.

Table 3. Hazard ratios for rheumatoid arthritis (RA) serologic phenotypes according to smoking status and years since cessation*

		Past-smoked					
	Never-smoker	Quit ≥30 years ago	Quit 20 to <30 years ago	Quit 10 to <20 years ago	Quit 5 to <10 years ago	Quit 0 to <5 years ago	Current-smoker
All RA							
Cases/person-years	675/3,262,927	123/397,373	129/445,185	160/551,108	90/262,059	94/282,440	257/836,060
Age-adjusted model	1.00 (Ref.)	1.22 (1.00–1.50)	1.17 (0.97–1.42)	1.36 (1.14–1.62)	1.65 (1.32–2.06)	1.58 (1.27–1.96)	1.45 (1.25–1.68)
Multivariable model	1.00 (Ref.)	1.25 (1.02–1.53)	1.19 (0.99–1.45)	1.37 (1.15–1.64)	1.63 (1.30–2.04)	1.57 (1.26–1.95)	1.47 (1.27–1.72)
Seropositive RA							
Cases/person-years	415/3,254,327	77/396,579	82/443,916	99/549,283	58/261,118	69/281,311	169/832,888
Age-adjusted model	1.00 (Ref.)	1.26 (0.97–1.62)	1.21 (0.95–1.53)	1.42 (1.13–1.77)	1.78 (1.35–2.35)	1.95 (1.51–2.53)	1.60 (1.33–1.93)
Multivariable model	1.00 (Ref.)	1.30 (1.01–1.68)	1.25 (0.98–1.59)	1.45 (1.16–1.81)	1.79 (1.36–2.37)	1.99 (1.54–2.58)	1.67 (1.38–2.01)
Seronegative RA							
Cases/person-years	260/3,254,901	46/396,731	47/443,877	61/549,236	32/261,042	25/281,204	88/832,099
Age-adjusted model	1.00 (Ref.)	1.17 (0.84–1.63)	1.12 (0.82–1.53)	1.28 (0.97–1.70)	1.45 (1.00–2.10)	1.03 (0.68–1.55)	1.22 (0.95–1.56)
Multivariable model	1.00 (Ref.)	1.18 (0.85–1.64)	1.11 (0.81–1.53)	1.25 (0.94–1.67)	1.41 (0.97–2.04)	0.98 (0.65–1.49)	1.20 (0.93–1.55)

* Multivariable models were adjusted for age, questionnaire period, cohort, oral contraceptive use (ever, never), parity/breastfeeding in months (nulliparous, parous/<1 month, parous/1–11 months, parous/≥12 months), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/never, postmenopausal/ever), body mass index category (underweight/normal, overweight, obese), sedentary physical activity, median household income (quartiles), alcohol intake (none to <5 gm/day, 5 to <10 gm/day, ≥10 gm/day). Except where indicated otherwise, values are the hazard ratio (95% confidence interval).

risk for any RA serologic phenotype for >0 to 10 pack-years. However, those who smoked for 10–20 pack-years had a significantly increased risk of all RA (HR 1.38, 95% CI 1.17–1.64) and seropositive RA (HR 1.54, 95% CI 1.25–1.89) but not seronegative RA. Those who smoked for >40 pack-years had a nearly 2-fold risk of developing all RA (HR 1.83, 95% CI 1.52–2.20) and seropositive RA (HR 2.25, 95% CI 1.80–2.82) but not seronegative RA (HR 1.27, 95% CI 0.92–1.74).

Smoking cessation and RA risk. Table 3 shows the associations of smoking cessation with RA, comparing current-smokers and past-smokers according to years since quitting to the reference group of never-smokers. Compared to never-smokers, past-smokers who quit 0 to <5 years ago had a significantly increased risk of all RA (HR 1.57, 95% CI 1.26–1.95) and seropositive RA (HR 1.99, 95% CI 1.54–2.58) but not seronegative RA (HR 0.98, 95% CI 0.65–1.49). This point of estimate of increased risk of developing RA started to decline among past-smokers who quit 10 to <20 years ago (HR 1.37, 95% CI 1.15–1.64) for all RA and those who quit 5 to <10 years ago (HR 1.79, 95% CI 1.36–2.37) for seropositive RA. However, modestly increased RA risk was still detectable even 30 years after quitting smoking for both all RA (HR 1.25, 95% CI 1.02–1.53) and seropositive RA (HR 1.30, 95% CI 1.01–1.68).

Figure 2 shows the restricted cubic spline curves among the subset of past-smokers for the association between RA and years since smoking cessation (reference, 0–2 years since quitting). There was a trend showing a statistically significant decreasing risk for developing RA with increasing years since smoking cessation for both all RA (P = 0.009) and seropositive RA (P = 0.002) but not seronegative RA (P = 0.78).

Table 4 shows categories of years since smoking cessation and RA risk. Women who quit smoking ≥30 years ago had a suggestive reduced risk (HR 0.78, 95% CI 0.58–1.05) for all RA compared to those who quit 0 to <5 years ago. The risk of seropositive RA was significantly reduced by 37% (HR 0.63, 95% CI 0.44–0.90) in women who quit smoking ≥30 years ago compared to those who quit 0 to <5 years ago. There was no association of any category of time since smoking cessation and the risk of seronegative RA.

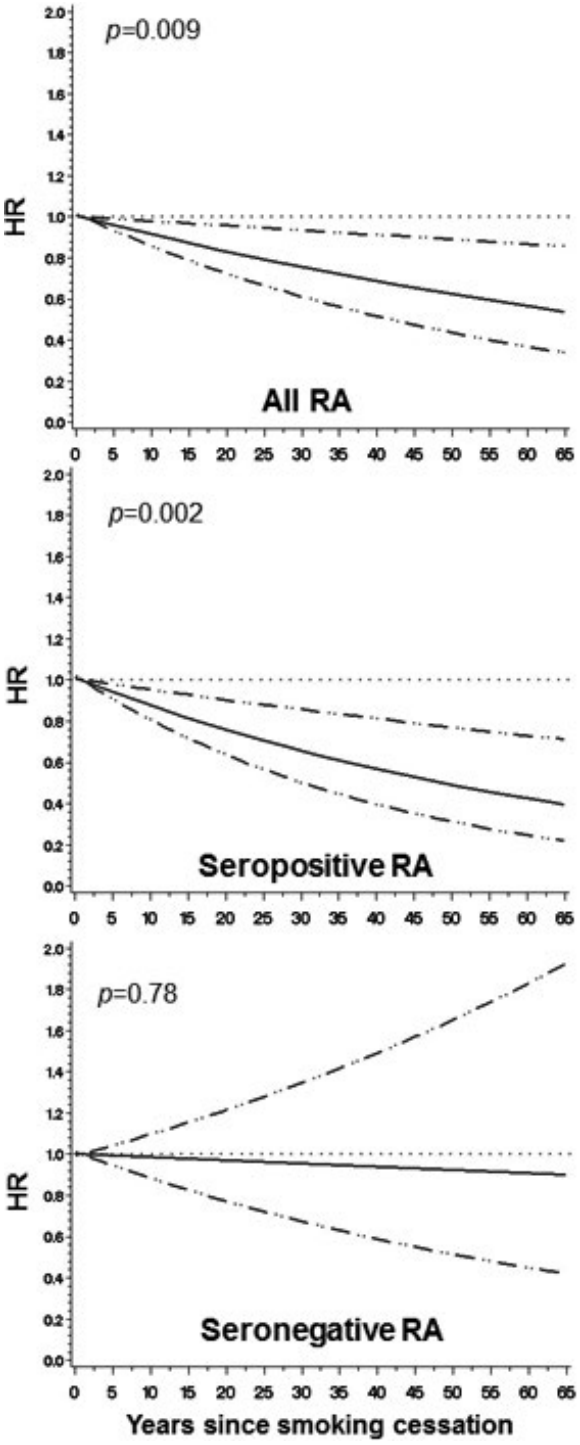


Figure 2. Restricted cubic spline curves showing HRs (solid lines) and 95% CIs (broken lines) for RA serologic phenotypes among the subset of past smokers, according to years since smoking (reference group, 0–2 years since quitting). Curves were adjusted for the covariates listed in Table 2. Broken lines represent the 95% CIs. *P* values are for trend. See Figure 1 for abbreviations.

DISCUSSION

In this large prospective study of women in the NHS and NHS II, we observed that sustained smoking cessation reduced the risk of seropositive RA compared to the risk in recent quitters, suggesting that this behavior change can delay or even prevent the onset of seropositive RA. The risk of seropositive RA was reduced by 37% among those who sustained smoking cessation for ≥30 years compared to those who recently quit smoking. Further, we showed an increased risk of RA, particularly the seropositive phenotype, in past-smokers and current-smokers, particularly in current-smokers with a high intensity of smoking or those with many pack-years of smoking and a strong dose-response, confirming the results of previous studies (4,7,25–27).

Table 4. Hazard ratios for rheumatoid arthritis (RA) serologic phenotypes according to smoking status and years since cessation among past smokers*

	Quit ≥30 years ago	Quit 20 to <30 years ago	Quit 10 to <20 years ago	Quit 5 to <10 years ago	Quit 0 to <5 years ago
All RA					
Cases/person-years	123/397,373	129/445,185	160/551,108	90/262,059	94/282,440
Age-adjusted model	0.75 (0.56–1.00)	0.75 (0.57–0.99)	0.86 (0.67–1.12)	1.09 (0.81–1.46)	1.00 (Ref.)
Multivariable model	0.78 (0.58–1.05)	0.78 (0.59–1.02)	0.88 (0.68–1.14)	1.09 (0.81–1.46)	1.00 (Ref.)
Seropositive RA					
Cases/person-years	77/396,579	82/443,916	99/549,283	58/261,118	69/281,311
Age-adjusted model	0.62 (0.44–0.88)	0.62 (0.45–0.86)	0.72 (0.53–0.98)	0.93 (0.65–1.32)	1.00 (Ref.)
Multivariable model	0.63 (0.44–0.90)	0.63 (0.45–0.88)	0.73 (0.53–0.99)	0.91 (0.64–1.30)	1.00 (Ref.)
Seronegative RA					
Cases/person-years	46/396,731	47/443,877	61/549,236	32/261,042	25/281,204
Age-adjusted model	1.11 (0.66–1.86)	1.12 (0.68–1.83)	1.27 (0.79–2.04)	1.54 (0.91–2.62)	1.00 (Ref.)
Multivariable model	1.20 (0.71–2.02)	1.17 (0.71–1.93)	1.30 (0.81–2.09)	1.57 (0.92–2.66)	1.00 (Ref.)

* Multivariable models were adjusted for age, questionnaire period, cohort, oral contraceptive use (ever, never), parity/breastfeeding in months (nulliparous, parous/<1 month, parous/1–11 months, parous/≥12 months), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/never, postmenopausal/ever), body mass index category (underweight/normal, overweight, obese), sedentary physical activity, median household income (quartiles), alcohol intake (none to <5 gm/day, 5 to <10 gm/day, ≥10 gm/day). Except where indicated otherwise, values are the hazard ratio (95% confidence interval).

We did not find an association between smoking and seronegative RA risk despite a large sample size and lengthy follow-up, suggesting that seropositive and seronegative RA may be distinct phenotypes with distinct risk factors.

Many previous studies have investigated smoking status and RA risk. In a large meta-analysis that included 11 studies (7), 13,885 RA cases among a total of 593,576 individuals, current-smokers had an odds ratio (OR) of 1.64 for seropositive RA compared to never-smokers. Recent results from the French E3N cohort study in which 71,248 women were prospectively followed since 1990 showed that past-smokers (HR 1.32, 95% CI 1.06-1.64) and current-smokers (HR 1.57, 95% CI 1.13–2.19) had an increased RA risk compared to never-smokers (1). Our group previously investigated smoking performed only in the NHS using follow-up data from 1976 to 2002 and demonstrated that current- smokers had an HR of 1.46 (95% CI 1.20–1.79) for all RA and HR of 1.58 (95% CI 1.21–2.06) for seropositive RA compared to never-smokers (5). Our current study findings are consistent with and extend these prior findings. We confirmed a strong association between smoking status and seropositive RA risk but no clear association with seronegative RA.

Stolt et al investigated smoking intensity and RA risk and reported an OR of 2.4 (95% CI 1.5–3.7) for RA among current-smokers who smoked ≥20 cigarettes per day compared to never-smokers (16). A recent meta-analysis of 3 cohort and 7 case–control studies demonstrated a dose-response relationship between smoking pack-years and risk of RA, showing a statistically significant increased risk of developing RA with increasing pack-years up to 20 years, when the HR plateaued at ~2.0, compared to never-smokers (2). We observed a similar relationship between smoking pack-years and risk of all RA and extended those findings by also investigating seropositive RA, with the HR plateauing at ~2.3 after approximately 30 pack-years compared to never smoking.

Smoking status, intensity, and pack-years are all associated with the risk of RA, particularly seropositive RA, which implies that a behavior change of smoking cessation might reduce the risk of RA. Some previous studies have also investigated smoking cessation and RA risk. The Swedish Mammography Cohort study (15) followed 34,101 women from 1997 to 2010, using a baseline questionnaire on smoking behaviors, and identified 219 incident cases of RA. Past-smokers who quit smoking ≥15 years ago had an increased risk of RA (HR 1.99, 95% CI 1.23–3.20) compared to never-smokers, suggesting that a residual elevated risk of RA remained even after sustained cessation of smoking. When we analyzed only past-smokers, there was a suggestion that the risk of RA was reduced with increasing time since cessation compared to recent quitters. However, the findings in that study were limited due to a low number of events and only a single baseline assessment of smoking. In the Swedish Epidemiological Investigations of RA analyzing 679 cases and 847 controls (16), past-smokers who quit smoking ≥20 years had a similar risk of RA (OR 1.0, 95% CI 0.5–1.9) compared to never-smokers. However, that case–control study may have been limited by recall bias, and there were few cases in the sustained smoking cessation group, so there may have been limited power to detect a true difference.

Similarly, the previous study analyzing only the NHS suggested that past-smokers who quit smoking ≥20 years ago had a similar risk of RA (HR 1.14, 95% CI 0.88–1.48) compared to never-smokers (5). In our current study, although the risk of RA

risk decreased with time since cessation, a modestly elevated RA risk was detectable 30 years after quitting smoking (for all RA, HR 1.25 [95% CI 1.02–1.53]; for seropositive RA, HR 1.30 [95% CI 1.01–1.68] [reference, never smoking]). By extending follow-up in the NHS and adding the NHS II cohort, our current study is better powered to detect a modest statistical difference among women with long-term sustained cessation. Therefore, our study extends previous findings and provides evidence that women who smoke may have a modestly increased risk of RA for decades. This suggests that secular trends in smoking cessation may be followed by a decrease in RA incidence in future decades. Although smoking cessation may not decrease the risk of RA to the level of a never-smoker, our findings provide evidence that a behavior change of smoking cessation may delay or even prevent the onset of seropositive RA. These results could provide rationale for a smoking intervention trial among active smokers to prevent the formation of RA-related autoantibodies or to prevent the progression to RA among those at increased risk of seropositive RA.

We observed that the risk of RA among recent quitters (0 to <5 years since smoking cessation) was higher than that among current-smokers, perhaps due to many of the recent quitters being heavy smokers (>40 pack-years). These recent quitters may be more likely to start smoking again so may not have actually had sustained smoking cessation and may be similar to current-smokers. Moreover, recent quitters may have decided to quit smoking due to early symptoms of RA or other serious health conditions. The population-attributable risk of RA from smoking is 14% (28) and may contribute up to 35% of the risk of anti–citrullinated protein antibody–positive RA (29). Smoking may interact with shared epitope genes to increase the risk of seropositive RA (4,30,31).

Although the biologic mechanisms linking smoking to an increased risk of developing RA are still not clear, components in cigarette smoke, such as nicotine, hydrocarbons, and carbon monoxide, are known to have aberrant effects on the immune system (9,32). Smoking causes impaired T cell function (33,34) and humoral immunity (35,36) and raises systemic levels of inflammatory markers such as interleukin-6 and C-reactive protein (37,38). Smoking has also been shown to increase levels of citrullinated proteins and expression of peptidylarginine deiminase type 2 in pulmonary alveoli (39). Evidence has accumulated showing that, in the presence of HLA shared epitope genes, cigarette smoking may trigger immune responses against citrullinated proteins (11,31,40). The observed associations between smoking status/intensity as well as the dose-response between pack-years of smoking with RA risk in our study are compatible with this triggering mechanism. Moreover, the detectable increased risk of RA 30 years after smoking cessation suggests that, in some individuals, the immune system may be permanently altered, perhaps with resultant autoimmunity established once a threshold of smoking is reached, with progression to RA occurring many years later.

A major strength of our study is the use of 2 large cohorts to prospectively identify incident cases of RA with up to 38 years during more than 6 million person-years of follow-up. We had detailed data on smoking exposures including smoking status, intensity, cumulative pack-years, and years since smoking cessation as well as information on important potential confounders such as alcohol intake and reproductive factors updated prospectively every 2 years, allowing for time-updated analyses. Further, women who self-reported CTD including unconfirmed RA were censored at the time of self-report to ensure that the analyzed sample was free of RA or other CTD. We identified cases by medical record review to ensure that all fulfilled accepted criteria and were truly incident, while allowing for subphenotyping based on serologic phenotype.

Our study does have some limitations. Our study population, consisting of mostly healthy, well-educated, white US women working in the nursing professions at baseline may not be representative of the general population. Because detailed smoking data were self-reported, there is the potential for recall bias. However, self-report of smoking has been demonstrated to be valid, and these repeated measures were collected prospectively prior to RA onset, so a differential bias between RA cases and non-RA cases is unlikely (41). Because smoking was assessed only every 2 years, we might not have captured intervening smoking behavior changes. In addition, there may be potential for misclassification by RA serologic phenotype. The serologic phenotype in our incident RA cases was determined by the combination of RF and anti-CCP tests obtained through routine clinical care. Many cases were diagnosed prior to the early 2000s when anti-CCP testing began to be used widely in the US. Thus, for earlier RA cases, medical records only contained data on RF. It is therefore possible that some of the women in whom RA was diagnosed before the early 2000s who were RF-negative may have actually been anti-CCP positive but misclassified as seronegative in our study. Because RF and anti-CCP are correlated, we expect that the misclassification of seropositivity in RA patients is relatively uncommon. We previously observed that only ~2% of our seronegative RA cases were initially misclassified based on negative RF but positive anti-CCP assays among a subset of women who provided blood samples in 1989 and had these tests for research purposes (42).

In conclusion, we observed that past-smokers had a significantly reduced risk of seropositive RA according to time since sustained smoking cessation, providing evidence that this behavior change may decrease or even prevent the onset of RA. We detected a slightly increased risk of seropositive RA even 30 years after smoking cessation in ever-smokers compared to never-smokers, suggesting that a minority of ever-smokers may have permanent immune alterations even after smoking cessation. We found no association of smoking with seronegative RA, suggesting a different pathogenesis for sero-

positive RA. Our study findings provide evidence that a behavior change of sustained smoking cessation may reduce the risk of seropositive RA.

ACKNOWLEDGMENTS

We thank the participants in the NHS and NHS II for their dedicated participation in this longitudinal study as well as the staff members at the Channing Division of Network Medicine (Department of Medicine, Brigham and Women's Hospital and Harvard Medical School).

AUTHOR CONTRIBUTIONS

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

Acquisition of data. Tedeschi, Barbhaiya, Leatherwood, Costenbader, Karlson, Sparks.

Analysis and interpretation of data. Liu, Tedeschi, Barbhaiya, Leatherwood, Speyer, Lu, Costenbader, Karlson, Sparks.

REFERENCES

1. Seror R, Henry J, Gusto G, Aubin HJ, Boutron-Ruault MC, Mariette X. Passive smoking in childhood increases the risk of developing rheumatoid arthritis. Rheumatology (Oxford) 2018. E-pub ahead of print.
2. Di Giuseppe D, Discacciati A, Orsini N, Wolk A. Cigarette smoking and risk of rheumatoid arthritis: a dose-response meta-analysis. Arthritis Res Ther 2014;16:R61.
3. De Rooy DP, van Nies JA, Kapetanovic MC, Kristjansdottir H, Andersson ML, Forslind K, et al. Smoking as a risk factor for the radiological severity of rheumatoid arthritis: a study on six cohorts. Ann Rheum Dis 2014;73:1384–7.
4. Karlson EW, Deane K. Environmental and gene-environment interactions and risk of rheumatoid arthritis. Rheum Dis Clin North Am 2012;38:405–26.
5. Costenbader KH, Feskanich D, Mandl LA, Karlson EW. Smoking intensity, duration, and cessation, and the risk of rheumatoid arthritis in women. Am J Med 2006;119: 503 e1–9.
6. Hutchinson D, Shepstone L, Moots R, Lear JT, Lynch MP. Heavy cigarette smoking is strongly associated with rheumatoid arthritis (RA), particularly in patients without a family history of RA. Ann Rheum Dis 2001;60:223–7.
7. 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.
8. Chang K, Yang SM, Kim SH, Han KH, Park SJ, Shin JI. Smoking and rheumatoid arthritis. Int J Mol Sci 2014;15:22279–95.
9. Sopori M. Effects of cigarette smoke on the immune system. Nat Rev Immunol 2002;2:372–7.
10. Onozaki K. Etiological and biological aspects of cigarette smoking in rheumatoid arthritis. Inflamm Allergy Drug Targets 2009;8:364–8.
11. Klareskog L, Padyukov L, Alfredsson L. Smoking as a trigger for inflammatory rheumatic diseases. Curr Opin Rheumatol 2007;19:49–54.
12. Hedstrom AK, Klareskog L, Alfredsson L. Exposure to passive smoking and rheumatoid arthritis risk: results from the Swedish EIRA study. Ann Rheum Dis 2018;77:970–2.
13. Criswell LA, Merlino LA, Cerhan JR, Mikuls TR, Mudano AS, Burma M, et al. Cigarette smoking and the risk of rheumatoid arthritis among postmenopausal women: results from the Iowa Women's Health Study. Am J Med 2002;112:465–71.
14. Karlson EW, Lee IM, Cook NR, Manson JE, Buring JE, Hennekens CH. A retrospective cohort study of cigarette smoking and risk of rheumatoid arthritis in female health professionals. Arthritis Rheum 1999;42:910–7.
15. Di Giuseppe D, Orsini N, Alfredsson L, Askling J, Wolk A. Cigarette smoking and smoking cessation in relation to risk of rheumatoid arthritis in women. Arthritis Res Ther 2013;15:R56.
16. Stolt P, Bengtsson C, Nordmark B, Lindblad S, Lundberg I, Klareskog L, et al. Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. Ann Rheum Dis 2003;62: 835–41.
17. Hazes JM, Dijkmans BA, Vandenbroucke JP, de Vries RR, Cats A. Lifestyle and the risk of rheumatoid arthritis: cigarette smoking and alcohol consumption. Ann Rheum Dis 1990;49:980–2.
18. Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. J Womens Health 1997;6:49–62.
19. Karlson EW, Sanchez-Guerrero J, Wright EA, Lew RA, Daltroy LH, Katz JN, et al. A connective tissue disease screening questionnaire for population studies. Ann Epidemiol 1995;5:297–302.
20. 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.
21. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62: 2569–81.
22. Ainsworth BE, Haskell WL, Leon AS, Jacobs DR Jr, Montoye HJ, Sallis JF, et al. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sports Exerc 1993;25:71–80.
23. Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. Am J Clin Nutr 1999;69:243–9.
24. Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med 1989;8:551–61.
25. Klareskog L, Padyukov L, Ronnelid J, Alfredsson L. Genes, environment and immunity in the development of rheumatoid arthritis. Curr Opin Immunol 2006;18:650–5.
26. Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L. A gene–environment interaction between smoking and shared epitope genes in HLA–DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 2004;50:3085–92.
27. Krishnan E, Sokka T, Hannonen P. Smoking-gender interaction and risk for rheumatoid arthritis. Arthritis Res Ther 2003;5:R158–62.
28. Sparks JA, Chen CY, Hiraki LT, Malspeis S, Costenbader KH, Karlson EW. Contributions of familial rheumatoid arthritis or lupus and environmental factors to risk of rheumatoid arthritis in women: a prospective cohort study. Arthritis Care Res (Hoboken) 2014;66:1438–46.
29. Kallberg H, Ding B, Padyukov L, Bengtsson C, Ronnelid J, Klareskog L, et al. Smoking is a major preventable risk factor for rheumatoid

arthritis: estimations of risks after various exposures to cigarette smoke. Ann Rheum Dis 2011;70:508–11.

30. Karlson EW, Chang SC, Cui J, Chibnik LB, Fraser PA, De Vivo I, et al. Gene-environment interaction between HLA-DRB1 shared epitope and heavy cigarette smoking in predicting incident rheumatoid arthritis. Ann Rheum Dis 2010;69:54–60.
31. Kim K, Jiang X, Cui J, Lu B, Costenbader KH, Sparks JA, et al. Interactions between amino acid–defined major histocompatibility complex class II variants and smoking in seropositive rheumatoid arthritis. Arthritis Rheumatol 2015;67:2611–23.
32. George J, Levy Y, Shoenfeld Y. Smoking and immunity: an additional player in the mosaic of autoimmunity. Scand J Immunol 1997;45: 1–6.
33. Robbins CS, Dawe DE, Goncharova SI, Pouladi MA, Drannik AG, Swirski FK, et al. Cigarette smoke decreases pulmonary dendritic cells and impacts antiviral immune responsiveness. Am J Respir Cell Mol Biol 2004;30:202–11.
34. Hughes DA, Haslam PL, Townsend PJ, Turner-Warwick M. Numerical and functional alterations in circulatory lymphocytes in cigarette smokers. Clin Exp Immunol 1985;61:459–66.
35. Moszczynski P, Zabinski Z, Moszczynski P Jr, Rutowski J, Slowinski S, Tabarowski Z. Immunological findings in cigarette smokers. Toxicol Lett 2001;118:121–7.
36. Burton RC. Smoking, immunity, and cancer. Med J Aust 1983;2: 411–2.
37. Bermudez EA, Rifai N, Buring JE, Manson JE, Ridker PM. Relation between markers of systemic vascular inflammation and smoking in women. Am J Cardiol 2002;89:1117–9.
38. Tracy RP, Psaty BM, Macy E, Bovill EG, Cushman M, Cornell ES, et al. Lifetime smoking exposure affects the association of C-reactive protein with cardiovascular disease risk factors and subclinical disease in healthy elderly subjects. Arterioscler Thromb Vasc Biol 1997;17:2167–76.
39. Makrygiannakis D, Hermansson M, Ulfgren AK, Nicholas AP, Zendman AJ, Eklund A, et al. Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells. Ann Rheum Dis 2008;67:1488–92.
40. Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA–DR (shared epitope)–restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006;54:38–46.
41. Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: a review and meta-analysis. Am J Public Health 1994;84:1086–93.
42. Bengtsson C, Malspeis S, Orellana C, Sparks JA, Costenbader KH, Karlson EW. Association between menopausal factors and the risk of seronegative and seropositive rheumatoid arthritis: results from the Nurses' Health Studies. Arthritis Care Res (Hoboken) 2017;69: 1676–84.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 925–935
DOI 10.1002/acr.23723

American College of Rheumatology
Empowering Rheumatology Professionals

Effects of Language, Insurance, and Race/Ethnicity on Measurement Properties of the PROMIS Physical Function Short Form 10a in Rheumatoid Arthritis

Zara Izadi,[1] Patricia P. Katz,[1] Gabriela Schmajuk,[2] Julie Gandrup,[1] Jing Li,[1] Milena Gianfrancesco,[1] and Jinoos Yazdany[1]

Objective. Most studies that have evaluated patient-reported outcomes, such as those utilizing the Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function Short Form 10a (PF10a) in rheumatoid arthritis (RA), have been performed in white and English-speaking populations. The aim of our study was to assess the measurement properties of the PF10a in a racially/ethnically diverse population with RA and to determine the effect of non-English language proficiency, insurance status, and race/ethnicity on the validity and responsiveness of the PF10a.

Methods. Data were abstracted from electronic health records for all RA patients seen in a university-based rheumatology clinic between 2013 and 2017. We evaluated the use of the PF10a, floor and ceiling effects, and construct validity across categories of language preference, insurance, and race/ethnicity. We used standardized response means and linear mixed-effects models to evaluate the responsiveness of the PF10a to longitudinal changes in the Clinical Disease Activity Index (CDAI) across population subgroups.

Results. We included 595 patients in a cross-sectional analysis of validity and 341 patients in longitudinal responsiveness analyses of the PF10a. The PF10a had acceptable floor and ceiling effects and was successfully implemented. We observed good construct validity and responsiveness to changes in CDAI among white subjects, English speakers, and privately insured patients. However, constructs evaluated by the PF10a were less correlated with clinical measures among Chinese speakers and Hispanic subjects, and less sensitive to clinical improvements among Medicaid patients and Spanish speakers.

Conclusion. While the PF10a has good measurement properties and is both practical and acceptable for implementation in routine clinical practice, we also found important differences across racial/ethnic groups and those with limited English proficiency that warrant further investigation.

INTRODUCTION

Assessment of patient-reported physical function is important for monitoring individuals with rheumatoid arthritis (RA) and provision of patient-centered care. Integration of patient-reported physical function into routine clinical care has been shown to be a feasible mechanism for incorporating patient preferences into a treat-to-target approach for managing RA (1). Incorporation of patient-reported measures of physical function in RA is a nationally endorsed quality measure and recommended in American College of Rheumatology guidelines (2).

The Patient Reported Outcome Measurement Information System (PROMIS) was developed by the National Institute of Health to provide standard metrics for measuring patient-reported outcomes across chronic conditions. To date, a number of researchers have examined the properties of PROMIS measures in rheumatic conditions, with the largest concentration of work being on the physical functioning measures in RA (3–11). While PROMIS physical function measures have been evaluated in white and English-speaking populations with RA, no studies have examined their validity or responsiveness in other racial and ethnic groups, non-English speakers, or populations with low socioeconomic status.

[1]Zara Izadi, PharmD MAS, Patricia P. Katz, PhD, Julie Gandrup, MD, Jing Li, MPH, Milena Gianfrancesco, PhD, MPH, Jinoos Yazdany, MD, MPH: University of California, San Francisco; [2]Gabriela Schmajuk, MD, MSc: San Francisco Veterans Affairs Medical Center, San Francisco, California.

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

Address correspondence to Zara Izadi, PharmD MAS, Department of Epidemiology and Biostatistics, Mission Hall: Global Health & Clinical Sciences Building, 550 16th Street, 2nd Floor, Box #0560, San Francisco, CA 94158-2549. E-mail: zara.izadi@ucsf.edu.

Submitted for publication May 21, 2018; accepted in revised form August 7, 2018.

SIGNIFICANCE & INNOVATIONS

- There is growing interest nationally in using patient-reported outcomes in routine clinical care to engage patients, monitor outcomes and inform treatment decisions; however, most studies evaluating patient-reported outcomes have been performed in white, English-speaking populations.
- We studied the validity and responsiveness of a Patient Reported Outcome Measurement Information System (PROMIS) measure in rheumatoid arthritis across different languages, races/ethnicities, and insurance groups in a real-world clinic population.
- We found that constructs evaluated by PROMIS Physical Function Short Form 10a (PF10a) are less correlated with clinical outcomes among Chinese speakers and Hispanic patients and that the PF10a has less sensitivity to clinical improvements among Medicaid patients and Spanish speakers.

Previous studies have shown that sociodemographic factors can affect multiple aspects of care in RA, including mortality and disability, disease activity, prevalence of comorbidities, patient-reported outcomes, access to treatment/health services, treatment preferences and medication adherence, health literacy, and trust in providers (12–35). A better understanding of the effects of sociodemographic factors on the validity of PROMIS physical function measures will determine their generalizability across diverse communities. This study demonstrates the effect of language preference, race/ethnicity, and insurance status (as a proxy for low income) on measurement properties of the PROMIS Physical Function Short Form 10a (PF10a), including floor and ceiling effects, construct validity, and responsiveness to improvements and deteriorations in clinical disease activity over time.

PATIENTS AND METHODS

Data sources. Provider information and clinical and demographic data were extracted from the electronic health record (EHR) for all patients seen at the University of California, San Francisco (UCSF) rheumatology clinic, with at least 1 face-to-face encounter with a rheumatologist that was associated with an International Classification of Diseases, Ninth Edition code for RA between February 1, 2013 and October 31, 2017. The UCSF Committee on Human Research approved this study.

Study population. In order to assess cross-sectional validity, we included patients who had at least 1 encounter with complete data, including a pain score, PF10a score, Clinical Disease Activity Index (CDAI) score, serum C-reactive protein (CRP) level, and erythrocyte sedimentation rate (ESR). For the longitudinal analysis of responsiveness, we restricted analysis to the cohort of patients with at least 2 encounters with complete data. For inclusion in the analysis, encounters were required to be spaced between 1 and 12 months apart, in order to capture separate episodes of care. The baseline was defined as the first encounter with complete data.

Measures. Physical function was measured using the PF10a for all patients. The PF10a is a 10-item questionnaire that assesses current self-reported physical function. Raw scores range from 10 to 50 and can be translated into T scores, with a mean ± SD of 50 ± 10, for comparison with the US general population mean; for this study, all reported PF10a scores are T scores. A higher PROMIS-PF10a T score represents better physical function. Chinese and Spanish PF10a forms were obtained from www.nih.promis.gov and were utilized for patients who preferred these languages. All forms were scored and entered by clinic staff prior to the encounter.

RA disease activity was measured using the CDAI (19), a composite measure of patient global assessment (on a 0–100 mm visual analog scale [VAS]), evaluator global assessment (on a 0–100 mm VAS), 28-tender joint counts (TJC), and 28-swollen joint counts (SJC). Scores range from 0 to 76, with higher values reflecting more severe disease. All patients completed a VAS (0–100 mm) for pain at each encounter, and CRP level or ESR was measured at least every 3 months.

Other variables. Other time-varying variables included body mass index (BMI) and smoking status, which were recorded at each encounter. Baseline variables included demographics, number of comorbidities (Charlson comorbidity score), and medication use. Demographics included date of birth, sex, self-reported race/ethnicity, preferred language, and insurance category (private, Medicare, Medicaid). For medication use at baseline, physician medication orders for all oral or intravenous drugs, including biologic disease-modifying antirheumatic drugs (DMARDs), targeted small-molecule DMARDs, nonbiologic DMARDs, and glucocorticoids that were associated with a rheumatology encounter within 12 months before baseline were retrieved from the EHR.

Statistical analysis. Pearson's chi-square test, one-way analysis of variance, or Kruskal-Wallis test were selected for descriptive statistics, based on the type and sample distribution of the variable being analyzed. The proportion of individuals with floor (defined as worst score, 14.1) and ceiling (defined as best score, 61.7) effects for PF10a was calculated across different categories of language, insurance, and race/ethnicity.

Construct validity, the extent to which a test measures the concept or construct that it is intended to measure, was assessed by looking at convergent, discriminant, and known-group validity. Convergent and discriminant validity explain how a measure

conforms to a similar or different measure and were assessed by comparing correlation of the PF10a to that of patient global RA assessments of pain, SJC and TJC, ESR, and CRP level with Spearman's correlation coefficient, as not all scores were normally distributed. We hypothesized that the PF10a would correlate strongly ($r < -0.60$) with other patient-reported measures (patient global assessment VAS, pain VAS), and moderately ($-0.30 < r < -0.60$) with clinical outcome measures (28 TJCs and SJCs) (36). Spearman's correlation coefficient was compared across different categories of language, insurance, and race/ethnicity.

Known-group validity explains how the measure discriminates between groups that are known to be different. This was investigated by evaluating differences in mean PF10a scores among predefined groups by age or disease severity. The PF10a was hypothesized to show lower scores in older patients (ages ≥65 compared to age <50 years), and those with higher disease activity (CDAI score >22 [severe] compared to CDAI score ≤10 [remission or mild]). We used *t*-tests to compare mean group differences in each category by language, insurance, and race/ethnicity, and Cohen's *d* effect size (the difference in mean scores divided by the pooled SD) was calculated. Effect size values for dichotomous variables were categorized as small (<0.5), medium (0.5–0.8), or large (>0.8) (37).

Responsiveness. Responsiveness was determined by analyzing changes in PF10a scores in relation to changes in disease activity (CDAI). Previous studies have defined a minimally important difference in CDAI as a 12-point change (38). In order to estimate the standardized response mean (SRM), patients with PF10a scores recorded on 2 encounters that were 1–12 months apart were divided into 3 groups, including those with a 12-point decrease in CDAI (clinical improvement), including those with a 12-point increase in CDAI (clinical deterioration), and those with a <12-point change in CDAI (no change). We investigated the association between language preference, insurance status, and race/ethnicity, and mean score changes of the PF10a using a test for trend (39). We then calculated the ratio of the mean score change to the SD of that change (SRM) across subgroups. Values were categorized as small (<0.5), medium (0.5–0.8), and large (>0.8) (40) and were compared across different categories of language, insurance, and race/ethnicity.

Finally, we used multilevel linear mixed-effects regression to assess the responsiveness of the PF10a to changes over the follow-up period by modeling the relationship between changes in the PF10a and changes in CDAI among all patients with at least 2 encounters. We used a random effects model (Model 1), allowing each subject to have his/her own starting intercept and disease trajectory. Also, since there may have been systematic differences in how providers rate swollen and tender joints in the CDAI, we accounted for clustering by provider. Because most patients saw the same provider across all visits, a nested random effects model was used. The association between change in CDAI score and change in PF10a might be influenced by the magnitude of the initial PF10a score; we therefore adjusted for the initial PF10a score. Since different patients had follow-up visits at different times, we also incorporated time as a linear predictor in the model. In order to assess differences in PF10a responsiveness across population groups, we fitted 3 additional models, with an interaction term between change in CDAI score (since the previous encounter) and either language (Model 2), insurance (Model 3), or race/ethnicity (Model 4) in each model. To assess responsiveness to both improvements and deteriorations in CDAI, we fitted splines with a single knot at ΔCDAI of 0. In our fully-adjusted analyses, in addition to the above terms, we also included baseline covariates (age, sex, smoking status, Charlson comorbidity score [41], and medications), and time-dependent covariates (CRP level and BMI). Analyses were performed using Stata statistical software, release 14.

RESULTS

Data from 846 RA patients and 5,834 encounters (mean ± SD encounters 10 ± 5 per patient; range 1–31) were extracted from the EHR. PF10a scores were recorded for 833 patients (98%) at 5,174 encounters (89%). The final data set for cross-sectional analysis included 595 patients. Of these, 341 patients

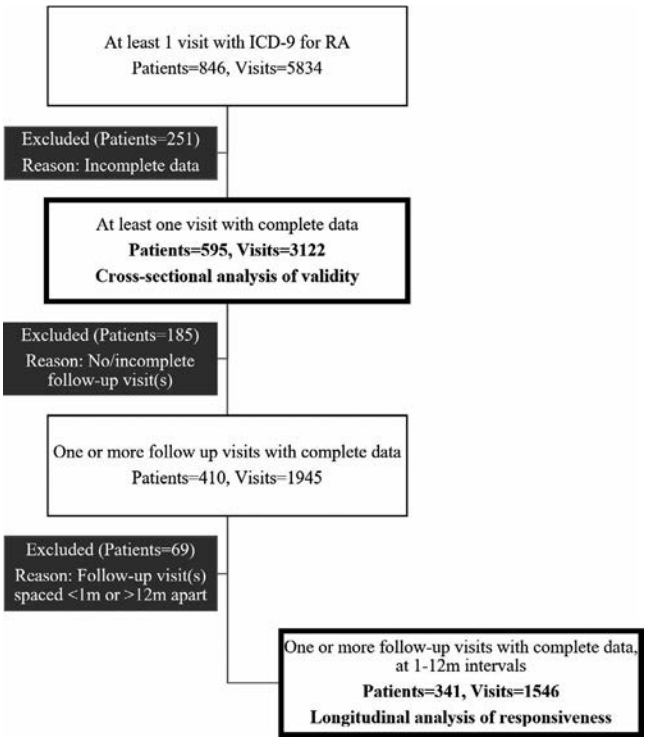


Figure 1. Data set for cross-sectional analysis of validity and longitudinal analysis of responsiveness. ICD-9 = International Classification of Diseases, Ninth Revision; RA = rheumatoid arthritis.

had complete data on at least 2 encounters that were 1–12 months apart (mean ± SD encounters 6 ± 3 per patient [range 2–15]) and were included in the longitudinal analysis of responsiveness (Figure 1). A total of 32 providers contributed data for analysis, with a mean ± SD RA encounters per provider of 23 ±11 (range 2–45).

Baseline clinical characteristics of patients included in the cross-sectional sample were representative of RA populations previously described (9,42) and similar to patients included in the longitudinal cohort (Table 1). The majority of patients were female (83%), with a mean ± SD age of 56 ± 15 years. The group was racially/ethnically diverse (50% nonwhite) and 14% preferred a

Table 1. Characteristics of the RA clinic population at baseline*

	Cross-sectional sample (n = 595)	Longitudinal cohort (n = 341)
Age, mean ± SD years	56.5 ± 15.3	55.8 ± 15.4
Female	493 (83)	282 (83)
BMI, mean ± SD	26.7 ± 6.4	26.2 ± 5.9
Race/ethnicity		
White	297 (50)	163 (48)
African American	38 (6)	19 (6)
Hispanic	27 (5)	14 (4)
Asian	115 (19)	73 (21)
Other	118 (20)	72 (21)
Preferred language		
English	512 (86)	288 (85)
Spanish	43 (8)	28 (8)
Chinese	40 (6)	25 (7)
Insurance type		
Private	236 (40)	142 (42)
Medicare	281 (47)	155 (45)
Medicaid	78 (13)	44 (13)
Smoking	137 (23)	79 (23)
Total Charlson score, median (IQR)	1 (1–2)	1 (1–2)
Medication		
DMARD and biologic naive	26 (5)	11 (3)
DMARD only	311 (52)	181 (53)
Biologic with or without DMARD	258 (43)	149 (44)
Clinical parameters		
RA disease activity		
CDAI remission	85 (14)	49 (14)
CDAI low	207 (35)	111 (33)
CDAI moderate	168 (28)	104 (30)
CDAI high	135 (23)	77 (23)
PhGA VAS, median (IQR)	23 (10–44)	24 (10–44)
PtGA VAS, median (IQR)	40 (15–64)	35 (15–62)
PF10a, mean ± SD	40.1 ± 10.7	40.8 ± 10.4
Pain VAS, median (IQR)	40 (15–68)	33 (12–65)
28-joint TJC, median (IQR)	2 (0–6)	2 (0–6)
28-joint SJC, median (IQR)	2 (0–5)	2 (0–6)
CRP mg/dl, median (IQR)	4 (2–9.2)	4 (2–8.6)
ESR mm/hr, median (IQR)	19 (9–36)	20 (10–35)

* Values are the number (%) unless indicated otherwise. RA = rheumatoid arthritis; BMI = body mass index; IQR = interquartile range; DMARD = disease-modifying antirheumatic drug; CDAI = clinical disease activity index; PhGA = physician global assessment; VAS = visual analog score; PtGA = patient global assessment; PF10a = PROMIS Physical Function Short Form 10a; TJC = tender joint count; SJC = swollen joint count; CRP = C-reactive protein level; ESR = erythrocyte sedimentation rate.

language other than English. Most patients had Medicare (47%) or private insurance (40%). The mean ± SD PF10a score was 40 ± 11, nearly 1 SD lower than the overall US population mean; about half had moderate or severe disease activity scores at baseline and the majority had received at least 1 nonbiologic DMARD (52%) or a biologic (43%) at baseline.

Preferred language, insurance, and racial/ethnic groups differed by age, disease activity, and number of comorbidities at baseline. Chinese speakers were on average older than Spanish or English speakers (68, 59, and 55 years, respectively; $P = 0.022$). Significantly more patients with Medicaid coverage had moderate-severe disease activity (CDAI ≥10), than those with Medicare or privately insured (73%, 53%, 41%, respectively; $P < 0.001$). Patients insured with Medicaid and African American subjects had a statistically significantly higher median Charlson comorbidity score at baseline (2), than other insurance groups (all 1; $P = 0.015$) and other race/ethnicities (all 1; $P < 0.001$). Baseline PF10a scores were lower among non-English speakers, Medicaid patients, and African American subjects (Table 1 and Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.23723/abstract).

Floor and ceiling effects. Clustering at the floor was low, ranging from 0%–7.5% across all language, insurance, and racial/ethnic groups (Table 2 and Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23723/abstract). Similarly, ceiling effects ranged from 0% among Chinese speakers to 11% among privately insured patients.

Validity. In examining convergent and discriminant validity, PF10a scores were strongly correlated (r ≤ –0.60) with patient global assessment of RA activity in all groups except Chinese speakers (r = –0.52) and Medicaid patients (r = –0.53) (Table 2). PF10a scores also had strong correlations with pain in all groups except Chinese speakers (r = –0.52) and African American subjects (r = –0.53). Correlations between PF10a and clinical outcomes (SJC and TJC) were moderate (–0.3 ≥ r > –0.6) in most groups; Chinese speakers and Hispanic subjects had weak correlations (r > –0.3) with both outcomes, while Spanish speakers and African American subjects had moderate correlations with TJC but weak correlations with SJC. Correlations between PF10a and inflammatory markers were weak or negligible among groups except English speakers, white patients, and privately insured patients who had moderate correlations with CRP levels.

In examining known-group validity, patients who had more active disease (CDAI >22), had significantly lower mean PF10a scores, as hypothesized (Table 3). In the group dichotomized by disease activity (CDAI >22 versus CDAI ≤10), effect size (Cohen's *d*) was large (>0.8) and statistically significant with respect to all sociodemographic variables except for Hispanic race, which

Table 2. Construct validity analysis, showing Spearman's correlation coefficients between PROMIS PF10a scores and patient-reported outcomes, physician-assessed outcomes, and inflammatory markers among different subgroups*

	Patient-reported outcomes		Physician-assessed outcomes		Inflammatory markers	
	PGA (VAS)	Pain (VAS)	TJC	SJC	CRP	ESR
Language						
English, n = 512	-0.71†	–0.68†	–0.45†	–0.34†	–0.31†	–0.29†
Spanish, n = 43	–0.60†	–0.61†	–0.35†	–0.24	–0.02	–0.12
Chinese, n = 40	–0.52†	–0.52†	–0.28	–0.19	–0.13	–0.08
Insurance						
Private, n = 236	–0.72†	–0.66†	–0.61†	–0.49†	–0.33†	–0.24†
Medicare, n = 281	–0.62†	–0.60†	–0.29†	–0.17†	–0.23†	–0.24†
Medicaid, n = 78	–0.53†	–0.61†	–0.42†	–0.31†	–0.14	–0.14
Race/ethnicity						
White, n = 297	–0.69†	–0.67†	–0.44†	–0.33†	–0.34†	–0.28†
African American, n = 38	–0.62†	–0.53†	–0.45†	–0.12	–0.04	–0.32
Hispanic, n = 38	–0.62†	–0.64†	–0.25	–0.05	0.05	0.08
Asian, n = 115	–0.64†	–0.65†	–0.40†	–0.33†	–0.26†	–0.13
Other, n = 118	–0.69†	–0.65†	–0.48†	–0.41†	–0.14	–0.31†

* Values are Spearman's correlation coefficients (r) with Patient-Reported Outcomes Measurement Information System Physical Function Short-Form 10a (PROMIS PF10a) scores. Data includes a cross-sectional sample (n = 595). r < –0.6 = strong correlation; –0.3 > r > –0.6 = moderate correlation; r > –0.3 = weak correlation. PGA = patient global assessment; VAS = visual analog scale; TJC = tender joint count; SJC = swollen joint count; CRP = C-reactive protein level; ESR = erythrocyte sedimentation rate.
† $P < 0.05$.

Table 3. Known-group validity for PF10a using Cohen's *d* effect size*

	Age <50 years	Age ≥65 years	Difference in mean PF10a†	Cohen's *d*‡	CDAI ≤10†	CDAI >22†	Difference in mean PF10a†	Cohen's *d*‡
Language								
English	176	155	2.5§	0.24§	263	113	13.3§	1.44§
Spanish	9	18	6.6	0.69	18	13	11.6§	1.21§
Chinese	3	26	4.2	0.44	11	9	10.9§	0.95§
Insurance								
Private	131	12	6.1§	0.64§	140	50	16.5§	2.15§
Medicare	22	184	−6.0§	−0.59§	131	55	9.9§	1.04§
Medicaid	35	3	−8.6	−1.00	21	30	9.6§	1.06§
Race/ethnicity								
White	94	94	2.3	0.23	161	54	13.3§	1.49§
African American	6	16	0.8	0.07	15	10	9.9§	0.96§
Hispanic	12	3	0.9	0.09	13	6	5.3	0.48
Asian	28	50	6.6§	0.64§	47	28	13.6§	1.37§
Other	48	36	3.8	0.36	56	37	13.7§	1.50§

* Values are the number of patients unless indicated otherwise. Data include a cross-sectional sample (n = 595). PF10a = Physical Function Short-Form 10a; CDAI = Clinical Disease Activity Index.
† Difference between mean PF10a among those ages <50 years compared to those ages ≥65 years or among those with CDAI ≤10 compared to those with CDAI >22.
‡ Difference in mean scores divided by the pooled SD; effect size (Cohen's *d*) values are categorized as small (<0.5), medium (0.5–0.8), or large (>0.8).
§ $P < 0.05$ using Student's *t* test; for Cohen's *d*, 95% confidence intervals do not cross 0.

had a small and nonsignificant effect size. Older patients also had lower mean PF10a scores compared to younger patients; however, differences were not clinically or statistically significant in most groups. As expected, younger Medicare patients had significantly worse physical functioning.

Responsiveness. Of the 341 patients with at least 2 encounters, the median (interquartile range [IQR]) interval between visits was 126 days (IQR 97–202). Patients with 2 encounters were divided into 3 subgroups based on whether they had a 12-point change in CDAI (clinical improvement, no change, and clinical deterioration). Mean PF10a scores decreased with clinical deteriorations, remained constant with no clinical change, and increased with clinical improvements over time (Table 3 and Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23723/abstract). Mean score changes differed significantly between groups ($P < 0.05$) in all groups of language, insurance, and race/ethnicity except Chinese speakers, and African American and Asian subjects. In the improvement group, the SRM was large (>0.8) in English and Chinese speakers, those with private insurance, and white and Hispanic subjects; small (<0.5) in Spanish speakers and African American subjects, and medium (0.5–0.8) in all other groups. In the deterioration group, SRM was large or medium in all language, insurance, and race/ethnicity groups with sufficient numbers for analysis.

Linear mixed-effects regression showed that both clinical improvements and deteriorations were associated with changes in PF10a scores over time ($P < 0.001$), suggesting that PF10a is responsive to changes in clinical disease activity. In a model without interaction terms (Model 1), a 12-point increase in CDAI was associated with a 2.93-point decrease (95% confidence interval [95% CI] 2.06, 3.80) in PF10a, and a 12-point decrease in CDAI was associated with a 2.70 point increase (95% CI 2.00, 3.41) in PF10a scores. We constructed 3 additional models (Model 2–4), incorporating an interaction term between CDAI and either language, insurance, or race/ethnicity. Model 2 showed that PF10a is more responsive to clinical deteriorations among Chinese speakers than English speakers (a 12-point increase in CDAI was associated with a 5.96-point decrease in PF10a among Chinese speakers and a 2.92-point decrease in PF10a among English speakers; all $P = 0.036$) (Table 4, Figure 2a). The PF10a was less responsive to clinical improvements among Spanish speakers than English speakers (a 12-point decrease in CDAI was associated with a 0.67-point increase in PF10a among Spanish speakers and a 3.08-point increase in PF10a among English speakers; $P = 0.029$). Among Chinese speakers, PF10a appeared to be more responsive to clinical deteriorations than clinical improvements, although this shift in responsiveness was not statistically significant ($P = 0.065$). Model 3 showed highest sensitivity to changes in disease activity among individuals with private insurance. The responsiveness of the PROMIS

Table 4. Effect of language preference, insurance status, and race/ethnicity on responsiveness of PF10a to changes in clinical disease activity*

	Clinical improvement (12-point decrease in CDAI)			Clinical deterioration (12-point increase in CDAI)		
	Change in PF10a score (β)†	95% CI	*P*‡	Change in PF10a score (β)†	95% CI	*P*‡
Model 2						
Language						
English (Ref.)	3.08	2.31, 3.86	Ref.	–2.92	–3.93, –1.92	Ref.
Spanish	0.67	–1.37, 2.70	0.029	–0.83	–3.10, 1.44	0.098
Chinese	1.84	–0.73, 4.42	0.363	–5.96	–8.61, –3.31	0.036
Model 3						
Insurance						
Private	3.63	2.59, 4.66	Ref.	–5.33	–6.74, –3.92	Ref.
Medicare	2.21	1.17, 3.25	0.055	–1.62	–2.91, –0.34	<0.001
Medicaid	0.70	–1.27, 2.67	0.005	–1.79	–3.80, 0.22	0.005
Model 4						
Race/ethnicity						
White (Ref.)	3.41	2.37, 4.44	Ref.	–2.96	–4.41, –1.51	Ref.
African American	2.47	–0.17, 5.12	0.520	–0.77	–4.36, 2.82	0.268
Hispanic/Latino	2.79	–0.24, 5.82	0.708	–2.81	–5.99, 0.37	0.934
Asian	1.50	–0.12, 3.14	0.052	–3.49	–5.15, –1.82	0.636
Other	2.36	0.95, 3.78	0.241	–2.92	–4.74, –1.09	0.974

* Data include a longitudinal cohort (n = 341; encounters = 1,546). Results are from the linear mixed-effects regression and adjusted for baseline Physical Function Short-Form 10a (PF10a) score and time. Models 2, 3, and 4 incorporate interaction terms between changes in the Clinical Disease Activity Index (CDAI) score (from previous visit) and preferred language, insurance status, or race/ethnicity, respectively. 95% CI = 95% confidence interval; Ref. = reference.
† Magnitude of change in PF10a score.
‡ Value for statistical significance of effect modification by non-English language, nonprivate insurance, or nonwhite race.

PF10a to clinical deteriorations was lower among patients with Medicare (–1.62; $P < 0.001$) and Medicaid (–1.79; $P = 0.005$) than privately insured patients (–5.33) (Figure 2B). Responsiveness to clinical improvements was also lower among Medicaid patients than privately insured patients (0.70 versus 3.63; both $P = 0.005$). Model 4 showed highest responsiveness to clinical deteriorations among Asian subjects (–3.49) and highest responsiveness to clinical improvements among white subjects (3.41). Differences in responsiveness to clinical improvements or deteriorations between white and nonwhite subjects did not reach statistical significance. In all models, consistent results were obtained after adjustment for age, sex, BMI, smoking, baseline medications, baseline total Charlson comorbidities index and CRP level (data not shown).

DISCUSSION

Consistent with prior research, PF10a has strong measurement properties and is responsive to longitudinal changes in disease activity among English speakers and white and privately insured patients. However, our study highlights important differences across racial/ethnic groups and in those with limited English proficiency.

Impressively, PF10a scores were recorded for 98% of eligible patients at 89% of encounters, even among those with non-English language proficiency. This finding demonstrates that PF10a can be collected efficiently and consistently over a prolonged period in a busy clinic that provides care to a diverse community. Fewer ceiling effects were noted in this clinical sample than in some research samples (22). Given that floor and ceiling effects were below the commonly accepted criteria of 15% (43) across all categories of language, insurance, and race/ethnicity, the PF10a seems both practical and acceptable for use in a general practice setting.

In our evaluations of convergent and discriminant validity, PF10a scores generally correlated strongly with other PROs, moderately with clinical measures, and weakly with laboratory measures. Although findings are consistent with prior research (8,9) and reflective of the instrument's convergent and discriminant validity, we observed some deviations among language, insurance, and race/ethnicity groups. Most notably, non-English speakers and Hispanic and African American subjects had weaker correlations between PF10a scores and clinical outcomes such as the TJC and SJC. While some of these correlations may have been limited by small samples, unraveling the contributions of other factors that may contribute to these findings is important. Some literature

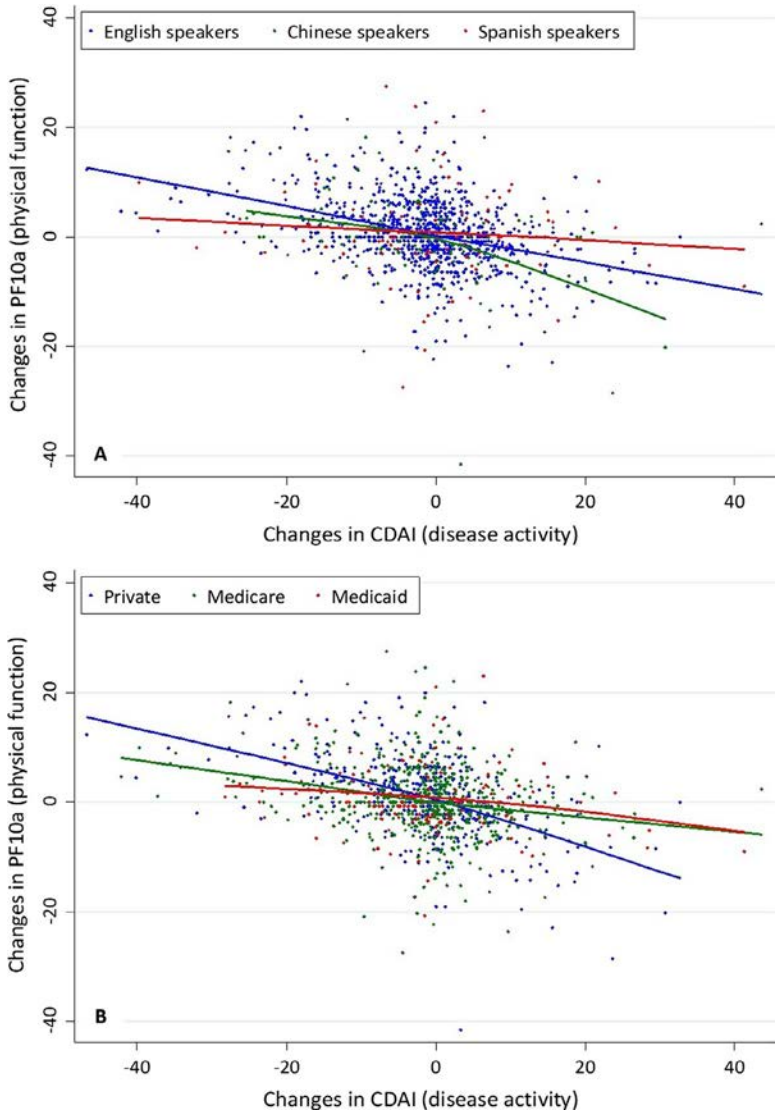


Figure 2. Linear mixed-effects regression models showing the effect of language and insurance status on the Physical Function Short Form 10a (PF10a) responsiveness to clinical improvements and deteriorations over time. CDAI = Clinical Disease Activity Index.

suggests that a higher prevalence of depression and chronic pain in some populations may hinder correlations between patient- and physician-reported outcomes in RA (22,27,44). Further, we found that Medicaid patients had weaker correlations between PF10a and patient global assessment of RA disease activity. Median baseline comorbidity scores were significantly higher among Medicaid patients than patients with private insurance. Since PF10a items are generic and do not specifically address RA-related impairments in physical functioning (45), weaker correlations between PF10a and patient global assessment may be attributed to non-RA comorbidities that were more prevalent among Medicaid patients.

Known-group differences by disease activity largely performed as hypothesized. A smaller effect size observed among Hispanics is likely the result of weak correlations between PF10a and clinical measures among this group, coupled with small sample sizes. Known-group differences by age did not perform as hypothesized and PF10a score differences among those <50 years compared to those ≥65 years were mostly small or not statistically significant. PF10a scores demonstrated in our study were better than expected among older individuals (data not shown), which may indicate a response-shift that is reflective of how patients adapt to and report their level of physical functioning over time. Response-shifts occur as patients recalibrate as they learn to live with RA. For instance, some patients with RA have reported that when they record a score of "0" on a questionnaire, this does not necessarily represent the absence of a symptom, but instead reflects a new baseline of "what is normal for me" (46).

SRMs, which were obtained in our evaluations of responsiveness, showed that the PF10a captured expected change and stability in scores across language, insurance, and race/ethnicity groups with sufficient numbers for analysis. While

another approach to evaluating responsiveness relies on patient self-reported change anchors obtained at a fixed time point, this was not possible in our retrospective analysis of clinical data. Mixed-effects modeling has been used previously to assess longitudinal responsiveness of a measure (47) and was used to model the relationship between changes in CDAI and changes in PF10a score. Among our entire eligible clinic population, we found that a 12-point increase in CDAI was associated with a 2.93-point decrease in PF10a, and a 12-point decrease in CDAI was associated with a 2.70-point increase in PF10a. Importantly, these findings are quantitatively consistent with prior evaluations of the responsiveness of PROMIS physical function measures anchored by deteriorations in clinical disease activity (9) or using patient-reported change anchors (7). However, we found that PF10a responsiveness to clinical improvements and deteriorations varied among population subgroups and was most notably influenced by insurance type and language preference. Patients with Medicaid coverage had worse baseline RA disease activity. Worse general health states among non-English speakers and low socioeconomic groups have been described previously (48). One possible explanation for poor responsiveness of the PF10a to clinical improvements among Medicaid patients and Spanish speakers may be average time spent in ill-states (11). It is possible that patients underreport their physical function during periods of clinical improvement because they reference their usual state rather than their improved state. Future research should examine responsiveness of the PF10a among non-English speakers and low socioeconomic groups to patient-reported change using validated change anchors. Responsiveness may also be dependent upon patients' physical functioning at baseline. Spanish speakers, Medicaid patients, and African American subjects had small PF10a score changes in response to CDAI worsening because their baseline PF10a scores were already poor and could not deteriorate much more (22).

Strengths of our study include representation of RA patients across the spectrum of RA disease activity, and inclusion of data from a large, real-world cohort. However, our study has some limitations. First, we were not able to examine individual item characteristics of the PF10a, which might inform internal consistency of items across language, insurance, and race/ethnicity groups. Second, the use of the CDAI as an anchor for changes in clinical disease activity may have introduced incorporation bias because of strong correlations between the PF10a score and patient global assessment, which is a component of the CDAI. Incorporation bias occurs when a reference standard is used that incorporates some of the test that is the subject of investigation. The result is a bias toward stronger associations between the PF10a and the CDAI among subgroups in whom the PF10a is strongly correlated with patient global assessment. However, correlations between the PF10a and patient global assessment varied only slightly among our population subgroups, making incorporation bias less likely in this study. Finally, sample size was modest among non-English speakers and nonwhite subjects, and we were underpowered to examine these subgroups in some analyses.

In order to optimize RA treatment, reliable, precise, and accurate measurement of symptoms and functional status across the continuum of disease activity has never been more important, given that remission or low disease activity is the current target for management (2,49). While ongoing efforts are in place to investigate the cross-cultural validity of PROMIS measures, our study is, to our knowledge, the first to evaluate the validity and responsiveness of a PROMIS measure across different languages, races/ethnicities, and insurance groups in a real-world clinic population and serves as an important step in the ongoing evaluation of the PROMIS Physical Function item bank. Our study demonstrated constructs evaluated by the PF10a were less correlated with clinical measures among Chinese speakers and Hispanic subjects, and less sensitive to clinical improvements among Medicaid patients and Spanish speakers.

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. Izadi 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. Izadi, Katz, Yazdany.

Acquisition of data. Izadi, Gandrup, Li.

Analysis and interpretation of data. Izadi, Katz, Schmajuk, Gianfrancesco, Yazdany.

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. 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.
3. Fries JF, Cella D, Rose M, Krishnan E, Bruce B. Progress in assessing physical function in arthritis: PROMIS short forms and computerized adaptive testing. J Rheumatol 2009;36:2061–6.
4. Fries JF, Krishnan E, Rose M, Lingala B, Bruce B. Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory. Arthritis Res Ther 2011;13:R147.
5. Bjorner JB, Rose M, Gandek B, Stone AA, Junghaenel DU, Ware JE Jr. Method of administration of PROMIS scales did not significantly impact score level, reliability, or validity. J Clin Epidemiol 2014;67:108–13.
6. Bartlett SJ, Orbai AM, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al. Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. PloS One 2015;10:e0138543.

7. Hays RD, Spritzer KL, Fries JF, Krishnan E. Responsiveness and minimally important difference for the patient-reported outcomes measurement information system (PROMIS) 20-item physical functioning short form in a prospective observational study of rheumatoid arthritis. Ann Rheum Dis 2015;74:104–7.
8. Oude Voshaar MA, Ten Klooster PM, Glas CA, Vonkeman HE, Taal E, Krishnan E, et al. Validity and measurement precision of the PROMIS physical function item bank and a content validity-driven 20-item short form in rheumatoid arthritis compared with traditional measures. Rheumatol (Oxford) 2015;54:2221–9.
9. Wahl E, Gross A, Chernitskiy V, Trupin L, Gensler L, Chaganti K, et al. Validity and responsiveness of a 10-item patient-reported measure of physical function in a rheumatoid arthritis clinic population. Arthritis Care Res (Hoboken) 2017;69:338–46.
10. Cook KF, Jensen SE, Schalet BD, Beaumont JL, Amtmann D, Czajkowski S, et al. PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. J Clin Epidemiol 2016;73:89–102.
11. Schalet BD, Hays RD, Jensen SE, Beaumont JL, Fries JF, Cella D. Validity of PROMIS physical function measured in diverse clinical samples. J Clin Epidemiol 2016;73:112–8.
12. Baldassari AR, Cleveland RJ, Luong MN, Jonas BL, Conn DL, Moreland LW, et al. Socioeconomic factors and self-reported health outcomes in African Americans with rheumatoid arthritis from the Southeastern United States: the contribution of childhood socioeconomic status. BMC Musculoskelet Disord 2016;17:10.
13. Barton JL, Schmajuk G, Trupin L, Graf J, Imboden J, Yelin EH, et al. Poor knowledge of methotrexate associated with older age and limited English-language proficiency in a diverse rheumatoid arthritis cohort. Arthritis Res Ther 2013;15:R157.
14. Barton JL, Trupin L, Tonner C, Imboden J, Katz P, Schillinger D, et al. English language proficiency, health literacy, and trust in physician are associated with shared decision making in rheumatoid arthritis. J Rheumatol 2014;41:1290–7.
15. Chen CH, Huang KY, Wang JY, Huang HB, Chou P, Lee CC. Combined effect of individual and neighbourhood socioeconomic status on mortality of rheumatoid arthritis patients under universal health care coverage system. Fam Pract 2015;32:41–8.
16. Constantinescu F, Goucher S, Weinstein A, Fraenkel L. Racial disparities in treatment preferences for rheumatoid arthritis. Med Care 2009;47:350–5.
17. Constantinescu F, Goucher S, Weinstein A, Smith W, Fraenkel L. Understanding why rheumatoid arthritis patient treatment preferences differ by race. Arthritis Rheum 2009;61:413–8.
18. Greenberg JD, Spruill TM, Shan Y, Reed G, Kremer JM, Potter J, et al. Racial and ethnic disparities in disease activity in patients with rheumatoid arthritis. Am J Med 2013;126:1089–98.
19. Hughes LB, Beasley TM, Patel H, Tiwari HK, Morgan SL, Baggott JE, et al. Racial or ethnic differences in allele frequencies of single-nucleotide polymorphisms in the methylenetetrahydrofolate reductase gene and their influence on response to methotrexate in rheumatoid arthritis. Ann Rheum Dis 2006;65:1213–8.
20. Jacobi CE, Mol GD, Boshuizen HC, Rupp I, Dinant HJ, Van Den Bos GA. Impact of socioeconomic status on the course of rheumatoid arthritis and on related use of health care services. Arthritis Rheum 2003;49:567–73.
21. Johnson BK, Bahce-Altuntas A. Too little too late: effect of poor access to biologics for patients with rheumatoid arthritis. J Rheumatol 2017;44:1765–6.
22. Katz PP, Barton J, Trupin L, Schmajuk G, Yazdany J, Ruiz PJ, et al. Poverty, depression, or lost in translation? Ethnic and language variation in patient-reported outcomes in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:621–8.
23. Kerr GS, Swearingen C, Mikuls TR, Yazici Y. Use of biologic therapy in racial minorities with rheumatoid arthritis from 2 US health care systems. J Clin Rheumatol 2017;23:12–18.
24. Lapcevic M, Vukovic M, Gvozdenovic BS, Mioljevic V, Marjanovic S. Socioeconomic and therapy factor influence on self-reported fatigue, anxiety and depression in rheumatoid arthritis patients. Rev Bras Reumatol Engl Ed 2017;57:545–56.
25. Lee SJ, Kremer J, Kavanaugh A. Treatment disparity related to race/ethnicity and education in rheumatoid arthritis patients: comment on the article by Constantinescu et al [letter]. Arthritis Rheum 2009;61:1141–2.
26. Li X, Sundquist J, Sundquist K. Socioeconomic and occupational risk factors for rheumatoid arthritis: a nationwide study based on hospitalizations in Sweden. J Rheumatol 2008;35:986–91.
27. Margaretten M, Barton J, Julian L, Katz P, Trupin L, Tonner C, et al. Socioeconomic determinants of disability and depression in patients with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2011;63:240–6.
28. McBurney CA, Vina ER. Racial and ethnic disparities in rheumatoid arthritis. Curr Rheumatol Rep 2012;14:463–71.
29. Mikuls TR, Kazi S, Cipher D, Hooker R, Kerr GS, Richards JS, et al. The association of race and ethnicity with disease expression in male US veterans with rheumatoid arthritis. J Rheumatol 2007;34:1480–4.
30. Molina E, Del Rincon I, Restrepo JF, Battafarano DF, Escalante A. Association of socioeconomic status with treatment delays, disease activity, joint damage, and disability in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2015;67:940–6.
31. Parks CG, D'Aloisio AA, DeRoo LA, Huiber K, Rider LG, Miller FW, et al. Childhood socioeconomic factors and perinatal characteristics influence development of rheumatoid arthritis in adulthood. Ann Rheum Dis 2013;72:350–6.
32. Solomon A, Christian BF, Dessein PH. Potential determinants of poor disease outcome in socioeconomically disadvantaged patients with rheumatoid arthritis. J Rheumatol 2008;35:1895–6.
33. Solomon DH, Ayanian JZ, Yelin E, Shaykevich T, Brookhart MA, Katz JN. Use of disease-modifying medications for rheumatoid arthritis by race and ethnicity in the National Ambulatory Medical Care Survey. Arthritis Care Res 2012;64:184–9.
34. Yang G, Bykerk VP, Boire G, Hitchon CA, Thorne JC, Tin D, et al. Does socioeconomic status affect outcomes in early inflammatory arthritis? Data from a canadian multisite suspected rheumatoid arthritis inception cohort. J Rheumatol 2015;42:46–54.
35. Yazici Y, Kautiainen H, Sokka T. Differences in clinical status measures in different ethnic/racial groups with early rheumatoid arthritis: implications for interpretation of clinical trial data. J Rheumatol 2007;34:311–5.
36. Oude Voshaar MA, ten Klooster PM, Taal E, van de Laar MA. Measurement properties of physical function scales validated for use in patients with rheumatoid arthritis: a systematic review of the literature. Health Qual Life Outcomes 2011;9:99.
37. Arat S, Lenaerts JL, De Langhe E, Verschueren P, Moons P, Vandenberghe J, et al. Illness representations of systemic lupus erythematosus and systemic sclerosis: a comparison of patients, their rheumatologists and their general practitioners. Lupus Sci Med 2017;4:e000232.
38. Ward MM, Guthrie LC, Alba MI. Clinically important changes in individual and composite measures of rheumatoid arthritis activity: thresholds applicable in clinical trials. Ann Rheumatic Dis 2015;74:1691–6.
39. Vittinghoff E. Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. 2nd ed. New York: Springer; 2012. p. 82.
40. Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: a clarification. J Clin Epidemiol 1989;42:403–8.

41. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83.
42. Katz P, Pedro S, Michaud K. Performance of the patient-reported outcomes measurement information system 29-item profile in rheumatoid arthritis, osteoarthritis, fibromyalgia, and systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69:1312–21.
43. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.
44. Karpouzas GA, Dolatabadi S, Moran R, Li N, Nicassio PM, Weisman MH. Correlates and predictors of disability in vulnerable US Hispanics with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64:1274–81.
45. Lynch AD, Dodds NE, Yu L, Pilkonis PA, Irrgang JJ. Individuals with knee impairments identify items in need of clarification in the Patient Reported Outcomes Measurement Information System (PROMIS(R)) pain interference and physical function item banks—a qualitative study. Health Qual Life Outcomes 2016;14:77.
46. Orbai AM, Smith KC, Bartlett SJ, De Leon E, Bingham CO, 3rd. "Stiffness has different meanings, I think, to everyone": examining stiffness from the perspective of people living with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2014;66:1662–72.
47. Koster N, Knol DL, Uitdehaag BM, Scheltens P, Sikkes SA. The sensitivity to change over time of the Amsterdam IADL Questionnaire((c)). Alzheimers Dement 2015;11:1231–40.
48. Barton JL, Trupin L, Schillinger D, Gansky SA, Tonner C, Margaretten M, et al. Racial and ethnic disparities in disease activity and function among persons with rheumatoid arthritis from university-affiliated clinics. Arthritis Care Res (Hoboken) 2011;63:1238–46.
49. 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.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 936–948
DOI 10.1002/acr.23762

Biopsy-Proven Small-Fiber Neuropathy in Primary Sjögren's Syndrome: Neuropathic Pain Characteristics, Autoantibody Findings, and Histopathologic Features

Julius Birnbaum, Aliya Lalji, Aveen Saed, and Alan N. Baer

Objective. Painful small-fiber neuropathies (SFNs) in primary Sjögren's syndrome (SS) may present as pure or mixed with concurrent large-fiber involvement. SFN can be diagnosed by punch skin biopsy results that identify decreased intra-epidermal nerve-fiber density (IENFD) of unmyelinated nerves.

Methods. We compared 23 consecutively evaluated patients with SS with pure and mixed SFN versus 98 patients without SFN. We distinguished between markers of dorsal root ganglia (DRG) degeneration (decreased IENFD in the proximal thigh versus the distal leg) versus axonal degeneration (decreased IENFD in the distal leg versus the proximal thigh).

Results. There were no differences in pain intensity, pain quality, and treatment characteristics in the comparison of 13 patients with pure SFN versus 10 patients with mixed SFN. Ten patients with SFN (approximately 45%) had neuropathic pain preceding sicca symptoms. Opioid analgesics were prescribed to approximately 45% of patients with SFN. When compared to 98 patients without SFN, the 23 patients with SFN had an increased frequency of male sex (30% versus 9%; $P < 0.01$), a decreased frequency of anti–Ro 52 ($P = 0.01$) and anti–Ro 60 antibodies ($P = 0.01$), rheumatoid factor positivity ($P < 0.01$), and polyclonal gammopathy ($P < 0.01$). Eleven patients had stocking-and-glove pain, and 12 patients had nonstocking-and-glove pain. Skin biopsy results disclosed patterns of axonal (16 patients) and DRG injury (7 patients).

Conclusion. SS SFN had an increased frequency among male patients, a decreased frequency of multiple antibodies, frequent treatment with opioid analgesics, and the presence of nonstocking-and-glove pain. Distinguishing between DRG versus axonal injury is significant, especially given that mechanisms targeting the DRG may result in irreversible neuronal cell death. Altogether, these findings highlight clinical, autoantibody, and pathologic features that can help to define mechanisms and treatment strategies.

INTRODUCTION

Small-fiber neuropathies (SFNs) in patients with primary Sjögren's syndrome (SS) target nociceptive thinly myelinated Aδ and unmyelinated C-fiber nerves and are frequently associated with burning and allodynic pain (1–3). Pure SFNs in SS may occur without concurrent large-fiber involvement, and may therefore be associated with normal nerve-conduction findings. In addition, previous studies have demonstrated that up to 45% of patients with SS may present with mixed SFNs associated with large-fiber involvement on nerve conduction studies (4).

Measurement of decreased intra-epidermal nerve-fiber density (IENFD) of unmyelinated nerves in punch skin biopsy specimens is a well-validated and highly reproducible diagnostic biomarker of SFN (5–8). Other tests to diagnose SFN in SS require more invasive approaches (i.e., sural-nerve biopsies), may be associated with high interoperator and interpatient variability (i.e., quantitative sensory testing) (3), or have not been widely validated in other disease states (i.e., laser evoked potentials).

Furthermore, skin biopsy results can discriminate between patterns of axonal versus dorsal root ganglia (DRG) degeneration (4,9–13). Despite these advantages, there have been few studies reporting on characteristics of patients with SS with SFN docu-

The Johns Hopkins Rheumatic Disease Research Core Center is funded by the NIH (grant P30-AR-053503) and by the Jerome L. Greene Foundation. Dr. Birnbaum's work was supported by the NIH (grant K23-AR-064279). Dr. Baer's work was supported by the NIH (DE-12354-15A1).

Julius Birnbaum, MD, MHS, Aliya Lalji, MD, Aveen Saed, MD, and Alan N. Baer, MD: Johns Hopkins University School of Medicine, Baltimore, Maryland.

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

Address correspondence to Julius Birnbaum, MD, MHS, Division of Rheumatology/Department of Neurology, The Johns Hopkins University School of Medicine, 5200 Eastern Avenue, Mason F. Lord Building, Center Tower, Suite 4100, Baltimore, MD 21224. E-mail: jbirnba2@jhmi.edu.

Submitted for publication January 2, 2018; accepted in revised form September 11, 2018.

SIGNIFICANCE & INNOVATIONS

- Patients with primary Sjögren's syndrome (SS) with small-fiber neuropathies (SFNs) had an increased frequency of male sex and a decreased frequency of multiple autoantibodies, compared to patients without neuropathy.
- Opioid analgesics were frequently used to treat patients with SS with SFN.
- Whereas the majority of patients with SFN with length-dependent pain (i.e., stocking-and-glove distribution) had punch skin biopsy markers of axonal degeneration, patients with nonlength-dependent pain (i.e., proximal, patchy, diffuse) were equally likely to have skin biopsy markers of dorsal root ganglia or axonal degeneration.

mented by skin biopsy results. Clinicians who care for patients with SS SFN should recognize these highly characteristic clinical and skin-biopsy findings. Therefore, in this study of a well-characterized cohort of 23 patients with SS with biopsy-proven SFN, we report on clinical features of neuropathic pain and skin biopsy findings, and compare demographic and immunologic features to those of 98 patients with SS without SFN.

PATIENTS AND METHODS

Study type. This was a cross-sectional, single-institution cohort of patients with SS enrolled between the years 2008 and 2015. All patients provided informed written consent to participate, and the study methods were approved by the Johns Hopkins Institutional Review Board. All patients were referred to the Jerome L. Greene Sjögren's Syndrome Center at Johns Hopkins. The center includes a neuro-rheumatology clinic, which is dedicated to patients affected by neurologic complications of this disease. Therefore, patients with SS with SFN were evaluated in this setting by 1 of the study authors (JB), who is board-certified as a neurologist and a rheumatologist. Patients with SS without initial suspicion of neuropathies were evaluated by the primary study rheumatologist (ANB). Among this subset, patients could subsequently undergo evaluation for a neuropathy if there was clinical suspicion. The patients with SFN included 2 patients initially evaluated by the study rheumatologist. Altogether, the study cohort consisted of 23 patients with SFN and 98 without SFN.

Inclusion and exclusion criteria. SS was defined by the revised 2002 American-European Consensus Group classification criteria (14), given that patients could be enrolled before publication of the 2012 American College of Rheumatology (ACR) or 2016 ACR/European League Against Rheumatism (EULAR) classification criteria for SS (15,16). Patients were diagnosed as having SFN based on an abnormal IENFD from punch skin biopsy results, as described below.

We excluded 19 patients for the following reasons: 4 with secondary SS, including systemic lupus erythematosus (2 patients) and sarcoidosis (2 patients); 6 with a comorbid disorder that is known to be associated with SFN, including diabetes (3 patients), vitamin B12 deficiency (1 patient), chemotherapy exposure to paclitaxel (1 patient), and varicella zoster virus re-activation (1 patient); and 9 with clinical characteristics of SFN who did not have skin biopsy studies, either due to loss of follow-up or to patient preference.

Serologic assays. Antibody assays for anti–Ro 52, anti–Ro 60, and anti-La/SSB antibodies were performed by the Johns Hopkins Rheumatic Disease Research Core Center, as previously described (17). Other assays were performed by the hospital or commercial laboratories.

Performance of punch skin biopsy. Skin biopsy was performed according to a standardized technique (5,6). Briefly, skin biopsy specimens were obtained using a 3-mm punch from 2 standardized sites, one 10 cm above the lateral malleolus and one from the proximal lateral thigh (5,6). Determination of the IENFD was performed on 50-μm frozen sections after immunostaining axons against the panaxonal protein PGP 9.5.

Evaluation of peripheral nerve status. Peripheral nerve status was characterized with a validated neuropathic pain symptom questionnaire, neurologic examination, nerve-conduction studies to assess large myelinated alpha-beta fibers, and punch skin biopsy results to assess IENFD. The Self-Administered Leeds Assessment of Neuropathic Symptoms and Signs (S-LANSS) pain scale was used (18). Although validated as a scale to help distinguish between nociceptive versus neuropathic pain (18), the S-LANSS provides a 7-item profile of neuropathic descriptors that are frequently used by patients with SFN. Patients report on the presence of burning, pins-and-needles, shock-like sensations, sensitivity to touch, mottling, tenderness upon pressure, and worsening of pain upon rubbing. The neurologic evaluation assessed for weakness, hyporeflexia/arreflexia, sensory ataxia, and abnormal sensory data suggestive of large-fiber findings (i.e., decreased vibration and proprioception) or small-fiber findings (i.e., decreased pinprick and thermal sensation). All patients underwent nerve-conduction studies (evaluation for large-fiber neuropathy) and skin biopsy (evaluation for SFN).

Ascertainment of pure and mixed SFN. A diagnosis of pure SFN was based on a combination of supportive symptoms and examination findings of small-fiber impairment (as described above), normal nerve-conduction studies, and decreased IENFD on skin biopsy studies. As previously described in patients with SS (4), we additionally identified patients as having mixed SFN (with large-fiber involvement)

Table 1. Characteristics of Sjögren's syndrome patients with small-fiber neuropathy (SFN)*

Variable	Cohort (n = 23)	Pure SFN (n = 13)	Mixed SFN (n = 10)	*P*
Age at sicca onset, mean ± SD years	44.0 ± 13.2	38.8 ± 10.3	50.8 ± 14.0	0.03
Age at neuropathic pain onset, mean ± SD years	45.0 ± 13.9	38.3 ± 11.0	56.9 ± 11.5	<0.01
Women	69.6 (16/23)	61.5 (8/13)	80.0 (8/10)	0.41
Ethnicity				0.33
White	73.9 (17/23)	76.9 (10/13)	70.0 (7/10)	–
African American	17.4 (4/23)	23.1 (3/13)	10.0 (1/10)	–
Asian	8.7 (2/23)	0.0 (0/13)	20.0 (2/10)	–
Neuropathic pain before onset of sicca	43.5 (10/23)	38.5 (5/13)	50.0 (5/10)	0.68
Onset of neuropathic pain†				0.72
Acute	17.4 (4/23)	23.1 (3/13)	10.0 (1/10)	–
Subacute	21.7 (5/23)	15.4 (2/13)	30.0 (3/10)	–
Chronic	60.9 (14/23)	61.5 (8/13)	60.0 (6/10)	–
Pain intensity, median (IQR)‡	6.0 (5)	7.0 (3)	5.0 (5.0)	0.80
Neuropathic pain descriptors				
Burning	91.3 (21/23)	100.0 (13/13)	80.0 (8/10)	0.18
Pins-and-needles	87.0 (20/23)	76.9 (10/13)	100.0 (10/10)	0.23
Shock-like sensations	65.2 (15/23)	53.8 (7/13)	80.0 (8/10)	0.38
Sensitivity to touch	60.9 (14/23)	53.8 (7/13)	70.0 (7/10)	0.67
Mottling	43.5 (10/23)	53.8 (7/13)	30.0 (3/10)	0.40
Tenderness upon pressure	39.1 (9/23)	23.1 (3/13)	60.0 (6/10)	0.10
Worsening pain upon rubbing	47.8 (11/23)	38.5 (5/13)	60.0 (6/10)	0.41
Antibody status				
Antinuclear antibody ≥1:320	47.8 (11/23)	38.5 (5/13)	60.0 (6/10)	0.41
Anti–Ro 52	69.6 (16/23)	69.2 (9/13)	70.0 (7/10)	1.00
Anti–Ro 60	52.2 (12/23)	33.3 (4/13)	80.0 (8/10)	0.04
Anti-La/SSB	26.1 (6/23)	15.4 (2/13)	40.0 (4/10)	0.34
Rheumatoid factor	26.1 (6/23)	15.4 (2/13)	40.0 (4/10)	0.34
Polyclonal gammopathy	13.0 (3/23)	7.7 (1/13)	20.0 (2/10)	0.56
Positive lip biopsy	50.0 (8/16)	50.0 (5/10)	50.0 (3/6)	1.00
Medications prescribed prior to SFN evaluation				
Polysymptomatic ≥2 medications	78.3 (18/23)	76.9 (10/13)	80.0 (8/10)	1.00
No. of medications tried, median (IQR)	3 (2)	4 (3)	2 (1)	1.00
Antiepileptic drugs	82.6 (19/23)	46.9 (10/13)	90.0 (9/10)	0.60
Tricyclic antidepressants	8.7 (2/23)	7.7 (1/13)	10.0 (1/10)	1.00
Selective serotonin reuptake inhibitors	17.4 (4/23)	7.7 (1/13)	30.0 (3/10)	0.28
Serotonin-norepinephrine reuptake inhibitor	43.5 (10/23)	53.9 (7/13)	30.0 (3/10)	0.40
Opioid analgesics	43.5 (10/23)	61.5 (8/13)	20.0 (2/10)	0.09
Corticosteroids	13.0 (3/23)	23.1 (3/13)	0.0 (0/10)	0.23

* Values are the percentage (number/total number) unless indicated otherwise. IQR = interquartile range.
† Acute = <1 week to the worst point, subacute = <1 month to the worst point, chronic = >1 month to the worst point.
‡ VAS (range 0–10).

if concurrently presenting with abnormal nerve-conduction studies. To investigate whether differences existed between pure and mixed SFN, we compared demographic, clinical (neuropathic pain intensity, pain quality, and treatment characteristics), and immunologic features.

Evaluation of skin biopsy findings. Biopsy-proven SFN was diagnosed when the IENFD was less than the fifth percentile of normative controls established by our Cutaneous Nerve Laboratory, when assessed at the proximal thigh and/or distal leg (6–8). We used the results of the paired skin biopsies from

each patient to distinguish between patterns of injury suggestive of DRG versus axonal degeneration (4,9–13). IENFD that is reduced or disproportionately decreased at the distal leg versus the proximal thigh is consistent with axonal degeneration (i.e., dying-back axonopathy). In contrast, IENFD that is decreased in the proximal thigh, or is more disproportionately reduced at the proximal thigh compared to the distal leg, is consistent with a primary pattern of DRG injury (4,9–13).

Neuropathic pain treatment. Patients often had difficulty remembering the start and stop dates of previously prescribed medications. Therefore, we defined cumulative treatment with medications as a situation in which patients recalled >1 month duration of therapy, at any time during their prior treatment course.

Statistical analysis. We first compared demographic, clinical, immunologic, and treatment characteristics of patients with pure versus mixed SFN. Differences were assessed by Student's *t*-test or Wilcoxon's rank sum test for continuous variables and by Fisher's exact test or chi-square analyses for categorical variables. We subsequently compared the combined group of 23 patients with SS with both pure and mixed SFN, versus the 98 patients with SS without neuropathy. To determine which among various patient factors were the most predictive of SFN, we used logistic regression with backwards covariate selection. We applied logistic regression to patients' SFN status, removing 1 factor at a time, until all covariates included in the model were significant, with a *P* value of less than 0.20. We expected that some factors would be strongly related to each other, especially the presence of autoantibodies, making this selection process critical for determining which among these predictive factors was marginally the most important. Secondary comparator groups included pure SFN versus mixed SFN. For all analyses, *P* values less than 0.05 (2-tailed) were considered statistically significant. The data analysis was performed using Stata software, version 11.0 (19).

RESULTS

Characteristics of patients with SS. Table 1 shows the neuropathic pain characteristics (intensity of pain, presence of neuropathic pain descriptors, acuity of neuropathic pain onset), demographic features, autoantibody features, and prior treatment regimens for the patients with SS with SFN. Table 2 individually lists these characteristics for each patient.

There were 23 patients with SS SFN, including 13 with pure SFN and 10 with mixed SFN. For these 23 patients, the mean ± SD age at onset of sicca symptoms was 44.0 ± 13.2 years, and there were 16 women (69.6%) and 17 white patients (73.9%). In 10 patients (43.5%), SFN symptoms antedated sicca symptoms by a median of 5.5 (interquartile range [IQR] 9.8) years. Among these 10 patients, there were 4 patients (40%) who presented with SFN symptoms antedating sicca symptoms by more than a decade. The median interval between onset of neuropathic pain symptoms and diagnosis was 4.4 (IQR 7) years. The median intensity of neuropathic pain was 6.0 (IQR 5) on a 0–10 numeric rating scale assessed over the past week (where 0 = no pain and 10 = pain as severe as it could be). Onset of neuropathic pain was reported as acute (<1 week to nadir) in 4 patients (17.4%), subacute (<1 month to nadir) in 5 patients (21.7%), and chronic (>1 month to nadir) in 14 patients (60.9%). Symptom descriptors from the S-LANSS included burning in 21 patients (91.3%), pins-and-needles in 20 patients (87.0%), shock-like sensations in 15 patients (65.2%), sensitivity to touch in 14 patients (60.9%), mottling in 10 patients (43.5%), tenderness upon pressure in 9 patients (39.1%), and worsening of pain upon rubbing in 11 patients (47.8%).

In a comparison of the patients with pure versus mixed SFN, Table 1 shows that patients with pure SFN were younger at the age of onset of sicca symptoms (mean ± SD 38.8 ± 10.3 years versus 50.8 ± 14.0 years; $P = 0.03$), and at the age of onset of neuropathic pain (mean ± SD 38.3 ± 11.0 years versus 56.9 ± 11.5 years; $P < 0.01$). There were no differences in frequency of female sex and white race. Similarly, there were no statistically significant differences with regard to neuropathic pain symptoms, intensity, or acuity of onset. Additionally, patients with pure and mixed SFN did not differ with regard to most immunologic characteristics, including frequencies of antinuclear antibody ≥1:320, anti–Ro 52 and anti-La/SSB antibodies, rheumatoid factor positivity, polyclonal gammopathy, and positive lip biopsy results. Patients with pure SFN did have a lower frequency of anti–Ro 60 antibodies (33.3% [4 of 13] versus 80.0% [8 of 10]; $P = 0.04$).

Table 2 further shows patterns of neurologic examination and nerve-conduction findings. As shown, 4 patients had concurrent large-fiber sensory neuropathies (patients 17, 19, 20, and 21) and presented with arreflexia, positive Romberg test, gait ataxia, and absent sensory nerve action potentials (SNAPs). Four patients (patients 1, 8, 9, and 15) had nerve-conduction studies consistent with symmetric axonal sensorimotor polyneuropathies. However, none of these 4 patients had detectable weakness on examination, indicating that involvement of motor nerves was subclinical. One patient (patient 3) had an axonal sensory polyneuropathy with large-fiber deficits and decreased sural SNAPs. One patient (patient 6) had presented with a mononeuritis multiplex 7 years previously, was treated with cyclophosphamide for 3 months, had resolution of lower-extremity dysesthesias and partial recovery of right foot drop (2 of 5 on the Medical Research Council Scale), and for more than 6 years had no new neurologic symptoms until the onset of bilateral burning foot pain 1 year prior to the current evaluation.

Table 2. Initial presentation of 23 Sjögren's syndrome patients with small-fiber neuropathy*

Demographics	Symptoms, duration	Neuropathic pain, characteristics, abnormal findings	Nerve conduction studies	Acuity of onset	Pain severity†	Peripheral neuropathy, therapy
Patient 1: age 41 years, F, Asian	2 years	LD distal legs; pins-and-needles; sensitivity to touch; shock-like sensations; burning; tenderness upon pressure; Vib-L‡, JP-L‡; PT reduced	Axonal sensorimotor polyneuropathy; SNAPs: sural: absent; median: reduced, 9–8μV; ulnar: reduced, 4–5μV; CMAPs: peroneal: reduced; tibial: reduced	Subacute	4	Gabapentin
Patient 2: age 47 years, F, white	3 years	LD distal feet; pins-and-needles; mottling; sensitivity to touch; shock-like sensations; burning; PT reduced	Normal	Chronic	2	None
Patient 3: age 72 years, F, white	2 years	LD legs and fingers; pins-and-needles; Vib-L§; PT reduced hyporeflexia (Achilles)	Axonal sensory neuropathy; SNAPs: sural: reduced, 3μV	Acute	9	Gabapentin; fentanyl
Patient 4: age 42 years, M, AA	2 years	LD distal feet; burning; PT reduced	Normal	Chronic	4	Pregabalin; oxycodone
Patient 5: age 51 years, M, white	32 years	LD distal feet; mottling; sensitivity to touch; burning; worsening of pain upon rubbing; Vib-L§; PT reduced	Normal	Acute	7	Gabapentin; prednisone ≥10 mg/day IVIG
Patient 6: age 66 years, M, white	1 year	LD distal legs; pins-and-needles; sensitivity to touch; shock-like sensations; burning; worsening of pain upon rubbing; tenderness upon pressure; weakness of tibialis anterior; on right 2/5 on MRC scale; Vib-L§, JP-L‡; PT reduced; areflexia (Achilles, patellar)	Asymmetric axonal sensorimotor polyneuropathy; SNAPs: sural: absent; ulnar: reduced (right); CMAPs: peroneal: reduced (right)	Chronic	10	Gabapentin; escitalopram
Patient 7: age 60 years, M, white	20 years	LD arms and legs; pins-and-needles; mottling; shock-like sensations; burning; worsening of pain upon rubbing; tenderness upon pressure; Vib-L¶, Vib-U¶, JP-L§; PT reduced; hyporeflexia (Achilles)	Normal	Chronic	5	Gabapentin; pregabalin; venlafaxine; escitalopram; oxycodone; fentanyl
Patient 8: age 63 years, M, white	5 years	LD distal feet; pins-and-needles; sensitivity to touch; worsening of pain upon rubbing; Vib-L§; PT reduced	Axonal sensorimotor polyneuropathy; SNAPs: sural: reduced, 2μV; CMAPs: peroneal: reduced	Chronic	4	Gabapentin; fentanyl
Patient 9: age 75 years, F, white	2 years	LD distal legs; pins-and-needles; sensitivity to touch; shock-like sensations; burning; worsening of pain upon rubbing; tenderness upon pressure; Vib-L‡, JP-L‡; PT reduced	Axonal sensorimotor polyneuropathy; SNAPs: sural: reduced, 3–4μV; median: reduced, 4μV; ulnar: reduced; CMAPs: peroneal: absent; tibial: reduced	Chronic	5	Pregabalin; duloxetine; oxycodone; mycophenolate

(continued)

Table 2. (*cont'd*)

Demographics	Symptoms, duration	Neuropathic pain, characteristics, abnormal findings	Nerve conduction studies	Acuity of onset	Pain severity†	Peripheral neuropathy, therapy
Patient 10: age 53 years, F, AA	3 years	LD legs and fingers; pins-and-needles; sensitivity to touch; shock-like sensations; burning; tenderness upon pressure; PT reduced	Normal	Chronic	10	Gabapentin; pregabalin; duloxetine; morphine sulfate
Patient 11: age 60 years, F, white	13 years	LD distal feet; pins-and-needles; mottling; shock-like sensations; burning; PT reduced	Normal	Chronic	8	Gabapentin; escitalopram; hydrocodone; cyclophosphamide
Patient 12: age 40 years, F, white	25 years	NLD face, arms, legs; pins-and-needles; mottling; sensitivity to touch; burning; worsening of pain upon rubbing; Vib-L¶, JP-L‡, JP-U§; PT reduced	Normal	Chronic	6	Gabapentin; pregabalin; venlafaxine; oxycodone; fentanyl
Patient 13: age 40 years, M, white	2 years	NLD distal feet and proximal thighs; pins-and-needles; mottling; sensitivity to touch; shock-like sensations; burning; worsening of pain upon rubbing; PT reduced; hyporeflexia (Achilles)	Normal	Chronic	4	Prednisone <10 mg/day
Patient 14: age 43 years, F, white	11 years	NLD thighs and feet; pins-and-needles; burning; PT reduced	Normal	Chronic	7	Pregabalin; gabapentin; duloxetine; amitriptyline; carbamazepine; opioid analgesics; prednisone ≥10 mg/day
Patient 15: age 42 years, F, white	6 years	NLD chest, back, hands and feet; pins-and-needles; shock-like sensations; burning; PT reduced	Axonal sensorimotor polyneuropathy; SNAPs: sural: absent; median: absent; ulnar: absent; CMAPs: peroneal: absent; tibial: reduced, 0.1mV	Chronic	7	Pregabalin; duloxetine; topiramate; hydrocodone as needed
Patient 16: age 60 years, F, white	22 years	NLD face and legs; shock-like sensations; burning; Vib-L‡; PT reduced	Normal	Acute	4	Gabapentin; duloxetine; oxycodone; rituximab
Patient 17: age 55 years, F, Asian	2 years	NLD face, arms, legs; pins-and-needles; burning; mottling; shock like sensations; tenderness upon pressure; Vib-L¶; PT reduced; areflexia (Achilles, patellar); Romberg/sensory ataxia	Sensory neuronopathy; SNAPs: sural: absent; median: absent; ulnar: absent; radial: absent	Subacute	1	Gabapentin; sertraline
Patient 18: age 40 years, M, white	2 years	NLD part of face, patchy thighs, legs; pins-and-needles; mottling; burning; sensitivity to touch; worsening of pain upon rubbing; tenderness upon pressure; PT reduced; hyporeflexia (Achilles)	Normal	Acute	9	Gabapentin; milnacipran; fentanyl

(continued)

Table 2. (*cont'd*)

Demographics	Symptoms, duration	Neuropathic pain, characteristics, abnormal findings	Nerve conduction studies	Acuity of onset	Pain severity†	Peripheral neuropathy, therapy
Patient 19: age 65 years, F, white	5 years	NLD arms, asymmetric legs; pins-and-needles; sensitivity to touch; shock like sensations; burning; worsening of pain upon rubbing; Vib-L¶, JP-L§; PT reduced; areflexia (Achilles); Romberg/sensory ataxia	Sensory neuronopathy; SNAPs: sural: absent; median: absent; radial: absent; ulnar: absent	Chronic	5	Gabapentin; venlafaxine; topiramate
Patient 20: age 59 years, F, white	2 years	NLD distal feet, shins; pins-and needles; mottling; sensitivity to touch; shock-like sensations; burning; worsening of pain upon rubbing; tenderness upon pressure; Vib-L¶, Vib-U‡, JP-L§; PT reduced; areflexia (Achilles, patellar, upper limb); Romberg/ sensory ataxia	Sensory neuronopathy; SNAPs: sural: absent; median: absent; radial: reduced, 6μV; ulnar: absent	Chronic	9	Gabapentin IVIG
Patient 21: age 70 years, F, AA	5 years	NLD asymmetric feet and bilateral hands; pins-and-needles; mottling; sensitivity to touch; shock-like sensations; burning; worsening of pain upon rubbing; tenderness upon pressure; Vib-L¶, Vib-U¶, JP-L¶; PT reduced; areflexia (Achilles); Romberg/sensory ataxia	Sensory neuronopathy; SNAPs: sural: reduced; median: absent; radial: absent; ulnar: absent; CMAPs: ulnar: reduced, 40 meters/second	Subacute	5	Pregabalin; amitriptyline; sertraline
Patient 22: age 58 years, F, AA	6 years	NLD arms, hands, legs, feet; pins-and-needles; sensitivity to touch; shock-like sensations; burning; PT reduced; areflexia (Achilles)	Normal	Subacute	8	Gabapentin; tramadol
Patient 23: age 32 years, F, white	1 year	NLD over eyes, hands, feet; pins-and-needles; burning; PT reduced; hyporeflexia (Achilles)	Normal	Subacute	6	Pregabalin; venlafaxine XR; celecoxib

* Only abnormal physical examination findings are reported on this table. F = female; LD = length-dependent pain distribution; Vib-L = vibration of lower limb; JP-L = joint position in lower limb; PT = pinprick and temperature sensation; SNAPs= sensory nerve action potentials, CMAPs= compound motor action potentials; AA = African American; M = male; IVIG = intravenous immunoglobulin; MRC = Medical Research Council scale; Vib-U = vibration of the upper limb; NLD = nonlength-dependent pain distribution; JP-U = joint position in upper limb.
† VAS (range 0–10).
‡ Reduced at the level of the ankle/MCP.
§ Reduced at the level of the foot/DIP.
¶ Reduced above the ankle/wrist.

Prior treatment of patients with pure and mixed SFN. By the time of evaluation, 78.3% of patients (18 of 23) had cumulatively required symptomatic treatment with ≥2 medications (Table 1). Antiepileptic drugs were the most frequently prescribed and taken by 82.6% of patients (19 of 23). Opioid analgesics were previously prescribed for 43.5% of patients (10 of 23) and were prescribed more frequently than selective serotonin reuptake inhibitors (17.4% of patients [4 of 23]) and at a similar frequency as serotonin-norepinephrine reuptake inhibitors (43.5% of patients [10 of 23]). Tricyclic antidepressants were taken by 8.7% of patients (2 of 23). Opioid analgesics were prescribed more frequently for patients with pure SFN versus mixed SFN (61.5% [8 of 13] versus 20.0% [2 of 10]; $P = 0.09$). There were otherwise no differences in the frequency of polysymptomatic therapy and the classes of medications prescribed for neuropathic pain.

As shown in Table 2, 5 patients had previously been prescribed immunomodulatory therapy for neuropathic pain. This group included 3 patients treated with corticosteroids (patients

5, 13, and 14), who did not report any improvement in pain intensity when treated with prednisone at maximal dosages ≥10 mg per day (patients 5 and 14), or who did not report any benefit from a maximal prednisone dosage of <10 mg per day (patient 13). One patient (patient 20) was prescribed intravenous immunoglobulin (IVIG) previously but discontinued it after 6 months due to lack of improvement in pain. Another patient (patient 5) who was prescribed monthly IVIG therapy (2 gm/kg) reported a 50% reduction in pain intensity after 4 months and was therefore prescribed monthly IVIG maintenance therapy. Another patient (patient 16) received 4 courses of rituximab (1,000 mg × 2 every 2 weeks, repeated at 6-month intervals), but reported pain relief mainly within 2 months of drug administration, which worsened over the ensuing 4 months prior to subsequent infusions.

Characterization of patients with SS with SFN versus patients with SS without SFN. Table 3 shows the association of demographic features and immunologic characteristics in the combined 23 patients with SS with pure and mixed SFN versus the 98 patients without SFN. No difference was found with regard to mean age at the time of sicca onset and frequency of white race. However, the combined SFN patient group had an increased frequency of male sex (30.4% versus 9.2%; $P < 0.01$), decreased frequency of anti–Ro 52 (69.6% versus 89.7%; $P = 0.01$), anti–Ro 60 (52.2% versus 79.7%; $P = 0.01$), anti-La/SSB (26.1% versus 50.0%; $P = 0.04$), rheumatoid factor (26.1% versus 61.9%; $P < 0.01$), and polyclonal gammopathy (13.0% versus 48.0%; $P < 0.01$). On multivariate analysis, male sex ($P = 0.04$), anti–Ro 52 ($P = 0.01$), and rheumatoid factor ($P = 0.04$) were independently associated with combined SFN.

Table 3. Demographic, clinical, autoantibody, and other immunologic characteristics in Sjögren's patients with small fiber neuropathy (SFN) and Sjögren's patients without SFN*

Variable	SFN (n = 23)	Non-SFN (n = 98)	*P*
Age onset of sicca symptoms, mean ± SD years	44.0 ± 13.2	42.3 ± 16.8	0.63
Women	69.6 (16)	90.8 (89)	0.007
Ethnicity			
White	73.9 (17)	80.6 (79)	0.48
African American	17.4 (4)	8.2 (8)	–
Hispanic	0.0 (0)	8.2 (8)	–
Asian	8.7 (2)	3.1 (3)	
Dry eyes	91.3 (21)	95.9 (94)	0.32
Dry mouth	95.7 (22)	90.8 (89)	0.69
Antinuclear antibody			0.23
≤1:40	26.1 (6)	19.4 (19)	–
1:80	8.7 (2)	3.1 (3)	–
1:160	17.4 (4)	5.1 (5)	–
1:320	8.7 (2)	17.4 (17)	–
≥1:640	39.1 (9)	55.1 (54)	–
Anti–Ro 52	69.6 (16)	89.7 (87/97)†	0.01
Anti–Ro 60	52.2 (12)	79.7 (78)	0.01
Anti-La/SSB	26.1 (6)	50.0 (48/96)†	0.04
Rheumatoid factor	26.1 (6)	61.9 (60/97)†	0.002
Polyclonal gammopathy	13.0 (3)	48.0 (47)	0.002
Monoclonal gammopathy	4.4 (1)	14.3 (14)	0.30
C3 hypocomplementemia, % (no./total no.)	0.0 (0/22)	4.2 (4/95)	1.00
C4 hypocomplementemia, % (no./total no.)	4.6 (1/22)	9.5 (9/95)	0.69
Positive lip biopsy, % (no./total no.)	50.0 (8/16)	69.8 (44/63)	0.14

* Values are the percentage (number) unless indicated otherwise. Altogether, 8 of the 98 patients without SFN underwent nerve-conduction and/or skin-biopsy studies. These nerve-conduction findings were either normal (2 patients) or evidenced focal, noninflammatory, and structural disorders, including entrapment neuropathies (2 patients with median neuropathies, 1 with ulnar neuropathy), cervical radiculopathies (2 patients), and a traumatic peroneal mononeuropathy at the ankle (1 patient). Among these 8 patients, 4 underwent skin-biopsy studies with normal results. These findings were therefore not consistent with a large-fiber polyneuropathy or small-fiber neuropathy.
† Values are the percentage (no./total no.).

Patterns of neuropathic pain and skin biopsy findings. As shown in Figure 1A, 11 patients with SFN presented with length-dependent pain. This pattern is analogous to a stocking-and-glove pain distribution. In patients with a length-dependent pattern of pain, the presumed anatomic site of injury is the most distal axon (4–9,13). Among these 11 patients, 10 did have corroborating skin-biopsy findings indicative of axonal degeneration, and only 1 patient had findings consistent with DRG degeneration (Figures 2A and 2B). In these patients with length-dependent pain, findings consistent with axonal injury

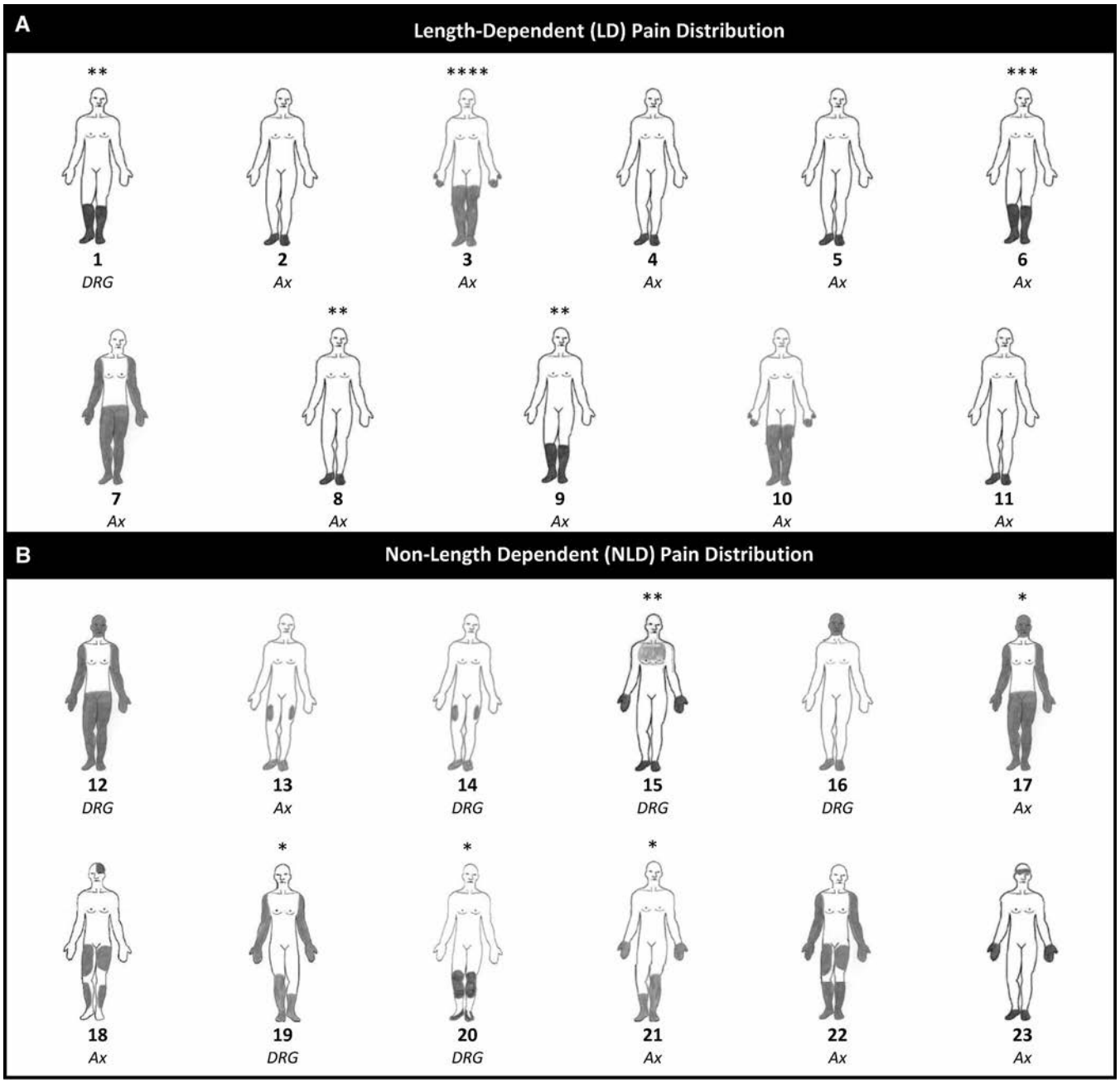


Figure 1. Distribution of neuropathic pain in patients with abnormal skin biopsy results. **A,** 11 patients presented with a stocking-and-glove pattern of neuropathic pain, frequently referred to as a length-dependent pattern of pain. In this pattern, the clinical site of injury may be suspected to reflect axonal (Ax) damage. In 10 of these 11 patients, the skin biopsy results similarly showed patterns of axonal injury. **B,** Twelve patients described a contrasting pattern of neuropathic pain termed nonlength-dependent pain distribution, occurring in a proximal, patchy, and/or asymmetric pattern. In this pattern, the clinical site of injury may be suspected to reflect dorsal root ganglia (DRG) injury. In 6 of 12 patients, skin biopsy results showed findings of DRG degeneration. In contrast, the remaining 6 of the 12 patients with nonlength-dependent pain had biopsy results suggestive of axonal injury. Co-occurring peripheral nerve syndromes: * = sensory neuropathy; ** = axonal sensorimotor (SM) polyneuropathy; *** = asymmetric axonal SM polyneuropathy; **** = axonal sensory neuropathy.

included reduced IENFD at the distal leg with normal IENFD at the proximal thigh (70% of patients [7 of 10]), or more severely and disproportionately reduced IENFD compared to reduced IENFD at the proximal thigh (30% of patients [3 of 10]) (4,8–13,20).

Figure 1B also shows that the 12 remaining patients with SS SFN presented with a contrasting pattern of nonlength-dependent (NLD) pain. This pattern is analogous to pain occurring in a nonstocking-and-glove distribution. Such NLD neuropathic pain could be asymmetric (3 patients), and affect the face (5 patients), upper extremities (7 patients), and torso (1 patient), and could occur simultaneously with or without pain in the feet. In patients presenting with this NLD pattern of pain, the presumed anatomic site of injury is the DRG (4,9–13). Among these

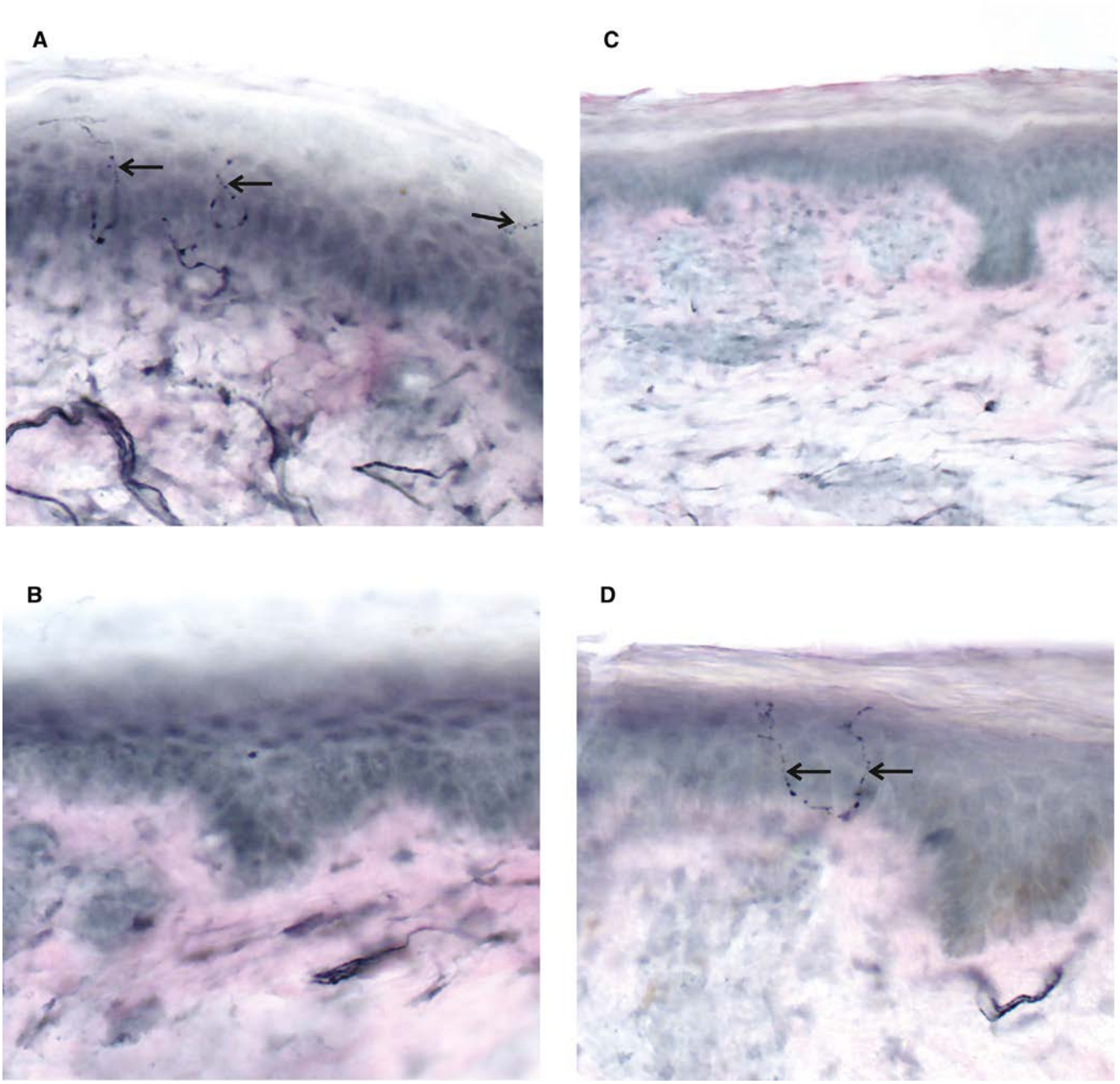


Figure 2. Skin biopsy specimens associated with patterns of neuropathic pain. Arrows indicate unmyelinated nerves immunostained against the panaxonal marker PGP 9.5. Skin biopsy specimens taken from the proximal thigh (**A**) and distal leg (**B**) in a patient with a length-dependent pattern of neuropathic pain. Decreased intra-epidermal nerve-fiber density (IENFD) was noted in the distal leg versus the proximal thigh, consistent with a pattern of axonal degeneration. Skin biopsy specimens taken from the proximal thigh (**C**) and distal leg (**D**) in a patient with a nonlength-dependent pattern of neuropathic pain. Decreased IENFD was noted in the proximal thigh versus the distal leg, consistent with a pattern of dorsal root ganglia degeneration.

12 patients, 6 (50%) had skin biopsy results with corroborative findings of DRG degeneration. These patients had decreased IENFD that was disproportionately more severe at the proximal thigh compared to the distal leg (Figures 2C and 2D). However, 6 of these patients (50%) who presented with an NLD pain distribution and clinical suspicion of DRG injury instead had skin biopsy findings suggestive of axonal degeneration. Mechanistic implications of these findings are discussed below. Additionally, a representative skin-biopsy figure from a patient without SFN SS is shown in Supplementary Figure 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23762/abstract.

DISCUSSION

To the best of our knowledge, this study was the largest cohort of patients with SS who had biopsy-proven SFN. Among the 23 patients with SS SFN, notable features included a higher prevalence of male patients, frequent prescriptions for opioid analgesics, a decreased frequency of autoantibodies, and clinicopathologic patterns of neuropathic pain that suggest mechanisms and anatomic sites of injury. These findings and treatment implications are considered below.

A previously unreported finding is the increased frequency of male sex among those with combined SFN, which is notable given that SS is otherwise a disease in which >90% of patients are female. Therefore, greater scrutiny for the possibility of SS may be warranted in male patients with SFN. The revised 2016 ACR/EULAR SS classification criteria can especially facilitate diagnosis when male patients with SS SFN present without sicca symptoms or potentially diagnostic antibodies (16). Neurologists can therefore play an important role in identifying SS among patients otherwise considered to have idiopathic SFN.

The increased frequency of male sex in patients with SS with SFN may reflect hormonal as well as immune-mediated influences. For example, an increased frequency of male sex has been noted among patients with chronic pain syndromes associated with antibodies that may cross-react with proteins upregulated in the prostate (i.e., anti-CASPR2 antibodies) (20,21). Such antibodies have been identified in patients with SFN (22,23). Therefore, further studies are warranted to evaluate whether the increased frequency of SFN in male patients with SS may reflect a different clinical spectrum and pathogenesis compared to that found in female patients with SS SFN.

Overall, our findings identify several characteristics of patients with SFN that can explain why diagnosis can be elusive. Similar to the results of prior studies (24–26), we detected that approximately 40% of patients with SS SFN experienced SFN symptoms before the emergence of sicca symptoms. These SFN symptoms preceded the diagnosis of SS by a median of 5 years. This finding also has important implications. Even when patients with SFN lack sicca symptoms and do not initially satisfy criteria for SS, our findings suggest that clinical vigilance for the possibility of SS is warranted if patients develop sicca symptoms detectable upon careful longitudinal appraisal. Additionally, when SS SFN presents as an NLD distribution, clinicians may deem this pain distribution as not being plausible or real (4,8–13). This result emphasizes the need for rheumatologists and other clinicians who care for patients with SS SFN to be cognizant of this NLD presentation, as well as the more familiar length-dependent, stocking-and-glove pattern.

Similar to results from Chai et al (4), we found that approximately 45% of patients with SS presented with mixed SFN. Whereas all of our patients with a sensory neuronopathy had a sensory ataxia, none of our patients with electrodiagnostic evidence of axonal sensorimotor polyneuropathies had weakness. These findings support the idea that patients with mixed SFN could have subclinical large-nerve fiber involvement, with neuropathic pain symptoms and findings primarily reflecting a clinically relevant SFN. Patients with mixed versus pure SFN were older at the onset of sicca and neuropathic pain symptoms, and they experienced an increased frequency of anti–Ro 60 antibodies. We otherwise did not find other discriminating features in patients with pure versus mixed SFN.

We noted that approximately 45% of patients with SS SFN were previously treated with opioid analgesic therapy. Opioid analgesics may contribute to a centrally sensitized pain state that may be refractory to other symptomatic approaches (27). Potential mechanisms include activation of microglial cells (27) and the central glutaminergic system (28). Therefore, our findings similarly suggest that opioid therapy in SS SFN should be prescribed cautiously and after treatment with other symptomatic therapies has been exhausted.

There are 2 primary advantages of skin biopsy. First, skin biopsy has desirable diagnostic features. Compared to other tests used in evaluating SFN (quantitative sensory testing, laser-evoked potentials) (25,26,29), skin biopsy has a comparably increased diagnostic efficiency (80–90% sensitivity, 90% specificity) (3,6) and higher validity. The intraobserver and interobserver intraclass correlation coefficient in our Cutaneous Nerve Laboratory is 0.90. Neuropathology laboratories that can interpret skin-biopsy studies are available in many academic centers. However, for clinicians practicing in locales in which there are no academic neurocutaneous laboratories, there are commercial neuropathology laboratories (such as Therapath Neurobiology) that can process and interpret skin-biopsy studies.

In addition, knowledge of neuropathic pain patterns and skin-biopsy findings may suggest anatomic regions of injury and underlying mechanisms. For example, we identified the fact that 10 of 11 patients with length-dependent pain had skin-biopsy patterns of axonal degeneration. This finding is consistent with other studies noting axonal degeneration occurring in patients with length-dependent pain. In contrast, we discovered that 6 of 12 patients with NLD pain had skin-biopsy results showing DRG degeneration. This finding is compatible with previous reports,

which show how this NLD pain pattern is associated with skin-biopsy markers of DRG degeneration in other immune-mediated SFN diseases (9–13).

By contrast, 6 of the remaining 12 patients with NLD pain (suspected as having DRG degeneration) instead had skin-biopsy markers of axonal injury. In these patients, there was no skin-biopsy evidence of DRG degeneration. In this situation, non-neurotoxic mechanisms that do not cause DRG degeneration may sensitize viable DRGs. One such important mechanism is DRG neuronal hyperexcitability, which may occur due to gain-of-function mutations or may be mediated by antineuronal antibodies (30,31). Prolonged excitation may cause a secondary axonal injury (23) and may account for apparent axonal degeneration in some NLD pain patients.

Our findings suggest therapeutic challenges and treatment strategies. First, there may only be a limited window to intervene before irreversible damage to the DRG occurs. A postmortem biopsy result from an SS patient with NLD pain showed extensive CD8+ T cell–mediated cytotoxic injury to the DRG (32). Such cytotoxic injury, which may be irreversible, reflects how earlier recognition of SFN and the performance of skin-biopsy studies may identify SFN at an earlier and potentially more treatable stage. Additionally, neuropathic pain medications that may be effective in other SFN disorders may worsen SS-associated sicca symptoms, fatigue, and cognitive impairment. Such side effects may be mitigated by slowly increasing dosages of medications and by using medications that have different mechanisms of action.

Limitations of our study included its cross-sectional design. Therefore, we are planning longitudinal studies to evaluate the predictive role of skin-biopsy results with different outcomes in patients with SS SFN. Further studies should also evaluate the association between SFN and other lip biopsy findings, including ectopic lymphoid follicles, patterns of cellular infiltration, and measurements of duct blockage (33). Additionally, our findings suggest the rationale for developing computer programs to standardize pain topography occurring across a large number of patients with SS SFN.

Our findings therefore collectively indicate how SFN is associated with highly characteristic clinical and skin-biopsy findings. In summary, we have described the largest cohort of patients with SS with biopsy-proven SFN and have described the increased frequency of male sex, an association with decreased antibody frequencies, and overlapping patterns of DRG and axonal injury across different patterns of neuropathic pain.

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. Birnbaum 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. Birnbaum, Lalji, Saed, Baer.

Acquisition of data. Birnbaum, Lalji, Saed, Baer.

Analysis and interpretation of data. Birnbaum, Lalji, Saed, Baer.

REFERENCES

1. Birnbaum J. Peripheral nervous system manifestations of Sjogren syndrome: clinical patterns, diagnostic paradigms, etiopathogenesis, and therapeutic strategies. Neurologist 2010;16:287–97.
2. Pavlakis PP, Alexopoulos H, Kosmidis ML, Stamboulis E, Routsias JG, Tzartos SJ, et al. Peripheral neuropathies in Sjogren syndrome: a new reappraisal. J Neurol Neurosurg Psychiatry 2011;82:798–802.
3. Chan AC, Wilder-Smith EP. Small fiber neuropathy: getting bigger! Muscle Nerve 2016;53:671–82.
4. Chai J, Herrmann DN, Stanton M, Barbano RL, Logigian EL. Painful small-fiber neuropathy in Sjogren syndrome. Neurology 2005;65:925–7.
5. Ebenezer GJ, Hauer P, Gibbons C, McArthur JC, Polydefkis M. Assessment of epidermal nerve fibers: a new diagnostic and predictive tool for peripheral neuropathies. J Neuropathol Exp Neurol 2007;66:1059–73.
6. McArthur JC, Stocks EA, Hauer P, Cornblath DR, Griffin JW. Epidermal nerve fiber density: normative reference range and diagnostic efficiency. Arch Neurol 1998;55:1513–20.
7. Devigili G, Tugnoli V, Penza P, Camozzi F, Lombardi R, Melli G, et al. The diagnostic criteria for small fibre neuropathy: from symptoms to neuropathology. Brain 2008;131:1912–25.
8. Lauria G, Hsieh ST, Johansson O, Kennedy WR, Leger JM, Mellgren SI, et al. European Federation of Neurological Societies/Peripheral Nerve Society guideline on the use of skin biopsy in the diagnosis of small fiber neuropathy: report of a joint task force of the European Federation of Neurological Societies and the Peripheral Nerve Society. Eur J Neurol 2010;17:903–12.
9. Khan S, Zhou L. Characterization of non-length-dependent small-fiber sensory neuropathy. Muscle Nerve 2012;45:86–91.
10. Gemignani F, Giovanelli M, Vitetta F, Santilli D, Bellanova MF, Brindani F, et al. Non-length dependent small fiber neuropathy. a prospective case series. J Peripher Nerv Syst 2010;15:57–62.
11. Gorson KC, Herrmann DN, Thiagarajan R, Brannagan TH, Chin RL, Kinsella LJ, et al. Non-length dependent small fibre neuropathy/ganglionopathy. J Neurol Neurosurg Psychiatry 2008;79:163–9.
12. Brannagan TH III, Hays AP, Chin SS, Sander HW, Chin RL, Magda P, et al. Small-fiber neuropathy/neuronopathy associated with celiac disease: skin biopsy findings. Arch Neurol 2005;62:1574–8.
13. Oomatia A, Fang H, Petri M, Birnbaum J. Peripheral neuropathies in systemic lupus erythematosus: clinical features, disease associations, and immunologic characteristics evaluated over a twenty-five–year study period. Arthritis Rheumatol 2014;66:1000–9.
14. Vitali C, Bombardieri S, Jonsson R, Moutsopoulos HM, Alexander EL, Carsons SE, et al. Classification criteria for Sjogren's syndrome: a revised version of the European criteria proposed by the American-European consensus group. Ann Rheum Dis 2002;61:554–8.
15. Shiboski SC, Shiboski CH, Criswell LA, Baer AN, Challacombe S, Lanfranchi H, et al, for the Sjögren's International Collaborative Clinical Alliance (SICCA) Research Groups. American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance Cohort. Arthritis Care Res (Hoboken) 2012;64:475–87.
16. Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome: a consensus and data-driven

methodology involving three international patient cohorts. Arthritis Rheumatol 2017;69:35–45.
17. Nagaraju K, Cox A, Casciola-Rosen L, Rosen A. Novel fragments of the Sjögren's syndrome autoantigens α-fodrin and type 3 muscarinic acetylcholine receptor generated during cytotoxic lymphocyte granule–induced cell death. Arthritis Rheum 2001;44:2376–86.
18. Bennett MI, Smith BH, Torrance N, Potter J. The S-LANSS score for identifying pain of predominantly neuropathic origin: validation for use in clinical and postal research. J Pain 2005;6:149–58.
19. Stata statistical software: release 11.0. College Station (TX): StataCorp; 2011.
20. Van Sonderen A, Arino H, Petit-Pedrol M, Leypoldt F, Kortvelyessy P, Wandinger KP, et al. The clinical spectrum of Caspr2 antibody-associated disease. Neurology 2016;87:521–8.
21. Irani SR, Pettingill P, Kleopa KA, Schiza N, Waters P, Mazia C, et al. Morvan syndrome: clinical and serological observations in 29 cases. Ann Neurol 2012;72:241–55.
22. Klein CJ, Lennon VA, Aston PA, McKeon A, Pittock SJ. Chronic pain as a manifestation of potassium channel-complex autoimmunity. Neurology 2012;79:1136–44.
23. Lahoria R, Pittock SJ, Gadoth A, Engelstad JK, Lennon VA, Klein CJ. Clinical-pathologic correlations in voltage-gated Kv1 potassium channel complex-subtyped autoimmune painful polyneuropathy. Muscle Nerve 2017;55:520–5.
24. Birnbaum J, Duncan T, Owoyemi K, Wang KC, Carrino J, Chhabra A. Use of a novel high-resolution magnetic resonance neurography protocol to detect abnormal dorsal root ganglia in Sjogren patients with neuropathic pain: case series of 10 patients and review of the literature. Medicine (Baltimore) 2014;93: 121–34.
25. Mori K, Iijima M, Koike H, Hattori N, Tanaka F, Watanabe H, et al. The wide spectrum of clinical manifestations in Sjogren's syndrome-associated neuropathy. Brain 2005;128:2518–34.
26. Sene D, Cacoub P, Authier FJ, Haroche J, Creange A, Saadoun D, et al. Sjogren syndrome-associated small fiber neuropathy: characterization from a prospective series of 40 cases. Medicine (Baltimore). 2013;92:e10–8.
27. Miller RJ, Jung H, Bhangoo SK, White FA. Cytokine and chemokine regulation of sensory neuron function. Handb Exp Pharmacol 2009;194:417–49.
28. Mao J, Price D, Mayer D. Mechanisms of hyperalgesia and morphine tolerance: a current view of their possible interactions. Pain 1995;62:259–74.
29. Mori K, Iijima M, Sugiura M, Koike H, Hattori N, Ito H, et al. Sjogren's syndrome associated painful sensory neuropathy without sensory ataxia. J Neurol Neurosurg Psychiatry 2003;74:1320–2.
30. Faber CG, Hoeijmakers JG, Ahn HS, Cheng X, Han C, Choi JS, et al. Gain of function Nanu1.7 mutations in idiopathic small fiber neuropathy. Ann Neurol 2012;71:26–39.
31. Huang J, Han C, Estacion M, Vasylyev D, Hoeijmakers JG, Gerrits MM, et al. Gain-of-function mutations in sodium channel na(v)1.9 in painful neuropathy. Brain 2014;137:1627–42.
32. Kawagashira Y, Koike H, Fujioka Y, Hashimoto R, Tomita M, Morozumi S, et al. Differential, size-dependent sensory neuron involvement in the painful and ataxic forms of primary Sjogren's syndrome-associated neuropathy. J Neurol Sci 2012;319:139–46.
33. Fisher BA, Brown RM, Bowman SJ, Barone F. A review of salivary gland histopathology in primary Sjögren's syndrome with a focus on its potential as a clinical trials biomarker. Ann Rheum Dis 2015;74:1645–50.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 949–960
DOI 10.1002/acr.23707

American College of Rheumatology
Empowering Rheumatology Professionals

The Incidence and Prevalence of Adult Primary Sjögren's Syndrome in New York County

Peter M. Izmirly,[1] Jill P. Buyon,[1] Isabella Wan,[1] H. Michael Belmont,[1] Sara Sahl,[2] Jane E. Salmon,[3] Anca Askanase,[4] Joan M. Bathon,[4] Laura Geraldino-Pardilla,[4] Yousaf Ali,[5] Ellen M. Ginzler,[6] Chaim Putterman,[7] Caroline Gordon,[8] Charles G. Helmick,[9] and Hilary Parton[10]

Objective. Extant epidemiologic data of primary Sjögren's syndrome (SS) remains limited, particularly for racial/ethnic populations in the US. The Manhattan Lupus Surveillance Program (MLSP) is a population-based retrospective registry of cases of systemic lupus erythematosus and related diseases, including primary SS in Manhattan, New York. The MLSP was used to provide estimates of the incidence and prevalence of primary SS across major racial/ethnic populations.

Methods. MLSP cases were identified from hospitals, rheumatologists, and population databases. Three case definitions were used for primary SS, including physician diagnosis, rheumatologist diagnosis, and modified primary SS criteria. Rates among Manhattan residents were age-adjusted, and capture-recapture analyses were conducted to assess underascertainment of cases.

Results. By physician diagnosis, age-adjusted overall incidence and prevalence rates of primary SS among adult Manhattan residents were 3.5 and 13.1 per 100,000 person-years, respectively. Capture-recapture adjustment increased incidence and prevalence rates (4.1 and 14.2 per 100,000 person-years, respectively). Based on physician diagnosis, incidence and prevalence rates were approximately 6 times higher among women than men ($P < 0.001$). Incidence of primary SS was statistically higher among non-Latina Asian women (10.5) and non-Latina white women (6.2) compared with Latina women (3.2). Incidence was also higher among non-Latina Asian women compared with non-Latina black women (3.3). Prevalence of primary SS did not differ by race/ethnicity. Similar trends were observed when more restrictive case definitions were applied.

Conclusion. Data from the MLSP revealed disparities among Manhattan residents in primary SS incidence and prevalence by sex and differences in primary SS incidence by race/ethnicity among women. These data also provided epidemiologic estimates for the major racial/ethnic populations in the US.

INTRODUCTION

Sjögren's syndrome (SS) is a chronic systemic autoimmune disease that manifests as oral and ocular dryness and parotid gland enlargement due to lymphocytic infiltration of exocrine glands, in addition to multiorgan-system extraglandular involvement (1). This syndrome can occur in the absence (referred to as primary) or presence (referred to as secondary) of other systemic rheumatologic or autoimmune diseases, such as systemic lupus erythematosus (SLE). The epidemiology of SS remains limited with few published estimates for the general population and minimal data on multiracial/ethnic populations in the US (2,3).

The Manhattan Lupus Surveillance Program (MLSP) was initiated in 2010 as a collaboration between the New York City

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Supported by cooperative agreements between the Centers for Disease Control and Prevention and The New York City Department of Health and Mental Hygiene (grant U58/DP002827) and the New York City Department of Health and Mental Hygiene and New York University School of Medicine.

[1]Peter M. Izmirly, MD, MSc, Jill P. Buyon, MD, Isabella Wan, MD, MPH, H. Michael Belmont, MD: New York University School of Medicine, New York, New York; [2]Sara Sahl, MD, MPH: Harbor-University of California Medical Center, Los Angeles; [3]Jane E. Salmon, MD: Hospital for Special Surgery, Weill Cornell Medical College, New York, New York; [4]Anca Askanase, MD, MPH, Joan M. Bathon, MD, Laura Geraldino-Pardilla, MD, MSc: Columbia University College of Physicians & Surgeons, New York, New York; [5]Yousaf Ali, MD: Icahn School of Medicine at Mount Sinai, New York, New York; [6]Ellen M. Ginzler, MD, MPH: State University of New York Downstate College of Medicine, Brooklyn; [7]Chaim Putterman, MD: Albert Einstein College of Medicine, Bronx, New York; [8]Caroline Gordon, MD: University of Birmingham, Birmingham, United Kingdom; [9]Charles G. Helmick, MD: Centers for Disease Control and Prevention, Atlanta, Georgia; [10]Hilary Parton, MPH: New York City Department of Health and Mental Hygiene, Long Island City, New York.

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

Address correspondence to Peter M. Izmirly, MD, MSc, NYU School of Medicine, 550 First Avenue, MSB 625, New York, NY 10016. E-mail: Peter.Izmirly@nyumc.org.

Submitted for publication April 9, 2018; accepted in revised form July 17, 2018.

SIGNIFICANCE & INNOVATIONS

- To our knowledge, this is the first population-based multiracial/ethnic study in the US to report on the epidemiology of Sjögren's syndrome where existing data are sparse in the literature.
- Our study revealed disparities in Sjögren's syndrome incidence and prevalence by sex among Manhattan residents and differences in incidence by race/ethnicity among women.
- These data also provided epidemiologic estimates for the major racial/ethnic populations in the US.

Department of Health and Mental Hygiene (DOHMH) and the New York University School of Medicine (NYUSoM) (4). The primary goal of the MLSP was to determine incidence and prevalence of SLE among Manhattan residents. To accomplish this, a retrospective population-based registry was established that comprises extensive information obtained on SLE as well as other autoimmune rheumatic diseases, including SS. Leveraging this rich data source, we provide incidence (between 2007 and 2009) and prevalence (during 2007) estimates of primary SS among Manhattan residents across the major racial/ethnic populations (black, Latino, Asian, white).

MATERIALS AND METHODS

MLSP. The MLSP is 1 of 5 registries funded by the Centers for Disease Control and Prevention (CDC) to provide credible estimates for the incidence and prevalence of SLE (4–8). Details on the MLSP have been previously reported (4). In brief, medical records were reviewed under the health surveillance exemption to HIPAA privacy rules [45 CFR § 164.512(b)] and as authorized by New York City Charter Sections 556(c)(2) and (d)(2). No cases were contacted for this project. The CDC deemed the various SLE surveillance programs public health practice, which did not require institutional review board (IRB) review, and IRBs at both the DOHMH and the NYUSoM considered the MLSP to be a surveillance activity. When requested, additional IRB applications were completed and submitted to independent case-finding sources. The DOHMH IRB reviewed and approved secondary analyses on a de-identified data set.

The MLSP surveillance period was from January 1, 2007 to December 31, 2009. Manhattan was selected for reasons previously described (4). In 2010, based on US census data, there were 1,585,873 persons residing in Manhattan (48% non-Latino white, 13% non-Latino black, 25% Latino, and 11% non-Latino Asian) (9).

Case ascertainment, data collection, and quality control of data entry. The MLSP used rheumatologist practices (including pediatric rheumatologists), hospitals, and administrative hospitalization discharge and death registry databases to identify cases (4). Case-finding sources were queried retrospectively, as far back as 2004 when available, for evidence of residence in Manhattan and International Classification of Disease Ninth Revision Clinical Modification (ICD-9CM) billing codes specific for SLE, discoid lupus, and related conditions that may evolve into SLE or have related symptoms, including SS. The ICD-9CM codes used to identify cases included 710.0 (SLE), 695.4 (discoid lupus), 710.8 (other specified connective tissue disease), 710.9 (unspecified connective tissue disease), and 710.2 (Sicca syndrome, which is used for SS). Charts for every patient who lived in Manhattan and had one of the respective ICD-9CM codes were fully abstracted and final diagnosis was coded. Abstraction was completed in 90.5% of hospitals and 75.8% of rheumatologist practices by trained abstractors, all of whom had medical degrees and underwent extensive training and routine quality assurance as previously described (4).

Case definitions. The MLSP was constructed for surveillance of SLE, and data elements that were collected focused on 2 widely used classification schemes for SLE, including the American College of Rheumatology (ACR) (10,11) and the Systemic Lupus International Collaborating Clinics (SLICC) (12) criteria for SLE. Additional manifestations commonly associated with SLE (even if not specifically included as a criterion for classification) were also captured, allowing for the potential to identify evidence of SS. Given the overlapping nature of the clinical manifestations of these autoimmune diseases, several but not all of the American-European Consensus Group (AECG) (13) criteria for SS (most recent available criteria at the time of data dictionary development) were captured. For criteria regarding the diagnosis of SS not systematically captured, abstractors were trained to take detailed text notes, including results of minor salivary gland biopsies and objective results of ocular and oral tests.

Because this analysis focused on primary SS, we excluded cases diagnosed with other rheumatologic diseases such as SLE, despite having an additional diagnosis of SS. Also, given the rarity of childhood primary SS, we only included cases ages ≥18 years in our analyses (14).

The diagnosis of SS is usually made by a physician familiar with the disease, which is often, but not exclusively, a rheumatologist. Thus, our primary case definition for primary SS required documentation of a primary SS diagnosis by any physician, and our more conservative secondary case definition required documentation of a primary SS diagnosis by a rheumatologist. In the MLSP, few cases met the AECG (13), ACR (15), and the more recent ACR/European League Against Rheumatism (EULAR) criteria for primary SS (16). Thus, we developed a third, more restrictive, case definition, slightly modified from the recent ACR/EULAR criteria (16), requiring

Table 1. Incidence rates of primary SS among Manhattan residents ages 18 years and older, 2007–2009, overall and by race/ethnicity and sex*

	Crude rate (95% CI)	Age-adjusted rate (95% CI)	*P* (by χ^2)	No. of missed cases†	Rate (95% CI)†
Physician diagnosis					
Total	3.4 (2.9,4.0)	3.5 (2.9,4.1)	<0.001	28.7	4.1 (2.5,5.8)
Male	0.9 (0.5,1.5)	1.0 (0.6,1.5)		4.7	1.2 (0.2,2.1)
Female	5.6 (4.6,6.6)	5.7 (4.7,6.7)		24.0	6.7 (4.3,9.1)
Race/ethnicity			<0.001§		
Non-Latino white	3.7 (3.0,4.7)	3.8 (3.0,4.7)		21.2	4.8 (2.2,7.4)
Non-Latino black	2.2 (1.1,3.9)	2.2 (1.1,4.0)		1.1	2.4 (1.4,3.4)
Latino	1.9 (1.1,3.0)	2.0 (1.2,3.2)		1.7	2.1 (1.7,2.5)
Non-Latino Asian	5.6 (3.6,8.1)	6.2 (4.0,9.2)		4.3	6.5 (5.2,7.8)
Non-Latino other‡	–	–		0.4	–
Race/ethnicity by sex					
Male			0.859		
Non-Latino white	1.1 (0.6,2.0)	1.1 (0.5,2.0)		4.3	1.6 (0.1,3.1)
Non-Latino black	0.9 (0.1,3.2)	0.9 (0.1,3.2)		0.2	1.0 (0.3,1.7)
Latino	0.5 (0.1,1.7)	0.5 (0.1,1.7)		0.2	0.5 (0.1,0.9)
Non-Latino Asian	0.5 (0.0,2.7)	0.6 (0.0,3.6)		0.0	0.5 (0.4,0.6)
Female			<0.001¶		
Non-Latina white	6.1 (4.7,7.8)	6.2 (4.7,7.9)		16.9	7.7 (4.1,11.3)
Non-Latina black	3.2 (1.5,6.1)	3.3 (1.5,6.3)		0.9	3.5 (2.3,4.8)
Latina	3.2 (1.8,5.1)	3.2 (1.8,5.2)		1.5	3.5 (3.1,3.8)
Non-Latina Asian	9.5 (6.2,14.0)	10.5 (6.6,15.7)		4.3	11.1 (8.9,13.4)
Rheumatologist diagnosis					
Total	2.1 (1.7,2.6)	2.1 (1.7,2.6)	<0.001	34.0	2.9 (1.1,4.8)
Male	0.5 (0.2,0.9)	0.5 (0.2,0.9)		–	–
Female	3.5 (2.7,4.4)	3.5 (2.7,4.4)		–	–
Race/ethnicity			0.001§		
Non-Latino white	2.4 (1.8,3.2)	2.3 (1.7,3.1)		13.2	3.0 (2.1,4.0)
Non-Latino black	1.2 (0.4,2.6)	1.2 (0.5,2.7)		9.0	2.9 (–1.8,7.7)
Latino	1.0 (0.4,1.8)	0.9 (0.4,1.8)		0.6	1.0 (0.6,1.5)
Non-Latino Asian	3.8 (2.3,6.1)	4.1 (2.3,6.5)		10.7	6.1 (1.0,11.2)
Non-Latino other‡	–	–		0.5	–
Race/ethnicity, male			0.524		
Non-Latino white	0.6 (0.2,1.3)	0.5 (0.2,1.2)		–	–
Non-Latino black	0.9 (0.1,3.2)	0.9 (0.1,3.2)		–	–
Latino	–	–		–	–
Non-Latino Asian	0.5 (0.0,2.7)	0.6 (0.0,3.6)		–	–
Race/ethnicity, female			0.001#		
Non-Latina white	4.0 (2.9,5.4)	3.8 (2.8,5.2)		–	–
Non-Latina black	1.4 (0.4,3.6)	1.6 (0.4,4.1)		–	–
Latina	1.8 (0.8,3.4)	1.7 (0.8,3.3)		–	–
Non-Latina Asian	6.5 (3.8,10.3)	6.7 (3.8,11.1)		–	–

* Rates are per 100,000 Manhattan residents. Denominator data is based on 2007–2009 intercensal population estimates from the New York City Department of Health and Mental Hygiene, Bureau of Epi Services (2000–2014 files). Data are age adjusted to the US 2000 standard population. Cases were assigned to 1 of 5 mutually exclusive race/ethnicity categories. SS = Sjögren's syndrome; 95% CI = 95% confidence interval.
† Values are those yielded from the capture-recapture analyses; log-linear models were fit separately for by sex and race/ethnicity for physician-diagnosed cases and by race/ethnicity alone for cases diagnosed by a rheumatologist.
‡ Non-Latino cases identifying with more than one race were categorized as non-Latino other.
§ Latino cases differed from non-Latino white and non-Latino Asian cases. Non-Latino Asian cases also differed from non-Latino black cases.
¶ Latina cases differed from non-Latina white and non-Latina Asian cases. Non-Latina Asian cases also differed from non-Latina black cases.
# Latina cases differed from non-Latina white and non-Latina Asian cases. Non-Latina black cases also differed from non-Latina white and non-Latina Asian cases.

3 criteria to be met: primary SS diagnosis by any physician, documentation of dry eyes and/or dry mouth, and a positive test for anti-SSA antibody.

Statistical analysis. Incident cases were those who were ages ≥18 years, met a primary SS case definition, resided in Manhattan, and were first diagnosed with primary SS between January 1, 2007 and December 31, 2009. Prevalent cases were new or existing cases among those ages ≥18 years that met a primary SS case definition and resided in Manhattan between January 1 and December 31, 2007. DOHMH intercensal population estimates for Manhattan were used to calculate denominators (9).

Rates were calculated overall, by sex, and by race/ethnicity per 100,000 person-years and were age-adjusted to the US 2000 standard population using 10-year age groups within each racial/ethnic group (17). Although data on race and Latino ethnicity were collected separately during abstraction, cases were assigned to 1 of 5 mutually exclusive race/ethnicity categories, including non-Latino white, non-Latino black, non-Latino Asian, Latino, and non-Latino other (including non-Latino cases identified with >1 race). Chi-square or Fisher's exact tests were used to determine if age-adjusted primary SS proportions differed by sex and race/ethnicity. When significant differences were found by race/ethnicity, pairwise differences were evaluated using Z tests assuming a Poisson distribution and statistical significance at 0.05, with Bonferroni correction ($P = 0.008$).

Capture-recapture analyses were performed (18,19) in order to estimate underascertainment of cases; specific methods have been described elsewhere (4). Log-linear models were fit separately for incident and prevalent cases by sex and race/ethnicity for physician-diagnosed cases and by race/ethnicity alone for cases diagnosed by a rheumatologist or meeting the modified case definition due to small numbers.

All analyses were completed using SAS software version 9.4, and R version 3.3.0 (R Foundation for Statistical Computing).

RESULTS

Incidence rates. From 2007 to 2009, 138 incident cases had a physician diagnosis of primary SS and 84 had a rheumatologist diagnosis of primary SS. The overall crude and age-adjusted incidence rates for physician-diagnosed cases of primary SS were 3.4 (95% confidence interval [95% CI] 2.9,4.0) and 3.5 (95% CI 2.9,4.1) per 100,000 person-years, respectively (Table 1). The overall crude and age-adjusted incidence rates for rheumatologist-diagnosed cases of primary SS were 2.1 (95% CI 1.7,2.6) and 2.1 (95% CI 1.7,2.6) per 100,000 person-years, respectively. Age-adjusted rates differed by sex, and were approximately 6 to 7 times higher for women compared with men for both physician- and rheumatologist-diagnosed primary SS (both $P < 0.001$). The incidence of physician-diagnosed primary SS differed by race/ethnicity ($P < 0.001$), with higher rates among Asian (6.2 per 100,000 person-years; $P = 0.002$) and white (3.8; $P = 0.006$) cases compared with Latino cases (2.0). Incidence of physician-diagnosed primary SS was also higher among Asian cases compared with black cases (2.2 per 100,000 person-years; $P = 0.005$). Similarly, incidence rates also differed by race/ethnicity among women ($P < 0.001$) and were higher among Asian (10.5; $P = 0.002$) and white women (6.2; $P = 0.007$) compared with Latina women (3.2), and among Asian women compared with black women (3.3; $P = 0.003$). There was no significant difference in age-adjusted incidence of physician-diagnosed primary SS by race/ethnicity among men ($P = 0.859$).

Incidence of rheumatologist-diagnosed primary SS also differed by race/ethnicity overall and among women (both $P = 0.001$), with higher rates among Asian women compared with Latina ($P = 0.007$) and black women ($P = 0.006$). Capture-recapture adjustment estimated 166.7 incident cases of physician-diagnosed primary SS, indicating that 17.2% of cases were missed. Among those missed, 58.9% were white women. The resulting capture-recapture adjusted incidence rate increased to 4.1 per 100,000 person-years (95% CI 2.5,5.8).

The mean ± SD age at diagnosis among incident cases was 56.0 ± 19.1 years among Latino cases, 54.7 ± 18.5 years among white cases, 48.6 ± 12.0 years among black cases, and 47.4 ± 18.1 years among Asian cases.

Among Latino primary SS cases, 77.8% of those diagnosed by a physician and 77.8% of those diagnosed by a rheumatologist were also classified as white. Ethnicity information among Latino cases was often absent, with two-thirds having no further information available. For those with more detail, ethnicities included Central or South American, Dominican, Puerto Rican, and Spanish. Among Asian primary SS incident cases diagnosed by a physician or rheumatologist, approximately one-third and more than one-quarter had no further classification for Asian ethnicity, respectively. Among cases with information available, ethnicities included Chinese, Indian or Pakistani, Japanese, and Thai.

The serologic and clinical manifestations of primary SS captured in the MLSP for incident cases of physician- and rheumatologist-diagnosed primary SS are shown in Table 2. Data ascertainment was more complete for cases with a rheumatologist diagnosis. Antinuclear antibodies and anti-SSA/Ro were the most commonly found serologic manifestations among both physician- and rheumatologist-diagnosed cases. Extraglandular manifestations were present in 62.6% and 65.4% of physician- and rheumatologist-diagnosed cases, respectively, with lymphopenia and arthritis being the most common.

Prevalence rates for primary SS. In 2007, a total of 166 cases had a physician diagnosis of primary SS and 94

had a rheumatologist diagnosis of primary SS. The crude and age-adjusted prevalence rates of physician-diagnosed primary SS overall were 12.4 (95% CI 10.5,14.3) and 13.1 (95% CI 11.1,15.1) per 100,000 person-years, respectively (Table 3). The overall crude and age-adjusted prevalence rates of rheumatologist-diagnosed primary SS were lower, at 7.0 (95% CI 5.7,8.6) and 7.3 (95% CI 5.9,8.9) per 100,000 person-years. Age-adjusted rates were approximately 6 times higher among women compared with men for both physician- and rheumatologist-diagnosed primary SS (both $P < 0.001$). Trends in both physician- and rheumatologist-diagnosed primary SS were similar to incidence. The age-adjusted prevalence rates of physician-diagnosed primary SS were 23.8 among white women, 23.7 among Asian women, 16.1 among black women, and 15.0 among Latina women. For rheumatologist-diagnosed primary SS, Asian women had the highest rate, followed by white, black, and Latina women. However, there were no significant differences in physician- or rheumatologist-diagnosed prevalence rates by race/ethnicity overall, among women or men.

Capture-recapture analyses estimated an additional 24.2 cases of physician-diagnosed primary SS, indicating that 12.7% of cases may have been missed. Among cases missed, almost two-thirds (65.3%) were white women. With capture-recapture adjustment, the overall prevalence by physician diagnosis increased to 14.2 per 100,000 person-years (95% CI 12.3,16.1).

The mean ± SD ages among women and men with primary SS identified by physician diagnosis were 56.4 ± 17.5 years and 60.9 ± 16.9 years, respectively. In 2007, the mean ± SD age of physician-diagnosed primary SS was 60.3 ± 16.7 years among white cases, 53.1 ± 20.5 years among black cases, 52.9 ± 14.8 years among Latino cases, and 49.6 ± 19.6 years among Asian cases.

Among physician- and rheumatologist-diagnosed prevalent Latino cases, more than three-fourths were also identified as white. Information on Latino ethnicity was often absent, with no

Table 2. Frequency of specific manifestations among incident primary SS cases among Manhattan residents ages 18 years and older, 2007–2009*

	Primary SS w/ physician diagnosis (n = 138)		Primary SS w/ rheumatologist diagnosis (n = 84)		Primary SS–modified definition (n = 45)†	
	No. available	Positive	No. available	Positive	No. available	Positive
Glandular/serologies						
Sicca symptoms	122	91 (74.6)	82	72 (87.8)	45	45 (100.0)
Anti-SSA/Ro	96	63 (65.6)	78	52 (66.7)	45	45 (100.0)
Anti-SSB/La	90	37 (41.1)	75	30 (40.0)	42	23 (55.0)
Anti-SSA/Ro and Anti-SSB/La	90	33 (36.7)	75	26 (34.7)	42	23 (55.0)
ANA	90	71 (78.9)	72	61 (84.7)	39	37 (95.0)
ANA titer >1:320	45	30 (66.7)	39	25 (64.1)	25	16 (64.0)
Rheumatoid factor	67	27 (40.3)	53	22 (41.5)	31	17 (55.0)
Extraglandular						
Arthritis	137	29 (21.2)	83	19 (22.9)	44	10 (23.0)
Photo sensitivity	138	7 (5.1)	84	7 (8.3)	45	4 (9.0)
Lymphopenia	124	64 (51.6)	79	39 (49.4)	42	22 (52.0)
ILD	138	4 (2.9)	84	2 (2.4)	45	2 (4.0)
Pneumonitis	138	1 (0.7)	84	1 (1.2)	45	1 (2.0)
Transverse myelitis	138	0 (0.0)	84	0 (0.0)	45	0 (0.0)
Low complements	138	5 (3.6)	84	5 (6.0)	45	5 (11.0)
Raynaud's	138	13 (9.4)	84	9 (10.7)	45	3 (7.0)
Cutaneous vasculitis	138	0 (0.0)	84	0 (0.0)	45	0 (0.0)
Cranial or peripheral neuropathy	137	10 (7.3)	84	5 (6.0)	45	2 (4.0)
Myositis	137	0 (0.0)	83	0 (0.0)	44	0 (0.0)

* Values are the number (%) of cases unless indicated otherwise. SS = Sjögren's syndrome; ANA = antinuclear antibodies; ILD = interstitial lung disease.

† Case definition, slightly modified from the recent American College of Rheumatology/European League Against Rheumatism criteria, required documentation including primary SS diagnosis by any physician, documentation of dry eyes and/or dry mouth, and a positive test for anti-SSA antibody.

Table 3. Prevalence rates of primary SS among Manhattan residents ages 18 years and older, 2007, overall and by race/ethnicity and sex*

				Capture-recapture†	
	Crude rate (95% CI)	Age-adjusted rate (95% CI)	*P* (*by* χ^2)	No. of cases missed	Rate (95% CI)
Physician diagnosis					
Total	12.4 (10.5,14.3)	13.1 (11.1,15.1)	<0.001	24.2	14.2 (12.3,16.1)
Male	3.1 (1.9,4.8)	3.5 (2.1,5.5)		0.1	3.1 (3.0,3.2)
Female	20.5 (17.2,23.8)	21.1 (17.6,24.5)		24.1	23.9 (20.3,27.4)
Race/ethnicity			0.099		
Non-Latino white	13.9 (11.3,17.0)	14.6 (11.8,17.9)		15.9	16.3 (14.4,18.1)
Non-Latino black	9.4 (5.4,15.2)	9.4 (5.4,15.4)		0.5	9.6 (8.4,10.9)
Latino	8.6 (5.7,12.6)	9.1 (6.0,13.2)		0.6	8.8 (8.1,9.5)
Non-Latino Asian	13.1 (8.0,20.2)	14.3 (8.5,22.5)		6.1	17.1 (11.6,22.6)
Non-Latino other‡	–	–		1.1	–
Race/ethnicity, male			0.638		
Non-Latino white	4.0 (2.1,6.9)	4.3 (2.3,7.4)		0.1	4.0 (3.8,4.3)
Non-Latino black	1.3 (0.0,7.3)	1.7 (0.0,9.4)		0.0	1.3 (1.3,1.3)
Latino	1.4 (0.2,5.0)	1.5 (0.2,5.4)		0.0	1.4 (1.4,1.4)
Non-Latino Asian	1.5 (0.0,8.3)	2.2 (0.1,12.5)		0.0	1.5 (1.5,1.5)
Race/ethnicity, female			0.153		
Non-Latina white	22.9 (18.2,28.4)	23.8 (18.9,29.6)		15.8	27.3 (24.0,30.7)
Non-Latina black	15.8 (8.9,26.1)	16.1 (8.9,26.7)		0.5	16.3 (14.1,18.6)
Latina	14.9 (9.6,21.9)	15.0 (9.7,22.2)		0.6	15.2 (13.9,16.5)
Non-Latina Asian	22.2 (13.4,34.6)	23.7 (13.9,37.7)		6.1	29.3 (19.5,39.1)
Rheumatologist diagnosis					
Total	7.0 (5.7,8.6)	7.3 (5.9,8.9)	<0.001	27.3	9.1 (6.2,11.9)
Male	1.6 (0.8,3.0)	1.8 (0.9,3.4)		–	–
Female	11.7 (9.3,14.5)	11.9 (9.5,14.8)		–	–
Race/ethnicity			0.399		
Non-Latino white	7.2 (5.3,9.5)	7.5 (5.5,10.0)		17.5	9.7 (7.5,12.0)
Non-Latino black	5.3 (2.4,10.0)	5.6 (2.5,10.6)		1.3	6.0 (4.2,7.9)
Latino	5.4 (3.2,8.7)	5.6 (3.2,9.0)		2.0	6.1 (5.0,7.2)
Non-Latino Asian	9.2 (5.0,15.4)	9.7 (5.1,16.6)		3.6	11.5 (5.5,17.5)
Non-Latino other‡	–	–		2.9	–
Race/ethnicity, male			0.703		
Non-Latino white	2.2 (0.9,4.5)	2.4 (0.9,4.9)		–	–
Non-Latino black	-			–	–
Latino	0.7 (0.0,3.9)	0.8 (0.0,4.3)		–	–
Non-Latino Asian	1.5 (0.0,8.3)	2.2 (0.1,12.5)		–	–
Race/ethnicity, female			0.490		
Non-Latina white	11.7 (8.5,15.9)	12.1 (8.7,16.5)		–	–
Non-Latina black	9.5 (4.3,18.0)	10.4 (4.7,19.8)		–	–
Latina	9.5 (5.4,15.4)	9.4 (5.4,15.3)		–	–
Non-Latina Asian	15.2 (8.1,25.9)	15.6 (8.0,27.3)		–	–

* Rates are per 100,000 Manhattan residents. Denominator data is based on 2007–2009 intercensal population estimates from the New York City Department of Health and Mental Hygiene, Bureau of Epi Services (2000–2014 files). Data are age adjusted to the US 2000 standard population. Cases were assigned to 1 of 5 mutually exclusive race/ethnicity categories. SS = Sjögren's syndrome; 95% CI = 95% confidence interval.

† Values are those yielded from the capture-recapture analyses, log-linear models were fit separately for by sex and race/ethnicity for physician-diagnosed cases and by race/ethnicity alone for cases diagnosed by a rheumatologist.

‡ Non-Latino cases identifying with more than one race were categorized as non-Latino other.

further details available for more than two-thirds of the cases. Among Asian primary SS cases diagnosed by a physician or rheumatologist, more than one-fourth had no further classification for Asian ethnicity.

The occurrence of relevant serologic and clinical manifestations captured in the MLSP for prevalent physician- and rheumatologist-diagnosed primary SS cases is shown in Table 4. Similar to incident cases, data ascertainment on manifestations was more complete for cases with a rheumatologist diagnosis.

Incidence and prevalence of primary SS using modified criteria. Using the modified case definition of primary SS (Table 5), incorporating the presence of autoantibodies and documentation of dry eyes and/or dry mouth resulted in an overall age-adjusted incidence rate of 1.1 (95% CI 0.8,1.5) per 100,000 person-years and an overall age-adjusted prevalence rate of 3.3 (95% CI 2.4,4.4) per 100,000 person-years. As with the other case definitions, age-adjusted rates were higher among women compared with men ($P < 0.001$). Incidence rates differed by race/ethnicity overall ($P < 0.001$), with higher rates among Asian cases compared with white ($P = 0.007$) and Latino cases ($P = 0.005$), and among women ($P < 0.001$), with higher rates among Asian cases compared with black cases ($P = 0.003$). Prevalence of primary SS differed by race/ethnicity overall ($P = 0.001$) and among women ($P = 0.001$), but no significant differences were found by pairwise comparison.

Incident and prevalent cases of primary SS meeting criteria for SLE. Cases with a diagnosis of primary SS also met ≥4 of the ACR and/or SLICC criteria for SLE despite not being clinically diagnosed with SLE (Table 6). Depending on the case definition for primary SS, 4.3–10.2% of incident cases met the ACR criteria for SLE and 5.8–16.3% met the SLICC criteria. The modified case definition for primary SS had the highest percentage of incident cases meeting ACR and SLICC criteria for SLE. There was a higher percentage of prevalent primary SS cases meeting SLE criteria (6.6–14.9%, for ACR criteria; 14.5–34.0%, for SLICC criteria).

DISCUSSION

Our analysis of the MLSP data set provides incidence and prevalence rate estimates of primary SS among Manhattan residents. These data also provided epidemiologic estimates for the major racial/ethnic populations in the US. The age-standardized incidence and prevalence of physician-diagnosed primary SS in Manhattan were 3.5 (95% CI 2.9,4.1) and 13.1 (95% CI 11.1,15.1) per 100,000 person-years, respectively. Capture-recapture adjustment increased incidence rates by 17.2% and prevalence rates by 12.7%. By rheumatologist diagnosis, the age-adjusted incidence rate of primary SS was 2.1 (95% CI 1.7,2.6) and the prevalence rate was 7.3 (95% CI 5.9,8.9). Incidence was highest among Asian and white cases, though prevalence did not significantly differ by race/ethnicity, and there were substantial disparities in the prevalence and incidence of primary SS among Manhattan residents by sex. The current analysis also provides information on serologic and clinical manifestations among primary SS cases, including data on extraglandular manifestations. Additionally, these data reveal that up to one-third of prevalent cases diagnosed with primary SS also fulfill both ACR and SLICC criteria for SLE, even though they do not carry a diagnosis of SLE, reflecting commonalities in manifestations of the 2 diseases. Not surprisingly, these data suggest that in clinical practice physicians diagnose patients without formal application of disease criteria.

Previous studies on the epidemiology of SS span decades, come from different regions of the world, and have used varying methods of case identification (2,3). The few published estimates for the general population reveal annual incidence rates of 6.9–20.1 per 100,000 persons and markedly discrepant prevalence figures ranging from 11.3 to 3790.1 cases per 100,000 persons (3). Whether these estimates reflect genuine variability between different populations or differences in methodology and study design is unclear.

Existing data suggest that the disease is most common in middle-aged women, which is consistent with findings of our analysis (2,3). In line with these findings, our analyses were restricted to adults ages ≥18 years, though it is worth noting that the MLSP did identify pediatric cases of primary SS. However, including these pediatric cases into our prevalence and incidence estimates of physician-diagnosed primary SS would have decreased our estimates by at least 21%, given the small number of cases added to our numerator relative to the person-years added to our denominator.

In a meta-analysis of primary SS studies published to date (3), 21 population-based studies were identified, of which only 10 included a review of medical records; the rest were population-based surveys. Six studies (20–25) determined an incidence rate, only 1 of which was US-based (21); the authors calculated a pooled primary SS prevalence rate of 60.8 per 100,000 person-years and an incidence rate of 6.9 per 100,000, both of which were higher than our estimates. However, in line with the findings from our study, the authors found a higher pooled primary SS incidence rate among women compared with men (12.3 versus 1.5). A report limited to European-based studies using the AECG criteria for primary SS showed a European prevalence rate of primary SS at 38.95 per 100,000 population (26), and after being updated as a meta-analysis showed a point prevalence of 4.7 per 10,000 population (27).

A recent US-based study of primary SS that was conducted in Olmsted County, Minnesota (28), with a mostly white population, reported a population-based prevalence estimate for primary SS based on physician diagnosis of 10.3 per 10,000 residents, which is also higher than our estimate. Even using the conserva-

Table 4. Frequency of specific manifestations among prevalent primary SS cases among NYC Manhattan residents ages 18 and older, 2007*

	Primary SS w/physician diagnosis (n = 166)		Primary SS w/rheumatologist diagnosis (n = 94)		Primary SS–modified definition (n = 47)†	
	No. available	Positive	No. available	Positive	No. available	Positive
Glandular/serologies						
Sicca symptoms	152	110 (72.4)	91	81 (89.0)	44	44 (100.0)
Anti-SSA/Ro	102	60 (58.8)	77	48 (62.3)	44	44 (100.0)
Anti-SSB/La	100	47 (47.0)	77	37 (48.1)	44	32 (73.0)
Anti-SSA/Ro and anti-SSB/La	100	44 (44.0)	77	34 (44.2)	44	32 (73.0)
ANA	106	72 (67.9)	81	58 (71.6)	44	38 (86.0)
ANA titer >1:320	56	32 (57.1)	46	27 (58.7)	31	21 (68.0)
Rheumatoid factor	82	42 (51.2)	64	36 (56.3)	35	26 (74.0)
Extraglandular						
Arthritis	166	37 (22.3)	94	26 (27.7)	44	16 (36.0)
Photo sensitivity	166	5 (3.0)	94	5 (5.3)	44	3 (7.0)
Lymphopenia	149	103 (69.1)	87	61 (70.1)	44	35 (80.0)
ILD	166	10 (6.0)	94	5 (5.3)	44	3 (7.0)
Pneumonitis	166	3 (1.8)	94	1 (1.1)	44	1 (2.0)
Transverse myelitis	166	0 (0.0)	94	0 (0.0)	44	0 (0.0)
Low complements	166	10 (6.0)	94	10 (10.6)	44	7 (16.0)
Raynaud's	166	16 (9.6)	94	14 (14.9)	44	5 (11.0)
Cutaneous vasculitis	166	6 (3.6)	94	3 (3.2)	44	2 (5.0)
Cranial or peripheral neuropathy	165	17 (10.3)	94	12 (12.8)	44	5 (11.0)
Myositis	166	2 (1.2)	94	1 (1.1)	44	1 (2.0)

* Values are the number (%) of cases unless indicated otherwise. SS = Sjögren's syndrome; NYC = New York City; ANA = antinuclear antibodies; ILD = interstitial lung disease.
† Case definition, slightly modified from the recent American College of Rheumatology/European League Against Rheumatism criteria, required documentation including primary SS diagnosis by any physician, documentation of dry eyes and/or dry mouth, and a positive test for anti-SSA antibody.

tive AECG definition, the prevalence estimate was still higher than ours at 2.2 (95% CI 1.3,3.1) per 10,000 (28). A separate study of the same population also showed a higher annual incidence rate of physician-diagnosed primary SS at 5.9 per 100,000 population (95% CI 4.4,7.4) (29).

There are virtually no studies that present primary SS findings among diverse populations. The aforementioned meta-analysis by Qin et al (3) presented no information on race or ethnicity other than a few studies done in Taiwan (23–25), which found a pooled incidence rate of 6.6 per 100,000 person-years with significant heterogeneity. One study conducted in the greater Paris area of France reported population-based estimates of primary SS prevalence among a multiracial/multiethnic population (2). In line with our findings, in this study prevalence estimates of primary SS in adults ages ≥15 years ranged from 10.0 per 100,000 to 15.2 per 100,000, depending on the definition used. Prevalence was approximately 2 times higher for non-Europeans, although incidence and further breakdown on non-European origin was not reported (2). Our study did not find significant differences in prevalence by race/ethnicity but also had a different racial/ethnic makeup. We did, however, find significant differences in incidence by race/ethnicity.

There were several limitations regarding the development of the MLSP, which have been previously acknowledged (4). These analyses may have underestimated incident and prevalent cases, because 2 hospitals and one-fourth of rheumatologists in the catchment area, who practiced in predominantly white neighborhoods, declined to participate. The Veteran's Administration Hospital was one of the hospitals that declined to participate (the other was a cancer specialty hospital), which may have caused underidentification specifically of males diagnosed with primary SS in the present analysis. It is also possible that cases were missed if they lived in Manhattan but sought care in other boroughs or a neighboring state. We also did not include ophthalmologists, otolaryngologists, or primary care practices among our case-finding sources.

As previously detailed (4), additional limitations of the MLSP resulted from the tremendous differences across medical sys-

Table 5. Rates of primary SS by modified definition among Manhattan residents ages 18 and older, overall and by race/ethnicity and sex*

				Capture-Recapture†	
	Crude rate (95% CI)	Age-adjusted rate (95% CI)	*P* (by χ^2)	No. of missed cases	Rate (95% CI)
Incidence, 2007–2009					
Total	1.1 (0.8,1.5)	1.1 (0.8,1.5)	<0.001	14.1	1.5 (0.3,2.6)
Male	0.2 (0.1,0.6)	0.2 (0.1,0.5)		–	–
Female	1.9 (1.4,2.6)	1.8 (1.3,2.5)		–	–
Race/ethnicity			<0.001§		
Non-Latino white	0.9 (0.6,1.4)	0.9 (0.5,1.3)		2.0	1.0 (0.6,1.4)
Non-Latino black	0.8 (0.2,2.0)	0.8 (0.2,2.1)		2.4	1.3 (–1.1,3.6)
Latino	0.7 (0.3,1.5)	0.7 (0.3,1.5)		1.2	0.9 (0.2,1.5)
Non-Latino Asian	3.0 (1.6,5.0)	3.3 (1.6,5.2)		8.2	4.7 (0.8,8.7)
Non-Latino other‡	–			0.3	–
Race/ethnicity, male			0.004		
Non-Latino white	0.1 (0.0,0.6)	0.1 (0.0,0.5)		–	–
Non-Latino black	0.9 (0.1,3.2)	0.9 (0.1,3.2)		–	–
Latino	–			–	–
Non-Latino Asian	0.5 (0.0,2.7)	0.6 (0.0,3.6)		–	–
Race/ethnicity, female			<0.001¶		
Non-Latina white	1.7 (1.0,2.6)	1.5 (0.9,2.3)		–	–
Non-Latina black	0.7 (0.1,2.6)	0.7 (0.1,2.7)		–	–
Latina	1.4 (0.6,2.8)	1.4 (0.5,2.8)		–	–
Non-Latina Asian	4.9 (2.6,8.5)	4.9 (2.5,8.6)		–	–
Prevalence, 2007					
Total	3.3 (2.4,4.4)	3.3 (2.4,4.4)	<0.001	18.6	4.7 (1.4,8.0)
Male	0.5 (0.1,1.4)	0.5 (0.1,1.4)		–	–
Female	5.7 (4.1,7.8)	5.7 (4.1,7.8)		–	–
Race/ethnicity			0.001		
Non-Latino white	2.8 (1.7,4.4)	2.8 (1.6,4.3)		12.3	4.6 (0.6,8.6)
Non-Latino black	3.5 (1.3,7.6)	3.6 (1.3,7.8)		0.8	4.0 (2.6,5.3)
Latino	3.2 (1.5,5.9)	3.1 (1.5,5.7)		2.5	4.0 (2.4,5.6)
Non-Latino Asian	5.2 (2.3,10.3)	5.2 (2.1,10.5)		2.1	6.6 (3.0,10.2)
Non-Latino other‡	–			0.9	–
Race/ethnicity, male					
Non-Latino white	0.9 (0.2,2.7)	0.9 (0.2,2.8)		–	–
Non-Latino black	–	–		–	–
Latino	–	–		–	–
Non-Latino Asian	–	–		–	–
Race/ethnicity, female			0.001		
Non-Latina white	4.5 (2.6,7.3)	4.8 (2.5,7.2)		–	–
Non-Latina black	6.3 (2.3,13.8)	6.6 (2.4,14.5)		–	–
Latina	5.9 (2.9,10.9)	5.8 (2.7,10.7)		–	–
Non-Latina Asian	9.3 (4.0,18.4)	9.4 (3.9,19.1)		–	–

* Rates are per 100,000 Manhattan residents. Denominator data is based on 2007-2009 intercensal population estimates from the New York City Department of Health and Mental Hygiene, Bureau of Epi Services (2000–2014 files). Data are age adjusted to the US 2000 standard population. Cases were assigned to 1 of 5 mutually exclusive race/ethnicity categories. SS = Sjögren's syndrome; 95% CI = 95% confidence interval.
† For capture-recapture analyses, log-linear models were fit by race/ethnicity alone for cases meeting the modified case definition (which required documentation including primary SS diagnosis by any physician, documentation of dry eyes and/or dry mouth, and a positive test for anti-SSA antibody).
‡ Non-Latino cases identifying with more than one race were categorized as non-Latino other.
§ Non-Latino Asian cases differed from non-Latino white and Latino cases.
¶ Non-Latina Asian cases differed from non-Latina black cases.

Table 6. Frequency of meeting ACR and SLICC classification criteria for SLE among primary SS cases*

	Physician diagnosis		Rheumatologist diagnosis		Modified definition†	
	No.	Positive (%)	No.	Positive (%)	No.	Positive (%)
Incident cases, 2007–2009						
Overall	138	–	84	–	49	–
Meet ACR SLE criteria	–	4.3	–	7.1	–	10.2
Meet SLICC SLE criteria	–	5.8	–	9.5	–	16.3
Prevalent cases, 2007						
Overall	166	–	94	–	47	–
Meet ACR SLE criteria	–	6.6	–	10.6	–	14.9
Meet SLICC SLE criteria	–	14.5	–	23.4	–	34.0

* ACR = American College of Rheumatology; SLICC = Systemic Lupus International Collaborating Clinics; SLE = systemic lupus erythematosus; SS = Sjögren's syndrome.
† Case definition, slightly modified from the recent American College of Rheumatology/European League Against Rheumatism criteria, which required documentation including primary SS diagnosis by any physician, documentation of dry eyes and/or dry mouth, and a positive test for anti-SSA antibody.

tems and abstracting several years after the surveillance period. These limitations could have resulted in abstractors missing information such as results of minor salivary gland biopsies and objective results of ocular and oral tests. This could account for our rates using the modified case definition (16) being considerably lower than those by physician and rheumatologist diagnosis.

Another explanation for these lower rates using the modified case definition comes from feedback obtained from our abstractors while in the field. Documentation of salivary gland biopsies or objective evidence of dry eyes (positive Schirmer's test, rose bengal test score, or other ocular dye score) and dry mouth (positive unstimulated whole salivary flow test, parotid gland sialography, or salivary gland scintigraphy) were rare, which limited our planned ability to use the various primary SS criteria (13,15,16). In addition, when biopsies were performed, they were not reported in any standardized way (30).

This observation was corroborated in the recent study by Maciel et al that explored the prevalence of SS in Olmsted County, MN, where the rates for AECG-confirmed SS were considerably lower than rates for physician-diagnosed SS (28). Maciel and colleagues concluded that classification criteria do not accurately reflect the diagnosis of SS in clinical practice, in part because the criteria include invasive tests that are rarely performed in routine care (28). Importantly, these criteria sets were not developed for diagnostic use in routine clinical practice, but were designed to capture a more homogeneous patient population for the purpose of research and clinical trials (31).

Additional limitations of our analysis pertain to assigning race and ethnicity based on administrative and medical records. Though available information did reflect the major ethnic subgroups in Manhattan, specific ethnicity information was missing for most Latino cases and more than one-fourth of Asian cases. Categorized broadly, Latino or Asian race encompasses a number of heterogeneous groups and primary SS rates among them may differ. Given the already limited number of published studies on primary SS among Asian and Latino individuals, additional work is needed to better describe and understand the epidemiology of primary SS among specific ethnic subpopulations.

Despite these limitations, our analysis benefitted from the design and composition of the MLSP (4). First, the MLSP was designed as a population-based registry with a diverse population, which allowed us to estimate rates of primary SS among the major racial/ethnic categories. The partnership with the DOHMH allowed us to collect information from a number of case-finding sources, which facilitated more complete clinical information on many cases. In addition, we conducted capture-recapture analyses to estimate missed cases. Finally, our abstractors all had medical backgrounds, which helped during training and provided an advantage in identifying criteria and manifestations of primary SS during extensive review of medical records.

In conclusion, data from a large population-based registry revealed substantial disparities by sex in primary SS among Manhattan residents. Differences were also found in the incidence of primary SS by race/ethnicity, highlighting higher rates among Asian women that have not been documented previously in the US.

ACKNOWLEDGMENTS

We thank Benjamin Wainwright for assistance in preparing the manuscript, all of the rheumatologists and their practice managers who participated in the Manhattan Lupus Surveillance Program (MLSP), and all of the administrators in the medical records departments of the hospitals who participated in the MLSP for their assistance in obtaining

medical records. At the New York City Department of Health and Mental Hygiene (DOHMH), the MLSP acknowledges the contributions of past and current members, including Tamira Collins-Bowers, Manasi Joshi, Bonnie Kerker, Maushumi Mavinkurve, Angela Merges, Kyyon Nelson, Viren Shah, Joseph Slade, Lorna Thorpe, Talytha Utley, and Elizabeth Waddell. In addition, the MLSP acknowledges the hard work of their abstractors, Drs. Janice McFarlane, Nick Stefanopoulos, Zahira Zahid, Rukayatu Ibrahim, Saleh Massasati, and Simone Shrestha. Finally, we acknowledge the support and contribution of the principal investigators of the other CDC-funded surveillance sites, including Drs. Sam Lim, Cristina Drenkard, Emily Somers, Joe McCune, Maria Dall'Era, and Elizabeth Ferucci.

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. Izmirly 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. Izmirly, Gordon, Helmick, Parton.
Acquisition of data. Izmirly, Buyon, Wan, Belmont, Sahl, Salmon, Askanase, Bathon, Geraldino-Pardilla, Ali, Ginzler, Putterman, Parton.
Analysis and interpretation of data. Izmirly, Buyon, Helmick, Parton.

REFERENCES

1. Brito-Zerón P, Baldini C, Bootsma H, Bowman SJ, Jonsson R, Mariette X, et al. Sjögren syndrome. Nat Rev Dis Primers 2016. E-pub ahead of print.
2. Maldini C, Seror R, Fain O, Dhote R, Amoura Z, De Bandt M, et al. Epidemiology of primary Sjögren's syndrome in a French multiracial/multiethnic area. Arthritis Care Res (Hoboken) 2014;66:454–63.
3. Qin B, Wang J, Yang Z, Yang M, Ma N, Huang F, et al. Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. Ann Rheum Dis 2015;74:1983–9.
4. Izmirly PM, Wan I, Sahl S, Buyon JP, Belmont HM, Salmon JE, et al. The incidence and prevalence of systemic lupus erythematosus in New York County (Manhattan), New York: the Manhattan Lupus Surveillance Program. Arthritis Rheumatol 2017;69:2006–17.
5. Lim SS, Bayakly AR, Helmick CG, Gordon C, Easley KA, Drenkard C. The incidence and prevalence of systemic lupus erythematosus, 2002–2004: the Georgia Lupus Registry. Arthritis Rheumatol 2014;66:357–68.
6. Somers EC, Marder W, Cagnoli P, Lewis EE, DeGuire P, Gordon C, et al. Population-based incidence and prevalence of systemic lupus erythematosus: the Michigan Lupus Epidemiology and Surveillance Program. Arthritis Rheumatol 2014;66:369–78.
7. Ferucci ED, Johnston JM, Gaddy JR, Sumner L, Posever JO, Choromanski TL, et al. Prevalence and incidence of systemic lupus erythematosus in a population-based registry of American Indian and Alaska Native people, 2007–2009. Arthritis Rheumatol 2014;66:2494–502.
8. Dall'Era M, Cisternas MG, Snipes K, Herrinton LJ, Gordon C, Helmick CG. The incidence and prevalence of systemic lupus erythematosus in San Francisco County, California: the California Lupus Surveillance Project. Arthritis Rheumatol 2017;69:1996–2005.
9. New York City Department of Health and Mental Hygiene. EpiQuery: NYC interactive health data system- NYCDOHMH neighborhood population estimates, modified from the US Census Bureau vintage population estimates, 2007, 2008, 2009. 2016. https://a816-healthpsi.nyc.gov/epiquery/.
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. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.
12. Petri M, Orbai AM, Alarcon GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum 2012;64:2677–86.
13. Vitali C, Bombardieri S, Jonsson R, Moutsopoulos HM, Alexander EL, Carsons SE, et al. Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. Ann Rheum Dis 2002;61:554–8.
14. Lieberman SM. Childhood Sjögren syndrome: insights from adults and animal models. Curr Opin Rheumatol 2013;25:651–7.
15. Shiboski SC, Shiboski CH, Criswell L, Baer A, Challacombe S, Lanfranchi H, et al, for the Sjögren's International Collaborative Clinical Alliance (SICCA) Research Groups. American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance cohort. Arthritis Care Res (Hoboken) 2012;64:475–87.
16. Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome: a consensus and data-driven methodology involving three international patient cohorts. Ann Rheum Dis 2017;76:9–16.
17. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People 2010 Stat Notes 2001:1–10.
18. Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. Epidemiol Rev 1995;17:243–64.
19. Baillargeon S, Rivest L. Rcapture: loglinear models for Capture-Recapture in R. J Stat Softw 2007;19:1–31.
20. Alamanos Y, Tsifetaki N, Voulgari PV, Venetsanopoulou AI, Siozos C, Drosos AA. Epidemiology of primary Sjögren's syndrome in north-west Greece, 1982–2003. Rheumatology (Oxford) 2006;45:187–91.
21. Pillemer SR, Matteson EL, Jacobsson LT, Martens PB, Melton LJ III, O'Fallon WM, et al. Incidence of physician-diagnosed primary Sjögren syndrome in residents of Olmsted County, Minnesota. Mayo Clin Proc 2001;76:593–9.
22. Plesivcnik Novljan M, Rozman B, Hocevar A, Grmek M, Kveder T, Tomsic M. Incidence of primary Sjögren's syndrome in Slovenia. Ann Rheum Dis 2004;63:874–6.
23. See LC, Kuo CF, Chou IJ, Chiou MJ, Yu KH. Sex- and age-specific incidence of autoimmune rheumatic diseases in the Chinese population: a Taiwan population-based study. Semin Arthritis Rheum 2013;43:381–6.
24. Weng MY, Huang YT, Liu MF, Lu TH. Incidence and mortality of treated primary Sjögren's syndrome in Taiwan: a population-based study. J Rheumatol 2011;38:706–8.
25. Yu KH, See LC, Kuo CF, Chou IJ, Chou MJ. Prevalence and incidence in patients with autoimmune rheumatic diseases: a nationwide population-based study in Taiwan. Arthritis Care Res (Hoboken) 2013;65:244–50.
26. Cornec D, Chiche L. Is primary Sjögren's syndrome an orphan disease? A critical appraisal of prevalence studies in Europe. Ann Rheum Dis 2015;74:e25.

27. Nocturne G, Seror R, Mariette X, Devauchelle-Pensec V, Saraux A, Chiche L. Primary Sjögren's syndrome prevalence: what if Sjögren was right after all? Comment on the article by Maciel et al. Arthritis Care Res (Hoboken) 2018;70:951–3.
28. Maciel G, Crowson CS, Matteson EL, Cornec D. Prevalence of primary Sjögren's syndrome in a US population-based cohort. Arthritis Care Res (Hoboken) 2017;69:1612–6.
29. Maciel G, Crowson CS, Matteson EL, Cornec D. Incidence and mortality of physician-diagnosed primary Sjögren syndrome: time trends over a 40-year period in a population-based US cohort. Mayo Clin Proc 2017;92:734–43.
30. Fisher BA, Jonsson R, Daniels T, Bombardieri M, Brown RM, Morgan P, et al. Standardisation of labial salivary gland histopathology in clinical trials in primary Sjögren's syndrome. Ann Rheum Dis 2017;76:1161–8.
31. Vitali C, Del Papa N. Classification and diagnostic criteria in Sjögren's syndrome: a long-standing and still open controversy. Ann Rheum Dis 2017;76:1953–4.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 961–969
DOI 10.1002/acr.23715

American College of Rheumatology
Empowering Rheumatology Professionals

Early Self-Reported Pain in Juvenile Idiopathic Arthritis as Related to Long-Term Outcomes: Results From the Nordic Juvenile Idiopathic Arthritis Cohort Study

Ellen Dalen Arnstad,[1] Veronika Rypdal,[2] Suvi Peltoniemi,[3] Troels Herlin,[4] Lillemor Berntson,[5] Anders Fasth,[6] Susan Nielsen,[7] Mia Glerup,[4] Maria Ekelund,[8] Marek Zak,[7] Kristiina Aalto,[3] Ellen Nordal,[2] Pål Richard Romundstad,[9] and Marite Rygg,[10] on behalf of the Nordic Study Group of Pediatric Rheumatology

Objective. To study self-reported pain early in the disease course of juvenile idiopathic arthritis (JIA) as a predictor of long-term disease outcomes.

Methods. Consecutive cases of JIA with disease onset from 1997 to 2000 from defined geographical areas of Norway, Sweden, Finland, and Denmark were prospectively enrolled in this population-based cohort study. Self-reported, disease-related pain was measured on a 10-cm visual analog scale (VAS pain). Inclusion criteria were a baseline visit with a pain score 6 months after disease onset, followed by an 8-year study visit. Remission was defined according to Wallace et al (2004) preliminary criteria. Functional disability was measured by the Childhood Health Assessment Questionnaire and the Child Health Questionnaire Parent Form if the child was age <18 years and by the Health Assessment Questionnaire if age ≥18 years. Damage was scored using the Juvenile Arthritis Damage Index.

Results. The final study cohort consisted of 243 participants, and 120 participants (49%) had oligoarticular onset. At baseline, 76% reported a VAS pain score >0 compared to 57% reporting at 8 years. Half of those who reported baseline pain also reported pain at 8 years but at a lower intensity. Compared to no pain, higher pain intensity at baseline predicted more pain at 8 years, more functional disability, more damage, and less remission without medication. Baseline pain predicted more use of disease-modifying antirheumatic drugs/biologics during the disease course. Participants with oligoarticular JIA reporting pain at baseline were more likely to develop extended oligoarticular JIA or other JIA categories with an unfavorable prognosis.

Conclusion. Early self-reported, disease-related pain among children and adolescents with JIA is common and seems to predict persistent pain and unfavorable long-term disease outcomes.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is a diverse chronic disease with onset at age <16 years. This most common rheumatic disease among children is characterized by at least 6 weeks of continuous arthritis of unknown cause in 1 or more joints (1). The incidence rate in the Nordic countries is reported to be approximately 15–22 per 100,000 children (2–4). JIA is a heterogeneous disorder classified into 7 categories, based on defined criteria occurring during the first 6 months after disease onset (5). Among the different categories, and within each category, the disease course and outcome differ markedly (6). Persistent oligoarticular JIA has the best prognosis of all JIA categories (7). Extended oligoarticular JIA has a more unfavorable outcome, similar to that of polyarticular disease (8,9). Predicting outcome is challenging, and several studies have focused on associations between long-term outcome and clinical characteristics and biomarkers, such as the nature of joint involvement, the intensity

[1]Ellen Dalen Arnstad, MD: Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, and Norwegian University of Science and Technology, Trondheim, Norway; [2]Veronika Rypdal, MD, Ellen Nordal, MD, PhD: University Hospital of North Norway and Arctic University of Norway, Tromsø, Norway; [3]Suvi Peltoniemi, MD, Kristiina Aalto, MD, PhD: Hospital for Children and Adolescents, University of Helsinki, Helsinki, Finland; [4]Troels Herlin, MD, PhD, Mia Glerup, MD: Aarhus University Hospital, Aarhus, Denmark; [5]Lillemor Berntson, MD, PhD: Uppsala University, Uppsala, Sweden; [6]Anders Fasth, MD, PhD: University of Gothenburg, Gothenburg, Sweden; [7]Susan Nielsen, MD, Marek Zak, MD, PhD: Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; [8]Maria Ekelund, MD: Uppsala University, Uppsala, and Ryhov County Hospital, Jonkoping, Sweden; [9]Pål Richard Romundstad, MSc, PhD: Norwegian University of Science and Technology, Trondheim, Norway; [10]Marite Rygg, MD, PhD: Norwegian University of Science and Technology and St. Olavs Hospital, Trondheim, Norway.

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

Address correspondence to Ellen Dalen Arnstad, MD, Department of Pediatrics, Levanger Hospital, Nord-Trøndelag Hospital Trust, Pb 333, 7601 Levanger, Norway. E-mail: ellen.d.arnstad@ntnu.no.

Submitted for publication March 24, 2018; accepted in revised form July 24, 2018.

SIGNIFICANCE & INNOVATIONS

- Pain is a frequent symptom, tends to persist, and affects health-related quality of life for children and adolescents with juvenile idiopathic arthritis (JIA).
- In this study, we showed for the first time that an early pain report is associated with long-term nonremission, functional impairment, more use of disease-modifying antirheumatic drugs/biologics, and, for those with oligoarticular JIA, development into extended disease.
- The study adds to the increasing amount of evidence establishing the importance of pain assessment in routine care of children and adolescents with JIA.

of acute-phase response, and the existence of autoantibodies and genetic variables (10). Because none of those predictors are perfect, and in order to tailor treatment to reach the target of clinical remission, there is a need for more prospective longitudinal studies to evaluate early predictors using validated and multidimensional measures (10). More patient-centered measurements have been needed for assessment of the course and outcome of JIA (11). Among the 6 core variables endorsed by the American College of Rheumatology (ACR) (12), only the parent/patient assessment of overall well-being can be defined as a patient-reported measure. Patients, parents, and clinicians have pointed to more specific quality-of-life measures, and especially pain, as important measures when evaluating the course and outcome of JIA (11).

Pain is a frequent symptom among children and adolescents with JIA (13,14). Pain perception is highly subjective, and different self-reported measures are used to detect pain frequency and intensity (15,16). Pain assessment in young children is especially challenging, because pain reports are dependent on the parents' assumption of their child's pain (17). Both unidimensional and multidimensional tools are available for parent and child/adolescent assessment of pain in JIA (16). Among the unidimensional tools, the visual analog scale (VAS) is a commonly used and validated scoring instrument (18,19). The pathogenesis of pain in children and adolescents with JIA is multifactorial, including both biologic and psychosocial factors (16,20). Pain is a distressing symptom, and several studies have elucidated the relationship between pain, functional disability, and health-related quality of life (21–24), but information about pain as a predictor of long-term disease outcome is lacking.

In our Nordic population-based JIA cohort with comprehensive and prospectively sampled data, we have previously studied different aspects of JIA (2,8). In this project, we aimed to study self-reported pain early in the disease course and the association with long-term disease outcomes.

PATIENTS AND METHODS

Patients. The Nordic JIA cohort is a population-based cohort study. Consecutive cases of newly diagnosed JIA from defined geographical areas of Norway, Sweden, Finland, and Denmark with disease onset from January 1, 1997 to June 30, 2000, were prospectively included. Disease onset was defined as the day the child fulfilled the criteria for active arthritis according to information given by the parents/patient or by a physician. Participants were included consecutively and as soon as possible after the diagnosis was determined. However, the first extensive baseline visit was scheduled for 6 months after disease onset. This time point was chosen to enable classification of the disease into a JIA category, according to the International League Against Rheumatism Edmonton criteria (5). We have no registration of onset symptoms in the database. A detailed description of data collection and patient enrollment has been published previously (2,8). In the current study, participants were included if they had at least a baseline visit 6 months after disease onset with available pain scores, and participation in the 8-year follow-up visit. At both visits, we had data from clinical examinations, disease activity measures, previous and ongoing medication, and damage and remission status, as well as results from blood tests. Health-related quality of life was reported by the children or by their parents.

Measures. Self-reported, disease-related pain intensity during the previous week was measured on a 10-cm VAS for pain (where 0 = no pain and 10 = worst possible pain) by the child if age ≥9 years or by the parents if the child was age <9 years. VAS pain was assessed with the question "How do you rate your/your child's pain due to your/his or her illness in the past week?" As in previous studies on pain in JIA (15,24–26), pain analyses were explored both by categorization of VAS pain (0, >0 to 3, >3 to 7, or >7 to 10), and dichotomized into 0 (no pain) or >0. We performed subanalyses on VAS pain scores in participants age <9 years and ≥9 years to look for any discrepancies between parent- and patient-reported pain (27). Self-reported physical disability questionnaires were the disease-specific and validated Childhood Health Assessment Questionnaire (C-HAQ; where 0 = no difficulty and 3 = unable to do) if the child was age <18 years (28,29), and the Health Assessment Questionnaire (HAQ; where 0 = no difficulty and 3 = unable to do) if age ≥18 years (30). Children age ≥9 years filled out the C-HAQ, and parents filled out the questionnaire for those age <9 years.

For children age <18 years, the parent form of the generic health-related quality-of-life instrument, the Child Health Questionnaire Parent Form (CHQ-PF50, or simply CHQ), was answered by the parents, yielding a physical summary score and a psychosocial summary score (range 0–100, where 0 = worst, with a mean ± SD score of 50 ± 10) (28,31,32). This instrument is designed to capture the child's physical and psychosocial well-being

independent of his/her disease, and it is comparable to norm scores from the general US population. Damage was scored by experienced pediatric rheumatologists using the Juvenile Arthritis Damage Index (JADI) assessment of articular damage (JADI-A) (range 0–72, where 0 = no damage) and extraarticular damage (JADI-E) (range 0–17, where 0 = no damage) (8,32). Damage was defined as either JADI-A and/or JADI-E score >0. As in previous studies, C-HAQ/HAQ and JADI scores were dichotomized into 0 (no disability, no damage) or >0 (8,24,33). Physical and psychosocial summary scores of the CHQ were dichotomized into <40 (poor health) or ≥40 (better health) (28,31). Remission was defined according to the preliminary criteria described by Wallace et al (34). Remission status was dichotomized into remission without medication or not in remission without medication (33). The latter included active disease, inactive disease not yet in remission, and in remission while taking medication.

Ethics approval. Medical research ethics committees from each participating country gave their approval according to national practice and regulations in accordance with the Declaration of Helsinki. Written informed consent was obtained from children age ≥16 years and from their parents if age <16 years.

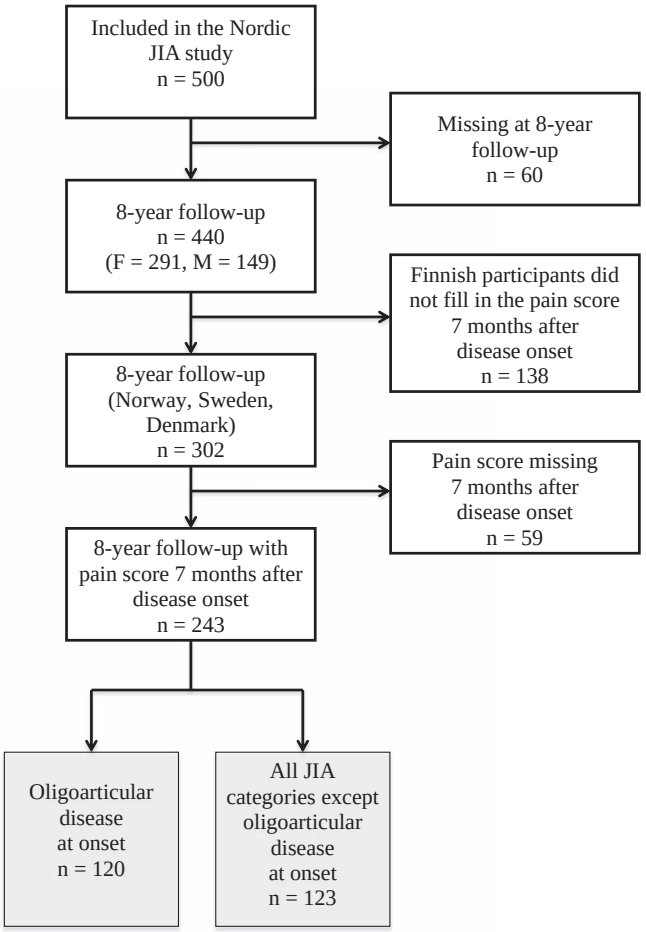


Figure 1. Flow-chart of the study population. JIA = juvenile idiopathic arthritis; F = female; M = male.

Statistical analysis. We used descriptive statistics with median and interquartile ranges (IQRs) for continuous variables, and absolute frequency percentage with 95% confidence intervals (95% CIs) for categorical variables. To evaluate the predictive value of pain at baseline for outcome measures after 8 years and medication during the disease course, model-based absolute risks were estimated after binominal regression using the post-estimation command lincom in Stata software, version 14. Sex adjustment was weighted 0.7 for girls to mimic the distribution in the population. In additional analyses, we also adjusted for age. To estimate absolute risks, we used the mean age at disease onset of 6.8 years. We used logistic regression to estimate the odds ratio with 95% CIs using VAS pain as a continuous variable. In further analyses, we made receiver operator characteristic (ROC) curves based on measures of sensitivity and specificity. The area

Table 1. Clinical characteristics of the juvenile idiopathic arthritis (JIA) study population*

Characteristic	Total no.	Values
Female	243	170 (70)
Oligoarticular JIA at onset	243	120 (49)
Age at disease onset, median (IQR) years	243	6.3 (2.9–10.3)
Age at 8-year follow-up, median (IQR) years	243	14.9 (11.1–18.5)
Disease duration at baseline visit, median (IQR) months	243	7 (6–9)
Disease duration at 8-year follow-up, median (IQR) months	243	97 (95–102)
VAS pain >0 at baseline visit†	243	185 (76)
VAS pain >0 at 8-year follow-up†	204	117 (57)
C-HAQ/HAQ >0 at 8-year follow-up	207	80 (39)
CHQ PhS <40 at 8-year follow-up	132	25 (19)
CHQ PsS <40 at 8-year follow-up	132	7 (5)
JADI >0 at 8-year follow-up	203	46 (23)
Not in remission at 8-year follow-up‡	236	135 (57)

* Values are the number (%) unless indicated otherwise. IQR = interquartile range; VAS = visual analog scale; C-HAQ = Childhood Health Assessment Questionnaire (used for age <18 years); HAQ = Health Assessment Questionnaire (used for age ≥18 years); CHQ PhS = Child Health Questionnaire physical summary score (range 0–100); CHQ PsS = Child Health Questionnaire psychosocial summary score (range 0–100); JADI = Juvenile Arthritis Damage Index.
† Self-reported pain was measured on a 10-cm VAS pain scale.
‡ Not in remission without medication according to the definition by Wallace et al (ref. 34).

under the curve was calculated with 95% CIs, and with the following interpretations: an area of 0.5 or lower was considered to be no discrimination, ≥0.7 to <0.8 as acceptable discrimination, ≥0.8 to <0.9 as excellent discrimination, and ≥0.9 as outstanding discrimination (35). Statistical analyses were carried out using STATA software, version 14.

RESULTS

Clinical characteristics of the study group. Of the 500 patients included from the 4 Nordic countries, 440 participated in the 8-year follow-up. Due to lack of baseline pain scores, all Finnish participants (n = 138) and an additional 59 participants from the other countries were excluded. The final study population consisted of 243 children (Figure 1) with a median baseline visit at 7 months and a final follow-up visit at 97 months (Table 1). Among these participants 70% were female, 49% had oligoarticular disease, and the median age was 6.3 years at disease onset and 14.9 years at follow-up (Table 1). The diagnostic delay was short, and the median interval between disease onset and diagnosis of arthritis by a physician was 50.5 days (IQR 14–101 days). Of these 243 participants, intraarticular glucocorticoid injections had been given to 91 participants, and for 34 of these the drug had been given within the last 3 months of the baseline visit (results not shown). At this baseline visit, none of the participants were taking biologics, but 20 were taking systemic steroids. Methotrexate was used by 31 of the participants, and of those, 8 had cumulative doses ≥100 mg. The 60 participants who did not participate in the 8-year study did not differ significantly in the proportion of oligoarticular JIA or with respect to sex, and had a median follow-up of 47 months (range 5–83 months). At their last registered visit, 30 participants (50%) had a pain assessment, including 21 with VAS pain scores >0, and 9 with VAS pain scores = 0. Participants excluded from the current study due to lack of pain data at baseline had a lower median age (5.1 versus 6.3 years) and the proportion of males was slightly higher (39% versus 30%). There was no difference in the proportion of oligoarticular JIA at onset and remission status at 8 years between the included and excluded participants.

Pain scores. More participants reported a VAS pain score >0 at the baseline visit (76%) than at 8 years (57%). The mean pain intensity score (VAS pain) of those reporting pain was higher at baseline, 3.0 (95% CI 2.6, 3.3) than at 8 years, 2.4 (95% CI 2.0, 2.8). The distribution of pain intensity scores at the baseline visit and at the 8-year follow-up is shown in Supplementary Figure 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23715/abstract. For participants age <9 years at baseline, 118 of 159 (74%) had a parent-reported pain score >0 with a mean intensity of 2.7 (95% CI 2.3, 3.0), while 67 of 84 participants (79%) age ≥9 years had a patient-reported pain score >0, with a mean intensity of 3.5 (95% CI 2.9, 4.1). Among participants with pain measures both at baseline and 8 years (n = 204), 50% reported a VAS pain score >0 at both visits, and 19% reported no pain at both visits (Figure 2). We divided this group into participants ages <9 and ≥9 years at baseline. Participants age <9 years had parent-reported pain scores at their first visit and patient-reported pain scores at their last visit, and 48% reported a VAS pain score >0 at both visits. Participants age ≥9 years had only patient-reported pain scores, and 55% reported a VAS pain score >0 at both visits (results not shown).

Baseline pain scores and long-term outcome measures. The association between baseline pain scores subdivided into 4 categories of pain intensity and long-term

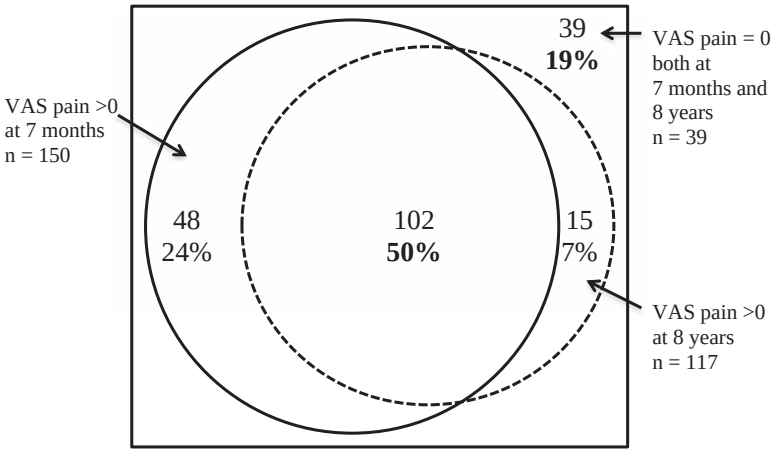


Figure 2. Venn diagram demonstrating pain persistency in the Nordic juvenile idiopathic arthritis study cohort. The cohort included 204 participants with pain measures 7 months after disease onset and at the 8-year follow-up. Disease-related pain was measured on a 10-cm visual analog scale (VAS pain) (where 0 = no pain and 10 = worst possible pain). A continuous circle represents a VAS pain score >0 at 7 months and a broken circle represents a VAS pain score >0 at 8 years.

Table 2. Association between baseline pain report at 7 months after disease onset and outcomes at 8-year follow-up in juvenile idiopathic arthritis*

Baseline VAS pain	VAS pain >0†	C-HAQ/ HAQ >0	CHQ PhS <40	JADI >0	Not in remission‡
0§	15/54, 28 (16, 40)	9/54, 18 (8, 28)	3/32, 11 (1, 21)	4/53, 8 (1, 15)	14/55, 26 (14, 37)
>0–3§	53/86, 62 (51, 72)	31/86, 35 (25, 45)	8/62, 12 (4, 21)	25/83, 30 (20, 40)	61/95, 64 (54, 74)
>3–7§	44/58, 74 (63, 85)	34/60, 56 (44, 69)	14/38, 37 (22, 52)	15/60, 25 (14, 36)	52/73, 71 (61, 81)
>7–10§	5/6, 90 (66, 114)	6/7, 94 (80, 107)	0 0	2/7, 28 (–5, 61)	8/13, 61 (35, 88)
Continuous, OR (95% CI)⁋	1.5 (1.2, 1.7)	1.4 (1.2, 1.6)	1.4 (1.1, 1.7)	1.1 (0.9, 1.2)	1.2 (1.1, 1.4)

* Values are the number/total number, percentage (95% confidence interval [95% CI]) unless indicated otherwise. VAS = visual analog scale; C-HAQ = Childhood Health Assessment Questionnaire (used for age <18 years); HAQ = Health Assessment Questionnaire (used for age ≥18 years); CHQ PhS = Child Health Questionnaire physical summary score (range 0–100); JADI = Juvenile Arthritis Damage Index; OR = odds ratio.
† Self-reported pain was measured on a 10-cm VAS pain scale.
‡ Not in remission without medication according to the definition by Wallace et al (ref. 34).
§ Self-reported pain was measured on a 10-cm VAS pain scale, adjusted for sex, weighted 0.7 for girls.
⁋ Self-reported pain was measured on a 10-cm VAS pain scale, analyzed with VAS pain as a continuous variable, adjusted for sex, weighted 0.7 for girls.

outcome measures is shown in Table 2. Participants reporting a VAS pain score >0 at the baseline visit more frequently reported pain and functional disability (C-HAQ/HAQ >0) at the 8-year follow-up. A distinct dose-response curve was observed with increasing pain intensity at baseline. Using VAS pain as a continuous variable, we observed an increased odds ratio for

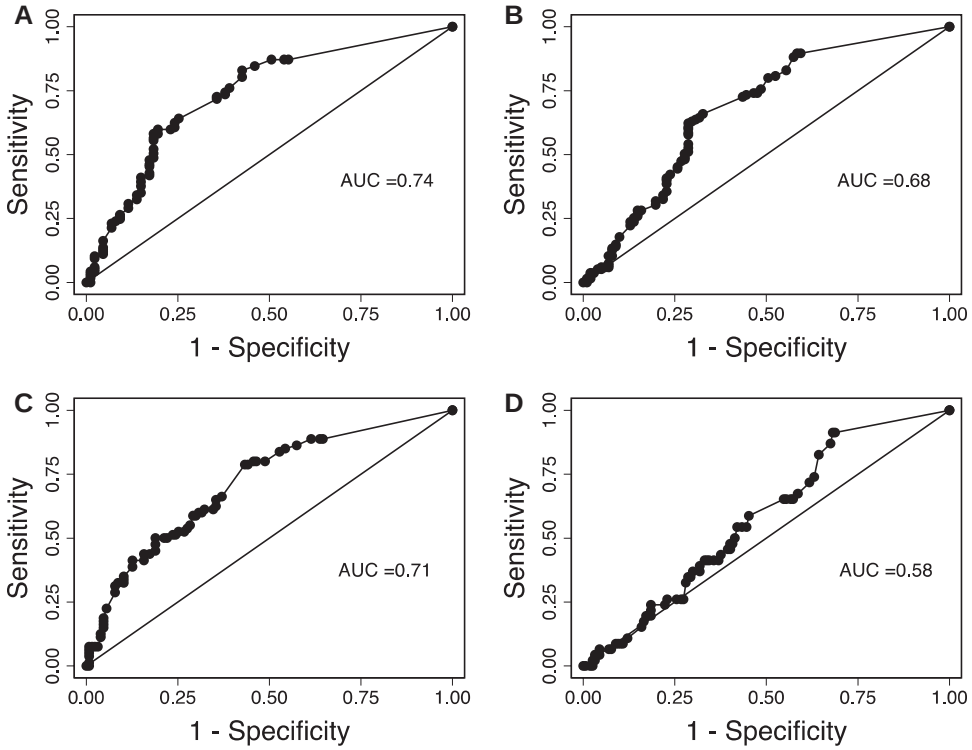


Figure 3. Receiver operator characteristic curves in the Nordic juvenile idiopathic arthritis study cohort for different disease outcomes after 8 years compared to self-reported disease-related pain at 7 months after disease onset, measured on a 10-cm visual analog scale (VAS; where 0 = no pain and 10 = worst possible pain). Remission was defined according to the preliminary criteria described by Wallace et al (34). Functional disability was measured with the Childhood Health Assessment Questionnaire/Health Assessment Questionnaire. Damage was measured with the Juvenile Arthritis Damage Index, articular and extraarticular. The area under the curve (AUC) values were 0.74 (95% CI 0.67, 0.80) for persistent pain, 0.68 (95% CI 0.61, 0.75) for not being in remission, 0.71 (95% CI 0.64, 0.79) for functional disability, and 0.58 (95% CI 0.50, 0.67) for joint damage. **A**, Persistent pain. **B**, Not in remission. **C**, Functional disability. **D**, Damage.

Table 3. Association between baseline pain report at 7 months after disease onset and outcomes at 8-year follow-up in juvenile idiopathic arthritis with oligoarticular onset*

Baseline VAS pain	Extend oligo/others†	VAS pain >0‡	C-HAQ/ HAQ >0	JADI >0	Not in remission§
0⁋	12/40, 30 (16, 44)#	10/37, 26 (12, 40)	7/37, 16 (7, 26)	3/36, 7 (0, 15)	11/38, 29 (14, 43)
>0⁋	39/80, 48 (38, 60)**	40/65, 61 (50, 73)	26/66, 39 (28, 51)	19/65, 29 (18, 40)	51/78, 65 (55, 76)
Continuous, OR (95% CI)††	1.3 (1.1, 1.6)	1.4 (1.1, 1.8)	1.4 (1.1, 1.8)	1.0 (0.8, 1.3)	1.3 (1.0, 1.5)

* Values are the number/total number, percentage (95% confidence interval [95% CI]) unless indicated otherwise. VAS = visual analog scale; Extend oligo = extended oligoarticular juvenile idiopathic arthritis (JIA); C-HAQ = Childhood Health Assessment Questionnaire (used for age <18 years); HAQ = Health Assessment Questionnaire (used for age ≥18 years); JADI = Juvenile Arthritis Damage Index; OR = odds ratio.
† Oligoarticular JIA at 6 months changed to either extended oligoarticular or other JIA categories at the 8-year follow-up.
‡ Self-reported pain was measured on a 10-cm VAS pain scale.
§ Not in remission without medication according to the definition by Wallace et al (ref. 34).
⁋ Self-reported pain was measured on a 10-cm VAS pain scale, adjusted for sex, weighted 0.7 for girls.
Others were 1 with enthesitis-associated arthritis and 2 with undifferentiated arthritis.
** Others were 2 with psoriatic arthritis, 6 with enthesitis-associated arthritis, and 2 with undifferentiated arthritis.
†† Self-reported pain was measured on a 10-cm VAS pain scale, analyzed with VAS pain as a continuous variable, adjusted for sex, weighted 0.7 for girls.

the different long-term outcomes. Functional disability as presented by the CHQ physical summary score demonstrated similar results. Participants reporting pain at baseline more frequently were not in remission without medication at follow-up, compared to those reporting no baseline pain. A similar association between increasing pain intensity at baseline and long-term remission status was observed, but the dose-response relationship tended to level out at the most extreme pain intensities. Participants reporting no pain at baseline rarely reported pain (28%) and functional disability (18%) at 8 years, and 74% were in remission without medication. In all analyses, we adjusted for sex, and additional adjustment for age did not change the results. Similar to the results on long-term remission, pain, and functional disability, baseline pain was associated with the use of disease-modifying antirheumatic drugs (DMARDs) and biologics during the 8-year disease course (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23715/abstract). In contrast to the other 8-year outcome measures, psychosocial health assessed with the CHQ psychosocial summary score did not show an association with the baseline pain score (results not shown). The predictive ability of baseline pain was also analyzed using ROC curves, giving acceptable discrimination between early pain scores and long-term outcomes of pain, functional disability, and not being in remission at 8 years, but no clear discrimination was observed for long-term damage (Figure 3).

Baseline pain scores and long-term outcomes in the oligoarticular category. Among participants with oligoarticular JIA reporting a VAS pain score >0 at the baseline visit, 48% (95% CI 38, 60) developed extended oligoarticular disease or other JIA categories during the course of the disease compared to 30% (95% CI 16, 44) of those reporting no pain (Table 3). Also, a higher proportion of participants with a VAS pain score >0 at baseline was not in remission without medication, 65% (95% CI 55, 76), and reported pain, 61% (95% CI 50, 73) at 8 years, compared to those reporting no pain at baseline. In contrast, 70% (95% CI 56, 84) of those reporting no pain at baseline remained in the persistent oligoarticular JIA category at the 8-year follow-up. The associations were strengthened using VAS pain as a continuous variable. In all analyses, we adjusted for sex, and additional adjustment for age did not change the results.

DISCUSSION

Among participants in the population-based Nordic JIA study, a higher proportion reported disease-specific pain 7 months after disease onset, and the pain was of higher intensity, compared to the 8-year follow-up. Half of the participants reporting pain at baseline also reported pain at 8 years. Self-reported pain early in the disease course predicted more pain, more functional disability, more damage, more use of DMARDs/biologics, and more long-term disease activity at 8 years. In addition, participants with oligoarticular JIA and a VAS pain score >0 at baseline more often developed an extended disease or other unfavorable JIA categories.

The strength of our study is the longitudinal and population-based design, a robust international cohort, and the use of validated and multidimensional outcome measures. The novelty in looking at associations between early pain report and long-term remission status, medication, and development into more unfavorable disease categories is also a major strength. Some limitations must be recognized. The exclusion of all the Finnish participants reduced the number

of study participants but did not change the population-based design of the study. The missing early pain scores from the other countries might have skewed the remaining cohort, but the distribution of JIA categories and the remission status were comparable among those with or without early pain scores. Even though the VAS pain instrument is disease-specific, we cannot rule out that other musculoskeletal co-conditions, such as generalized joint hypermobility, specific onset symptoms, or differences in timing of diagnosis, might have influenced the child's/parent's pain rating. However, these possibilities are a challenge to all pain research. Since pain is a subjective descriptor, our cohort is close to population-based, and because trained pediatric rheumatologists ascertained the JIA diagnosis, we do not think that these possibilities will seriously disturb the interpretation of our results. Similarly, we cannot ascertain the nature of bodily pain that the participants scored when filling out this question in the CHQ questionnaire. However, we only used this question in accordance with the CHQ instructions, as one of many items describing a summary of physical function, and not as a pain measure. Parents reporting their child's pain for children age <9 years constituted a majority of pain reports at the baseline visit, but a small minority at the 8-year follow-up. We cannot rule out some element of parent/child discordance, although the subanalyses on pain reports according to age at baseline seem to indicate that discordance was not a major problem. This result is in accordance with a study from 2006 showing moderate agreement between parent's and child's pain rating (17). Our results are not directly comparable, because the parent/child pain reports are not from the same visit. The early pain scores from the baseline study visit 7 months after disease onset were given by participants both while taking and not taking medication, but only a few had started DMARDs.

Consistent with previous research, we found pain as a frequent symptom among the participants in our study cohort (14,16,36). We found a reduction in the number of participants reporting pain from baseline (76%) to the 8-year follow-up (57%). A quite similar reduction was found by Lovell and Walco (37) in 1989, demonstrating pain frequency of 60% at baseline, 50% at 1-year follow-up, and 40% at 5-year follow-up. In a recent 30-year follow-up study of JIA in Norway, 66% of the participants reported pain of some degree (24).

In accordance with other studies, the intensity of pain was mainly in the mild-to-moderate range (19,37,38). Our results on early pain intensity with a mean VAS pain score of 3.0 are consistent with a recent cross-sectional study in children and adolescents with JIA from the southeastern region of the US, showing a mean VAS pain score of 2.6 (13). Our results on pain intensity at the 8-year follow-up appear to be lower compared to other studies (14,39). Those studies are, however, skewed to the severe end of the JIA spectrum, whereas our population-based study included the full disease spectrum. Also, to compare studies on pain, age and disease duration must be taken into account.

Even in the biologic era with generally good disease control, persistent pain during the course of the disease remains a concern (26,40,41). Half of our participants reporting pain at 7 months after disease onset also reported pain at the 8-year follow-up, indicating high pain persistency. This finding is in agreement with results from other studies showing that a significant number of children and adolescents with JIA continue to report pain during the course of disease and into adulthood (14,24,42). Notably, the proportion of pain persistence is fairly similar, whether the parents report their child's pain, or whether the pain is self-reported at baseline. Pain persistence despite a seemingly good treatment response supports theories that the causes of pain are multifactorial (41,43). Both psychosocial and biologic factors contribute to these children's subjective experience of pain (38,44,45).

Pain as a predictor of unfavorable health-related quality of life in children with JIA is widely studied (21,23,46,47). In a multinational quality-of-life study from the Pediatric Rheumatology International Trials Organization, pain was found to be a predictor of psychosocial well-being (21). In agreement with our study, previous studies have shown that pain at presentation was a strong predictor of persistent pain (42,48). In accordance with our results, pain as a predictor of functional disability was also found in a small cross-sectional study from the US (49). Except for health-related quality-of-life outcomes and functional disability, studies that specifically address pain as a predictor of other long-term outcomes, such as remission, damage, medication, and changing of JIA categories, are lacking. Our results demonstrate for the first time that early pain is associated with not achieving remission without medication in a long-term perspective. We also demonstrate, for the first time, that early pain reports predict a higher risk of development into extended oligoarticular or other unfavorable JIA categories during the course of the disease. This finding suggests that early pain may be an indicator of subclinical disease activity or a marker of a more severe disease category. This possibility is also supported by the fact that a higher proportion of participants with an early pain report used DMARDs/biologics during the course of the disease.

Even though pain assessment has been highlighted as a quality measure of pediatric arthritis care (50), pain scores are infrequently used as guiding tools in daily care of these patients (41). Our results demonstrate that pain in children with JIA at an early stage in their disease should be taken seriously, not just to relieve ongoing discomfort, but probably also as a sign of ongoing clinical or subclinical disease activity. This necessity emphasizes the importance of pain assessment in routine care of children and adolescents with JIA. In a Canadian study where patients, parents, and clinicians were asked what matters most in the care of JIA, pain was 1 of the 5 most important factors (11). The active joint count was the only 1 of these 5 factors that is included in the pediatric version of the ACR core variables for clinical care in

children with JIA (12). The association between early pain reports and long-term unfavorable outcome adds to the discussion on the validity of the ACR core variables, and on whether pain should be included in these variables. In conclusion, early self-reported pain in JIA is common, tends to persist, and seems to predict unfavorable long-term disease outcome in several outcome dimensions.

ACKNOWLEDGMENTS

The authors thank the participants and their parents for making this study possible. We also thank the other members of the Nordic Study Group of Pediatric Rheumatology: Gudmund Marhaug (Trondheim), Boel Anderson-Gäre (Jonköping), Freddy Karup Pedersen (Copenhagen), and Pekka Lahdenne and Pirkko Pelkonen (Helsinki) for their inspiring cooperation. We are also grateful to local participating physicians for collecting data: Nina Moe and Kjell Berntzen (Trondheim), Astri Lang, Nils Thomas Songstad, and Anne Elisabeth Ross (Tromsø), Eric Ronge (Skövde), Mikael Damgaard and Nils Olof Jonsson (Jönköping), Anna-Lena Nilsson (Östersund), Anders Berner and Hans Ekström (Karlstad), Agne Lind and Lars Hammarén (Borås), and Johan Robinsson (Trollhättan).

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. Arnstad 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. Herlin, Berntson, Fasth, Nielsen, Zak, Aalto, Nordal, Rygg.

Acquisition of data. Peltoniemi, Herlin, Berntson, Nielsen, Ekelund, Zak, Aalto, Nordal, Rygg.

Analysis and interpretation of data. Arnstad, Rypdal, Glerup, Nordal, Romundstad, Rygg.

REFERENCES

1. Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007;369:767–78.
2. Berntson L, Andersson Gäre B, Fasth A, Herlin T, Kristinsson J, Lahdenne P, et al. Incidence of juvenile idiopathic arthritis in the Nordic countries: a population based study with special reference to the validity of the ILAR and EULAR criteria. J Rheumatol 2003;30:2275–82.
3. Riise OR, Handeland KS, Cvancarova M, Wathne KO, Nakstad B, Abrahamsen TG, et al. Incidence and characteristics of arthritis in Norwegian children: a population-based study. Pediatrics 2008;121:e299–306.
4. Moe N, Rygg M. Epidemiology of juvenile chronic arthritis in northern Norway: a ten-year retrospective study. Clin Exp Rheumatol 1998;16:99–101.
5. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
6. Flato B, Lien G, Smerdel A, Vinje O, Dale K, Johnston V, et al. Prognostic factors in juvenile rheumatoid arthritis: a case-control study revealing early predictors and outcome after 14.9 years. J Rheumatol 2003;30:386–93.
7. Ravelli A. Toward an understanding of the long-term outcome of juvenile idiopathic arthritis. Clin Exp Rheumatol 2004;22:271–5.
8. Nordal E, Zak M, Aalto K, Berntson L, Fasth A, Herlin T, et al. Ongoing disease activity and changing categories in a long-term Nordic cohort study of juvenile idiopathic arthritis. Arthritis Rheum 2011;63:2809–18.
9. Selvaag AM, Aulie HA, Lilleby V, Flato B. Disease progression into adulthood and predictors of long-term active disease in juvenile idiopathic arthritis. Ann Rheum Dis 2016;75:190–5.
10. Van Dijkhuizen EH, Wulffraat NM. Early predictors of prognosis in juvenile idiopathic arthritis: a systematic literature review. Ann Rheum Dis 2015;74:1996–2005.
11. Guzman J, Gomez-Ramirez O, Jurencak R, Shiff NJ, Berard RA, Duffy CM, et al. What matters most for patients, parents, and clinicians in the course of juvenile idiopathic arthritis? A qualitative study. J Rheumatol 2014;41:2260–9.
12. Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. Preliminary definition of improvement in juvenile arthritis. Arthritis Rheum 1997;40:1202–9.
13. Bromberg MH, Connelly M, Anthony KK, Gil KM, Schanberg LE. Self-reported pain and disease symptoms persist in juvenile idiopathic arthritis despite treatment advances: an electronic diary study. Arthritis Rheumatol 2014;66:462–9.
14. Schanberg LE, Anthony KK, Gil KM, Maurin EC. Daily pain and symptoms in children with polyarticular arthritis. Arthritis Rheum 2003;48:1390–7.
15. Tupper SM, Rosenberg AM, Pahwa P, Stinson JN. Pain intensity variability and its relationship with quality of life in youths with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2013;65:563–70.
16. Weiss JE, Luca NJ, Boneparth A, Stinson J. Assessment and management of pain in juvenile idiopathic arthritis. Paediatr Drugs 2014;16:473–81.
17. Garcia-Munitis P, Bandeira M, Pistorio A, Magni-Manzoni S, Ruperto N, Schivo A, et al. Level of agreement between children, parents, and physicians in rating pain intensity in juvenile idiopathic arthritis. Arthritis Rheum 2006;55:177–83.
18. Stinson JN, Kavanagh T, Yamada J, Gill N, Stevens B. Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. Pain 2006;125:143–57.
19. Shiff NJ, Tupper S, Oen K, Guzman J, Lim H, Lee CH, et al. Trajectories of pain severity in juvenile idiopathic arthritis: results from the Research in Arthritis in Canadian Children Emphasizing Outcomes cohort. Pain 2018;159:57–66.
20. Stinson JN, Luca NJ, Jibb LA. Assessment and management of pain in juvenile idiopathic arthritis. Pain Res Manag 2012;17:391–6.
21. Gutierrez-Suarez R, Pistorio A, Cespedes Cruz A, Norambuena X, Flato B, Rumba I, et al. Health-related quality of life of patients with juvenile idiopathic arthritis coming from 3 different geographic areas: the PRINTO multinational quality of life cohort study. Rheumatology (Oxford) 2007;46:314–20.
22. Shaw KL, Southwood TR, Duffy CM, McDonagh JE. Health-related quality of life in adolescents with juvenile idiopathic arthritis. Arthritis Rheum 2006;55:199–207.
23. Klotsche J, Minden K, Thon A, Ganser G, Urban A, Horneff G. Improvement in health-related quality of life for children with juvenile idiopathic arthritis after start of treatment with etanercept. Arthritis Care Res (Hoboken) 2014;66:253–62.
24. Tollisen A, Selvaag AM, Aulie HA, Lilleby V, Aasland A, Lerdal A, et al. Physical functioning, pain, and health-related quality of life in adults with juvenile idiopathic arthritis: a longitudinal 30-year followup study. Arthritis Care Res (Hoboken) 2018;70:741–9.

25. Consolaro A, Vitale R, Pistorio A, Lattanzi B, Ruperto N, Malattia C, et al. Physicians' and parents' ratings of inactive disease are frequently discordant in juvenile idiopathic arthritis. J Rheumatol 2007;34:1773–6.
26. Lomholt JJ, Thastum M, Herlin T. Pain experience in children with juvenile idiopathic arthritis treated with anti-TNF agents compared to non-biologic standard treatment. Pediatr Rheumatol Online J 2013;11:21.
27. Lal SD, McDonagh J, Baildam E, Wedderburn LR, Gardner-Medwin J, Foster HE, et al. Agreement between proxy and adolescent assessment of disability, pain, and well-being in juvenile idiopathic arthritis. J Pediatr 2011;158:307–12.
28. Ruperto N, Ravelli A, Pistorio A, Malattia C, Cavuto S, Gado-West L, et al. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries: review of the general methodology. Clin Exp Rheumatol 2001;19 Suppl 23:S1–9.
29. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 1994;37:1761–9.
30. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137–45.
31. Landgraf JM, Abetz L, Ware JE. Child Health Questionnaire (CHQ): a user's manual. Boston: Health Institute, New England Medical Center; 1996.
32. Viola S, Felici E, Magni-Manzoni S, Pistorio A, Buoncompagni A, Ruperto N, et al. Development and validation of a clinical index for assessment of long-term damage in juvenile idiopathic arthritis. Arthritis Rheum 2005;52:2092–102.
33. Bertilsson L, Andersson-Gare B, Fasth A, Petersson IF, Forsblad-D'elia H. Disease course, outcome, and predictors of outcome in a population-based juvenile chronic arthritis cohort followed for 17 years. J Rheumatol 2013;40:715–24.
34. Wallace CA, Ruperto N, Giannini E. Preliminary criteria for clinical remission for select categories of juvenile idiopathic arthritis. J Rheumatol 2004;31:2290–4.
35. David W. Hosmer, Stanley Leweshow. Area under the ROC curve. In: Cressie NA, editor. Applied logistic regression. 2nd ed. New York: John Wiley; 2000. p. 160–4.
36. Oen K, Reed M, Malleson PN, Cabral DA, Petty RE, Rosenberg AM, et al. Radiologic outcome and its relationship to functional disability in juvenile rheumatoid arthritis. J Rheumatol 2003;30:832–40.
37. Lovell DJ, Walco GA. Pain associated with juvenile rheumatoid arthritis. Pediatr Clin North Am 1989;36:1015–27.
38. Schanberg LE, Lefebvre JC, Keefe FJ, Kredich DW, Gil KM. Pain coping and the pain experience in children with juvenile chronic arthritis. Pain 1997;73:181–9.
39. Guzman J, Henrey A, Loughin T, Berard RA, Shiff NJ, Jurencak R, et al. Predicting which children with juvenile idiopathic arthritis will have a severe disease course: results from the ReACCh-Out cohort. J Rheumatol 2017;44:230–40.
40. Anink J, Prince FH, Dijkstra M, Otten MH, Twilt M, ten Cate R, et al. Long-term quality of life and functional outcome of patients with juvenile idiopathic arthritis in the biologic era: a longitudinal follow-up study in the Dutch Arthritis and Biologicals in Children Register. Rheumatology (Oxford) 2015;54:1964–9.
41. Giancane G, Alongi A, Rosina S, Calandra S, Consolaro A, Ravelli A. Open issues in the assessment and management of pain in juvenile idiopathic arthritis. Clin Exp Rheumatol 2017;35 Suppl 107:123–6.
42. Thastum M, Herlin T. Pain-specific beliefs and pain experience in children with juvenile idiopathic arthritis: a longitudinal study. J Rheumatol 2011;38:155–60.
43. Malleson PN, Oen K, Cabral DA, Petty RE, Rosenberg AM, Cheang M. Predictors of pain in children with established juvenile rheumatoid arthritis. Arthritis Rheum 2004;51:222–7.
44. 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.
45. La Hausse de Lalouviere L, Ioannou Y, Fitzgerald M. Neural mechanisms underlying the pain of juvenile idiopathic arthritis. Nat Rev Rheumatol 2014;10:205–11.
46. Oen K, Guzman J, Dufault B, Tucker LB, Shiff NJ, Duffy KW, et al. Health-related quality of life in an inception cohort of children with juvenile idiopathic arthritis: a longitudinal analysis. Arthritis Care Res (Hoboken) 2018;70:134–44.
47. Haverman L, Grootenhuis MA, van den Berg JM, van Veenendaal M, Dolman KM, Swart JF, et al. Predictors of health-related quality of life in children and adolescents with juvenile idiopathic arthritis: results from a web-based survey. Arthritis Care Res (Hoboken) 2012;64:694–703.
48. Rashid A, Cordingley L, Carrasco R, Foster HE, Baildam EM, Chieng A, et al. Patterns of pain over time among children with juvenile idiopathic arthritis. Arch Dis Child 2018;103:437–43.
49. Feinstein AB, Forman EM, Masuda A, Cohen LL, Herbert JD, Moorthy LN, et al. Pain intensity, psychological inflexibility, and acceptance of pain as predictors of functioning in adolescents with juvenile idiopathic arthritis: a preliminary investigation. J Clin Psychol Med Settings 2011;18:291–8.
50. Lovell DJ, Passo MH, Beukelman T, Bowyer SL, Gottlieb BS, Henrickson M, et al. Measuring process of arthritis care: a proposed set of quality measures for the process of care in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011;63:10–6.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 970–976
DOI 10.1002/acr.23702

The Risk of Ischemic and Hemorrhagic Stroke in Patients With Idiopathic Inflammatory Myopathies: A Swedish Population-Based Cohort Study

John Moshtaghi-Svensson,[1] Ingrid E. Lundberg,[2] Mia Von Euler,[1] Elizabeth V. Arkema,[1] and Marie Holmqvist[2]

Objective. To study the occurrence of ischemic stroke and hemorrhagic stroke in patients with idiopathic inflammatory myopathies (IIMs) compared to that in the general population and to investigate how it varies by sex, age, clinical subdiagnosis, and time since IIM diagnosis.

Methods. All patients in Sweden with newly diagnosed IIM were identified from the National Patient Register, and general population comparators were identified from the Total Population Register. The study population was followed prospectively until death, emigration, December 2013, or first incident stroke. Incidence rates, rate differences, and hazard ratios (HRs) comparing patients with IIMs to the general population were estimated and stratified by age, sex, type of IIM, and time since diagnosis. To account for the competing risk of death, the subdistribution HR was estimated using Fine and Gray models.

Results. We observed 34 and 229 stroke events in 663 IIM patients and 6,673 comparators, respectively. The HR was elevated for ischemic stroke (HR 2.1 [95% confidence interval (95% CI) 1.4, 3.0]). Few hemorrhagic stroke events were identified, but an increased risk was observed (HR 1.9 (95% CI 0.7, 5.5]). The association remained elevated for both outcomes when taking the competing risk of death into account. For ischemic stroke, the rate difference was highest in the oldest age group (≥68 years), while the HR was highest in the youngest age group (<56 years).

Conclusion. Our findings indicate that the risk of both ischemic stroke and hemorrhagic stroke is increased in patients with IIMs, but it should be kept in mind that stroke is a rare event. Focus on prevention should be directed toward groups with the highest absolute risk, especially older patients.

INTRODUCTION

Idiopathic inflammatory myopathies (IIMs) are rare chronic rheumatic diseases mainly affecting skeletal muscle, causing weakness and low endurance. IIM is most commonly divided into 3 adult clinical subdiagnoses: polymyositis (PM), dermatomyositis (DM), and sporadic inclusion body myositis (IBM) (1). Advances in the treatment of IIM have improved the prognosis, but IIM is still associated with increased mortality and morbidity (2). Cardiovascular disease is a major cause of death in IIM patients, but clinically manifest heart disease is uncommon (3).

A few cohort studies have investigated the risk of stroke in patients with PM and those with DM (4–7), with all showing an increased risk. A recent meta-analysis including 3 of these studies demonstrated a pooled relative risk for ischemic stroke of 1.61 in IIM patients compared to non-IIM populations, but due to the heterogeneity of the included studies, the actual risk of stroke following IIM diagnosis is still uncertain (8). In addition, few of the studies published thus far have investigated the risk of the 2 main subtypes of stroke (ischemic and hemorrhagic) separately and the stroke risk over time in individuals with IIM. Rate differences, which are important for communicating absolute risk, have not been reported, and only 1 study has presented stratified results based on age and sex, making it difficult to identify high-risk groups. Last, previous studies have not investigated the association while taking into account the competing risk of death, which could be important from a prognostic perspective.

Supported by the Swedish Society of Medicine, the Swedish Research Council (grant K2014-52X-14045-14-3), the Swedish Rheumatism Association, King Gustaf V's 80-Year Foundation, the Karolinska Institutet, and the Stockholm County Council.

[1]John Moshtaghi-Svensson, PhD, Mia Von Euler, MD, PhD, Elizabeth V. Arkema, ScD: Karolinska Institutet, Stockholm, Sweden; [2]Ingrid E. Lundberg, MD, PhD, Marie Holmqvist, MD, PhD: Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden.

Drs. Arkema and Holmqvist contributed equally to this work.

Dr. Lundberg has received honoraria from Bristol-Myers Squibb (less than $10,000), and has received research support from Bristol-Meyers Squibb and AstraZeneca.

No other disclosures relevant to this article were reported. Address correspondence to John Moshtaghi-Svensson, PhD, D2:01 Karolinska University Hospital, 171 76 Stockholm, Sweden. E-mail: John.svensson@godjy.se.

Submitted for publication March 27, 2018; accepted in revised form July 10, 2018.

SIGNIFICANCE & INNOVATIONS

- The risk of ischemic stroke in patients with idiopathic inflammatory myopathies (IIMs) was twice that in the general population.
- The number of hemorrhagic stroke events was small, but the risk was doubled in patients with IIM compared to that in the general population.
- The absolute risk for ischemic stroke was highest in the oldest individuals, while the relative risk was the highest in the youngest individuals.
- Stroke is a rare event, and focus on prevention should be directed toward the groups with the highest absolute risk.

The overall goal of this study was to investigate whether there is an increased risk of ischemic stroke and hemorrhagic stroke following an IIM diagnosis compared to the risk in the general population. We sought to investigate when, in relation to disease onset, the risk of stroke is increased, and what demographic groups are at greatest risk. Because most previously published studies have used models that insufficiently account for death, we also aimed to model the relative risk of stroke, taking into account the competing risk of death.

PATIENTS AND METHODS

Study design. We conducted a population-based cohort study including patients with newly diagnosed IIM and general population comparators, in order to investigate the risk of incident stroke following a diagnosis of IIM. This study was approved by the regional ethics review board in Stockholm.

Setting. All Swedish residents have access to publically funded health care. Patients with IIMs are treated at hospital-based rheumatology or internal medicine units by specialists in rheumatology or at neurology units by specialists in neurology. Occasionally, patients with DM are treated in dermatology units. Linking between different health care and demographics register sources is possible through the use of each Swedish resident's unique personal identity number.

Study population. *IIM patients.* We used the National Patient Register (NPR) to identify all adults (ages ≥18 years) followed for IIM in Sweden. The NPR lists all non–primary care outpatient visits from 2001 and all hospitalizations from 1987 for all Swedish residents (9). For each visit, information is listed on the main diagnosis and up to 10 contributory diagnoses, using International Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10), codes used in Sweden since 1997. Individuals with ≥2 visits indicating IIM (See Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract for ICD codes) at a rheumatology, neurology, internal medicine, or dermatology unit were included in the study. We previously validated the ICD-10 codes used to identify IIM patients, using the Swedish Rheumatology Quality Register (SRQ), which includes the SweMyoNet register (10), as the gold standard. IIM patients are included in the SRQ by an IIM specialized rheumatologist. The positive predictive values for ICD-10 codes G724, M330, M331+M339, and M332 were all >80% (11).

Patients must have had their first-ever visit between 2002 and 2011 and a follow-up visit within 1–12 months, in order to exclude possible miscoded visits. This approach allows for a 12-month washout period before the study period to exclude prevalent cases and 12 months after the study period to allow sufficient time for a follow-up visit. The patient's follow-up visit was used as the index date. In addition, information on IIM patients was retrieved from the SRQ, which contains information on IIM-related clinical variables, including the diagnosis set by rheumatologists, since 2003.

Because there is no specific ICD-10 code for IBM, it is difficult to separate between PM and IBM using the NPR (see ref. 11). Therefore, included IIM patients were given the diagnosis "DM or other IIM." If the clinical subdiagnosis differed between the NPR and the SRQ, the SRQ was used as the gold standard.

General population comparators. Up to 10 individuals matched for age, sex, and place of residence were randomly selected from the Total Population Register (TPR) (12) and used as general population comparators. The general population comparators were assigned the same index date as their corresponding IIM patients. To be eligible as a general population comparator, each individual had to be living in Sweden at the time of matching (when the IIM patient was classified as having IIM).

Identification of outcome. Ischemic stroke and hemorrhagic stroke were identified from the NPR and the Cause of Death Register (CDR) (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract, for ICD codes). The CDR contains information on the main and contributory causes of death for almost all deaths in Sweden (13). The positive predictive value for stroke was shown to be 94% in the NPR and 87.3% in the CDR (14).

Exclusion criteria. All individuals with a history of stroke or stroke-related events (ischemic stroke, hemorrhagic stroke, or unspecified stroke), subarachnoid hemorrhage, sequelae of cerebrovascular disease and personal history of certain other diseases, or transient ischemic attack at baseline were excluded from the study (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract for ICD codes).

Follow-up. Follow-up started on the index date and ended at the time of the first stroke event under evaluation (ischemic stroke or hemorrhagic stroke), death, migration, or December 31,

2013. Ischemic stroke and hemorrhagic stroke were analyzed separately, allowing individuals who had an ischemic stroke to contribute person-time and events in the analysis of hemorrhagic stroke and vice versa. The death date was identified from the CDR and date of migration from the TPR. Two IIM patients and 6 comparators had both outcomes.

Covariates. Information on sex and date of birth was retrieved from the TPR. Educational level was retrieved from the Longitudinal Integration Database for Health Insurance and Labour Market Studies, which integrates information from the labor market, social sectors, and educational sectors and is updated yearly and categorized into <9 years, 10–12 years, and >12 years (15). Prevalent stroke risk factors were assessed at baseline. All individuals with ≥1 visit in the NPR indicating diabetes mellitus, hypertension, atrial fibrillation, or congestive heart disease prior to the index date were categorized as having that specific risk factor.

Statistical analysis. Crude incidence rates for ischemic stroke and hemorrhagic stroke were calculated per 1,000 person-years. Confidence intervals (CIs) were calculated assuming a Poisson distribution, using the exact method (16).

The risks of ischemic stroke and hemorrhagic stroke were assessed separately, because they might be caused by different mechanisms. The association between an IIM diagnosis and stroke was estimated using Cox proportional hazards models, with time since index date used as the time scale to calculate cause-specific hazard ratios (HRs) and 95% CIs as a measurement of relative risk. This cause-specific regression model, which assumes that the 2 competing events (in this case, stroke and death) are independent, estimates the HR in the population that has not yet experienced the competing event (death) or event of interest (stroke). To estimate the HR of stroke taking the competing risk of death into account, we calculated the subdistribution HRs using Fine and Gray (17) competing-risks regression models. Cumulative incidence was estimated for ischemic stroke at 1, 5, and 10 years after diagnosis, taking the competing risk of death into account. The proportional hazards assumption was tested by introducing an interaction term between the exposure and the log of the time scale. *P* values less than 0.05 were considered significant. In addition, age- and sex-adjusted rate differences between IIM and comparators were estimated using additive Poisson models.

Because the number of hemorrhagic stroke events was low, stratified estimates were performed for ischemic stroke only. Estimates for ischemic stroke were stratified by sex, age at diagnosis (in tertiles), and IIM subdiagnosis (DM and other IIMs). Furthermore, the relative risk of ischemic stroke stratified by time since index date (<1 year, 1 to <5 years, 5 to <12 years) was estimated using time-dependent covariates. Effect modification by time since index date was tested using a likelihood ratio test.

Sensitivity analyses. Estimates were adjusted for stroke risk factors (education level, diabetes mellitus, hypertension, atrial fibrillation, and congestive heart disease) to investigate whether our results were driven by these factors. We also estimated the subdistribution HR, with transient ischemic attack and other types of stroke as competing risks.

RESULTS

We identified 716 patients with newly diagnosed IIM and 7,100 age- and sex-matched general population comparators, among whom 53 (7.4%) and 425 (6%), respectively, were excluded due to prior stroke-related events (see Supplementary Table 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract). A total of 663 IIM patients and 6,673 general population comparators were included in the study population. In both cohorts, 56% were women, and the mean age at the start of follow-up was 61 years. One-third of the IIM patients had DM. The prevalence of stroke risk factors, hypertension, and congestive heart disease at baseline was higher in IIM patients compared to the general pop-

Table 1. Baseline characteristics of the IIM patients and matched general population comparators identified between 2002 and 2011*

	IIM patients (n = 663)	General population (n = 6,673)
Follow-up, median (IQR) years	4.6 (2.6–8.0)	6.0 (3.4–8.9)
Women	369 (56)	3,723 (56)
Age, mean ± SD years	61 ± 15	61 ± 14
Age group, years		
<56	219 (33)	2,211 (33)
56 to <68	215 (32)	2,210 (33)
≥68 to ≤90	229 (35)	2,252 (34)
Education, years		
<10	211 (32)	2,130 (32)
10–12	275 (41)	2,678 (40)
Missing	10 (2)	100 (1)
Diagnosis		
Dermatomyositis	219 (33)	–
Other inflammatory myopathy	444 (67)	–
History of comorbidities		
Diabetes mellitus	35 (5)	256 (4)
Hypertension	98 (15)	580 (9)
Atrial fibrillation	35 (5)	220 (3)
Congestive heart disease	25 (4)	129 (2)

* Values are the number (%) except where indicated otherwise. IIM = idiopathic inflammatory myopathy; IQR = interquartile range.

ulation (Table 1). However, when we adjusted for these factors, the estimates did not change (see Supplementary Table 3, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract).

The median follow-up was 4.6 years in IIM patients, among whom 178 (27%) were censored due to death and 2 (0.3%) due to emigration, compared to 6.0 years in the general population, among whom 641 individuals (10%) were censored due to death and 59 (0.9%) due to emigration.

We identified a total of 34 strokes in IIM patients: 30 ischemic strokes (88%) and 4 hemorrhagic strokes (12%). In the general population, 229 strokes occurred during the study period; 201 (88%) were ischemic strokes and 28 (12%) were hemorrhagic strokes. IIM patients had their first ischemic stroke at a younger age compared to the general population (66 years versus 72 years). The number of hemorrhagic strokes was small, but the age and sex distribution at the time of the first event was similar between IIM patients and the comparators (Table 2).

Table 2. Characteristics of the IIM patients and matched general population comparators at the time of the first ischemic or hemorrhagic stroke*

	IIM patients	General population
Ischemic stroke		
No. of events	30	201
Women	15 (50)	108 (54)
Age at event, mean ± SD years	66 ± 13	72 ± 9
Hemorrhagic stroke		
No. of events	4	28
Women	2 (50)	13 (46)
Age at event, mean ± SD years	72 ± 6	70 ± 12

* Values are the number (%) except where indicated otherwise.

Association between IIM and stroke. For hemorrhagic stroke, the crude incidence rates were 1.1 (95% CI 0.3, 2.9) per 1,000 person-years in IIM patients and 0.7 (95% CI 0.4, 1.0) per 1,000 person-years in the general population. The sex- and age-adjusted rate difference was 0.3 (95% CI –0.6, 1.1) per 1,000 person-years, and the sex-, age-, and place of residence–adjusted HR was 1.9 (95% CI 0.7, 5.5). For ischemic stroke, the crude incidence rate was 8.7 (95% CI 5.9, 12.4) per 1,000 person-years in IIM patients and 4.9 (95% CI 4.2, 5.6) in the general population. The age- and sex-adjusted rate difference was 3.8 (95% CI 1.0, 6.5) per 1,000 person-years, and the HR was 2.1 (95% CI 1.4, 3.0). The incidence rate of ischemic stroke was higher in women with IIM than in women from the general population (7.6 [95% CI 4.3, 12.6] versus 4.6 [95% CI 3.7, 5.5] per 1,000 person-years). Men with IIM were also at higher risk than men from the general population (10.1 [95% CI 5.6, 16.6] versus 5.3 [95% CI 4.3, 6.5] per 1,000 person-years). The rela-

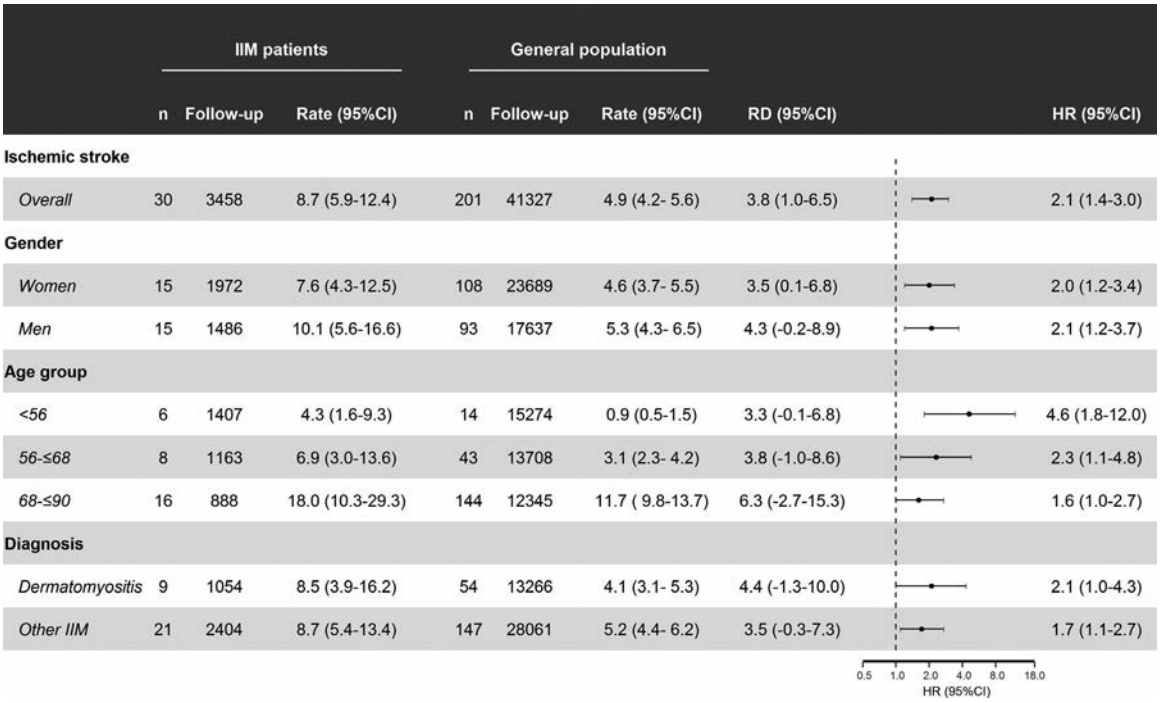


Figure 1. Number of strokes, years of follow-up in person-years, incidence rates, rate differences (RDs), and hazard ratios (HRs) for ischemic stroke overall and stratified by sex, age group, and diagnosis for patients with idiopathic inflammatory myopathies (IIMs) and general population comparators. 95% CI = 95% confidence interval.

tive risk was highest in the youngest age tertile (<56 years) (HR 4.6 [95% CI 1.8, 12.0]), but the age- and sex-adjusted rate difference was highest in the oldest age tertile (≥68 years) (rate difference 6.3 ([95% CI –2.7, 15.3]), respectively. The relative risk was increased for both patients with DM (HR 2.1) and those with other IIMs (HR 1.7) (Figure 1).

When we tested the proportional hazards assumption, the *P* value for the interaction term between exposure and the log of follow-up was <0.01, indicating that the hazards were nonproportional; therefore, the HRs stratified by time since diagnosis might better represent the association between IIM and stroke.

When the association between IIM and ischemic stroke was estimated using a competing risk model with death as a competing event, the subdistribution HR was lower (1.5 [95% CI 1.0, 2.2]) compared to the HR from the cause-specific model. The subdistribution HR for hemorrhagic stroke was also decreased (see Supplementary Table 4, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract). The cumulative incidence of ischemic stroke was almost doubled in IIM patients compared to the general population 1 year after diagnosis (0.9% and 0.4%, respectively) as well as after 5 years (4.0% and 2.2%, respectively) but was more similar after 10 years (5.8% and 4.6%, respectively) (Figure 2).

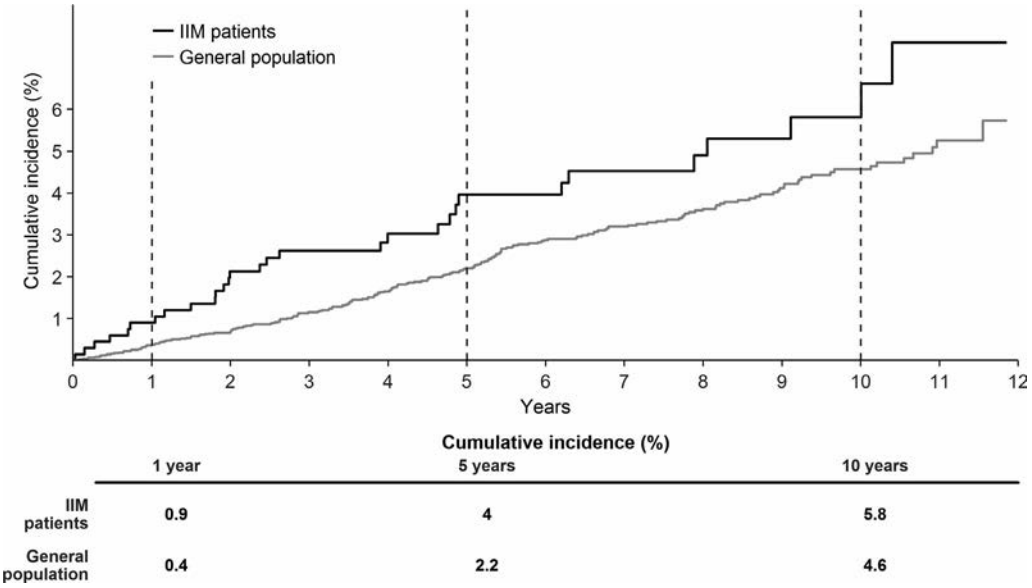
The cumulative incidence of ischemic stroke was small in both IIM patients and in the general population compared to the cumulative incidence of death (see Supplementary Table 5 and Supplementary Figure 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract).

Stroke in relation to time since diagnosis. After stratification by time since the start of follow-up, the relative risk was doubled in IIM patients compared to the general population, for both the first year after and 1–5 years after diagnosis, while a nonsignificant increase was observed after 5–12 years (Figure 3). No effect modification was observed by time since start of follow-up (*P* = 0.62 by likelihood ratio test).

Sensitivity analysis. After adjustment for baseline stroke risk factors, only small differences were observed compared to the primary analysis, and the overall estimated HRs remained the same for both hemorrhagic stroke and ischemic stroke. The largest differences were observed in the youngest age tertile (<56 years), but the HR remained increased (HR 3.5 [95% CI 1.3, 9.7]) (Supplementary Table 3, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract). Adding transient ischemic attack and hemorrhagic stroke as additional competing events to the Fine and Gray model did not alter the subdistribution HR (see Supplementary Table 4, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23702/abstract).

DISCUSSION

In this population-based nationwide cohort study of stroke following IIM, we observed an increased risk of ischemic stroke in IIM patients compared to that in the general population. The relative risk was doubled directly following diagnosis and up to 5 years later, but the difference decreased 5–12 years after diag-



Cumulative incidence (%)	1 year	5 years	10 years
IIM patients	0.9	4	5.8
General population	0.4	2.2	4.6

Figure 2. Cumulative incidence of ischemic stroke in patients with idiopathic inflammatory myopathies (IIMs) and general population comparators, taking into account the competing risk of death. Follow-up values are person-years. The incidence rate per 1,000 person-years and 95% confidence intervals (95% CIs) were estimated using a Poisson distribution. Hazard ratios were estimated by Cox models adjusted for age, sex, and county of residence at the index year. Rate differences were estimated using a Poisson model adjusted for age (tertiles) and sex and were not adjusted for age in the diagnosis subanalyses. n = number of strokes.

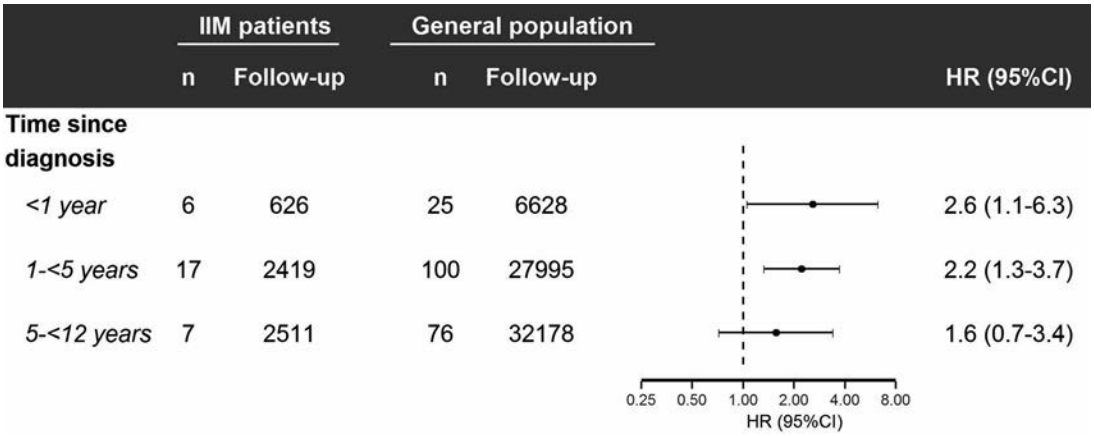


Figure 3. Number of ischemic strokes, follow-up in person-years, and HRs with 95% CIs for ischemic stroke stratified by time since diagnosis in patients with IIM and general population comparators. HRs were estimated using Cox models adjusted for age, sex, and county of residence at the index year. See Figure 1 for definitions.

nosis. The relative risk of ischemic stroke was highest in younger age groups, while the risk difference was highest in the oldest age group. We did observe an HR of 1.9 for hemorrhagic in stroke IIM patients compared to the general population, but due to the very few events, the 95% CI was wide and the estimate nonsignificant.

When we took into account the competing risk of death using the Fine and Gray model, the subdistribution HR for ischemic stroke was decreased compared to the HR in the cause-specific model. A similar difference has been observed previously (5). Cause-specific Cox models can be used to estimate the HR in individuals who have not yet experienced the main event or any competing event. Because it has previously been shown that IIM patients have an almost 4-fold higher mortality compared to the general population (2), and death is a competing event for stroke, the subdistribution HR might better describe the relative risk for these patients in the context of prognosis by also incorporating the association with the competing event. With the subdistribution HR, it is possible to predict the effect that a variable will have on the cumulative incidence function even in the presence of competing risks (18,19). The cumulative incidence of stroke was increased in IIM patients compared to the general population, especially up to 5 years after diagnosis, but at 10 years the incidence was more similar. Because there are few individuals with 10 years of follow-up, this could be attributable to lack of power, and it is difficult to know whether or not the risk is decreasing with time. Previous studies have primarily focused on the HR from cause-specific models, which is of greater interest in etiologic studies (i.e., when investigating casual effects as opposed to studying prognosis and allocating resources for prevention) (20).

Our estimate for ischemic stroke is similar to what was observed in a recent meta-analysis (8) (pooled risk ratio 1.61 [95% CI 1.28, 2.02]) and a recent Canadian study (5) (age- and sex-adjusted HRs of 2.46 [95% CI 1.38, 4.41] and 1.86 [95% CI 0.76, 4.32] for patients with PM and patients with DM, respectively). The majority of previous studies estimating the relative risk of stroke associated with IIM have failed to separate hemorrhagic stroke from ischemic stroke (4–6), and because these types of stroke have different etiologies, it is important to investigate these outcomes separately. Although the incidence of hemorrhagic stroke in our study was too low to make certain conclusions, it indicates an increased risk of hemorrhagic stroke. A novel aspect of our study was the use of a register-based algorithm, which made it possible to identify and separately analyze patients with newly diagnosed IIM in Sweden. We could demonstrate a higher incidence of stroke during the first years after diagnosis compared to that after 5 years. This finding is in contrast to those in other studies, in which both incident and prevalent IIM cases or only hospitalized IIM patients were included (6,7). We could also stratify our results for demographic variables (age, sex) and length of follow-up, which is important if we hope to be able to identify specific risk groups in clinical practice.

The general population comparators and use of prospectively collected and linked data on comorbidity and covariates from public registers with high coverage further increase both the internal and external validity of our findings. The risk of misclassification of the overall outcomes of ischemic stroke and hemorrhagic stroke is relatively low, because both outcomes are conditions that require hospitalization. It is possible that individuals with a minor stroke neglect to seek medical care, but they are found to a higher degree among IIM patients, due to more frequent contact with health care. This type of differential misclassification of outcome would therefore result in an overestimation of the true association. Also, unpublished data from Karolinska Institutet suggest that traumatic intracranial bleeding is sometimes misdiagnosed as hemorrhagic stroke, but this would likely be nondifferential and would have small effects on the estimated risk of hemorrhagic stroke.

A limitation of our study is the overlap of ICD codes used for IBM and PM, and that we have no ICD code to identify the newly identified subset of IIM, immune-mediated necrotizing myopathy.

Because of this limitation, we could analyze only DM separately and not other subgroups of IIM. The fact that we did not have information on the traditionally used classification criteria for different subsets of IIM is another limitation. Further, our aim was to examine whether IIM patients have an increased risk of stroke and to identify groups of IIM patients with the highest risk, and not primarily to identify why IIM patients have an increased risk above and beyond traditional risk factors for stroke; however, the lack of information on stroke risk factors such as hyperlipidemia and smoking is a limitation.

Identification of the underlying mechanism of stroke, in particular ischemic stroke, could be approached in IIM by assessing the risk in relation to traditional risk factors and other immunologic features (such as autoantibodies). Inflammatory markers such as C-reactive protein have previously been linked to an increased risk of stroke and could partly explain the increased risk of stroke in other rheumatic diseases such as systemic lupus erythematosus and rheumatoid arthritis (21). In IIM, however, the disease activity and main inflammatory marker used in clinical practice is serum levels of creatinine kinase, and we know very little about the role of creatinine kinase in the pathogenesis of cardiovascular diseases, including stroke.

In conclusion, our findings indicate that patients with IIMs have an increased risk of both hemorrhagic stroke and ischemic stroke. It should be kept in mind that even if the risk is elevated, stroke is still a rare event. Therefore, focus on prevention should be directed toward the groups with the highest absolute risk.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Lundberg 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. Moshtaghi-Svensson, Von Euler, Arkema, Holmqvist

Acquisition of data. Moshtaghi-Svensson, Lundberg, Holmqvist

Analysis and interpretation of data. Moshtaghi-Svensson, Lundberg, Von Euler, Arkema, Holmqvist

REFERENCES

1. Dalakas MC. Polymyositis, dermatomyositis, and inclusion-body myositis. N Engl J Med 1991;325:1487–98.
2. Dobloug GC, Svensson J, Lundberg IE, Holmqvist M. Mortality in idiopathic inflammatory myopathy: results from a Swedish nationwide population-based cohort study. Ann Rheum Dis 2018;77:40–7.
3. Lundberg I. Cardiac involvement in autoimmune myositis and mixed connective tissue disease. Lupus 2005;14:708–12.
4. Lai YT, Dai YS, Yen MF, Chen LS, Chen HH, Cooper R, et al. Dermatomyositis is associated with an increased risk of cardiovascular and cerebrovascular events: a Taiwanese population-based longitudinal follow-up study. Br J Dermatol 2013;168:1054–9.
5. Rai SK, Choi HK, Sayre EC, Aviña-Zubieta JA. Risk of myocardial infarction and ischaemic stroke in adults with polymyositis and dermatomyositis: a general population-based study. Rheumatology (Oxford) 2016;55:461–9.
6. Tisseverasinghe A, Bernatsky S, Pineau CA. Arterial events in persons with dermatomyositis and polymyositis. J Rheumatol 2009;36:1943–6.
7. Zöller B, Li X, Sundquist J, Sundquist K. Risk of subsequent ischemic and hemorrhagic stroke in patients hospitalized for immune-mediated diseases: a nationwide follow-up study from Sweden. BMC Neurol 2012;12:41.
8. Ungprasert P, Cheungpasitporn W, Wijarnpreecha K, Ahuja W, Ratanasrimetha P, Thongprayoon C. Risk of ischemic stroke in patients with polymyositis and dermatomyositis: a systematic review and meta-analysis. Rheumatol Int 2015;35:905–9.
9. National Board of Health and Welfare. The National Patient Register. URL: http://www.socialstyrelsen.se/register/halsodataregister/patientregistret/inenglish.
10. Svensson J, Holmqvist M, Tjärnlund A, Dastmalchi M, Hanna B, Magnusson BS, et al. Use of biologic agents in idiopathic inflammatory myopathies in Sweden: a descriptive study of real life treatment. Clin Exp Rheumatol 2017;35:512–5.
11. Svensson J, Arkema EV, Lundberg IE, Holmqvist M. Incidence and prevalence of idiopathic inflammatory myopathies in Sweden: a nationwide population-based study. Rheumatology (Oxford) 2017;56:802–10.
12. Statistics Sweden. Registret över Totalbefolkningen. URL: https://www.scb.se/sv_/Vara-tjanster/Bestalla-mikrodata/Vilka-mikrodata-finns/Registret-over-totalbefolkningen-RTB/.
13. National Board of Health and Welfare. Dödsorsaksregistret. URL: http://www.socialstyrelsen.se/register/dodsorsaksregistret.
14. Gilbert DT, Morgan O, Smikle M, Simeon D, Barton EN. HTLV-1 associated polymyositis in Jamaica. Acta Neurol Scand 2001;104:101–4.
15. Statistics Sweden. Longitudinal integration database for health insurance and labour market studies. URL: http://www.scb.se/en_/Services/Guidance-for-researchers-and-universities/SCB-Data/Longitudinal-integration-database-for-health-insurance-and-labour-market-studies-LISA-by-Swedish-acronym/.
16. Fay MP, Feuer EJ. Confidence intervals for directly standardized rates: a method based on the gamma distribution. Stat Med 1997;16:791–801.
17. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Associ 1999;94:496–509.
18. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. Am J Epidemiol 2009;170:244–56.
19. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. Stat Med 2017;36:4391–400.
20. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. Circulation 2016;133:601–9.
21. Kuo HK, Yen CJ, Chang CH, Kuo CK, Chen JH, Sorond F. Relation of C-reactive protein to stroke, cognitive disorders, and depression in the general population: systematic review and meta-analysis. Lancet Neurol 2005;4:371–80.

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 977–985
DOI 10.1002/acr.23746

American College of Rheumatology
Empowering Rheumatology Professionals

Novel Ultrasound Image Acquisition Protocol and Scoring System for the Pediatric Knee

Tracy V. Ting,[1] Patricia Vega-Fernandez,[2] Edward J. Oberle,[3] Deirdre De Ranieri,[4] Hulya Bukulmez,[5] Clara Lin,[6] David Moser,[7] Nicholas J. Barrowman,[8] Yongdong Zhao,[9] Heather M. Benham,[10] Laura Tasan,[11] Akaluck Thatayatikom,[12] Johannes Roth,[13] and the Childhood Arthritis and Rheumatology Research Alliance Juvenile Idiopathic Arthritis Ultrasound Workgroup

Objective. The use of musculoskeletal ultrasound is increasing among pediatric rheumatologists. Reliable scoring systems are needed for the objective assessment of synovitis. The aims of this study were to create a standardized and reproducible image acquisition protocol for B-mode and Doppler ultrasound of the pediatric knee, and to develop a standardized scoring system and determine its reliability for pediatric knee synovitis.

Methods. Six pediatric rheumatologists developed a set of standard views for knee assessment in children with juvenile arthritis. Subsequently, a comprehensive literature review, practical exercises, and a consensus process were performed. A scoring system for both B-mode and Doppler was then developed and assessed for reliability. Interreader reliability or agreement among a total of 16 raters was determined using 2-way single-score intraclass correlation coefficient (ICC) analysis.

Results. Twenty-one views to assess knee arthritis were initially identified. Following completion of practical exercises and subsequent consensus processes, 3 views in both B-mode and Doppler were selected: suprapatellar longitudinal and medial/lateral parapatellar transverse views. Several rounds of scoring and modifications resulted in a final ICC of suprapatellar view B-mode 0.89 (95% confidence interval [95% CI] 0.86–0.92) and Doppler 0.55 (95% CI 0.41–0.69), medial parapatellar view B-mode 0.76 (95% CI 0.68–0.83) and Doppler 0.75 (95% CI 0.66–0.83), and lateral parapatellar view B-mode 0.82 (95% CI 0.75–0.88) and Doppler 0.76 (95% CI 0.66–0.84).

Conclusion. A novel B-mode and Doppler image acquisition and scoring system for assessing synovitis in the pediatric knee was successfully developed through practical exercises and a consensus process. Study results demonstrate overall good-to-excellent reliability.

INTRODUCTION

Musculoskeletal ultrasound (MSUS) is a noninvasive, efficient modality for the assessment of inflammatory arthritis (1–3). In adult rheumatology, the clinical and research use of MSUS has become a routine point-of-care procedure; however, there are a limited number of studies of MSUS in juvenile idiopathic arthritis (JIA) and pediatric rheumatology (4,5).

Traditionally, the diagnosis of arthritis is made by clinical examination demonstrating swelling, pain, and/or limitation in range of motion of the joint. However, these findings may not be specific to synovitis, because trauma, infection, and pain amplification are other etiologies that can present similarly. Furthermore, in many children with synovitis, these findings may not be elicited on physical examination. Radiographic evaluation primarily assesses the bone and often shows only late manifestations of inflammatory

Supported by the Childhood Arthritis and Rheumatology Research Alliance, which is supported by the Arthritis Foundation. Dr. Zhao's work was supported by the Clinical Research Scholar Program (Seattle Children's Research Institute).

[1]Tracy V. Ting, MD, MSc, RhMSUS: Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; [2]Patricia Vega-Fernandez, MD, MSc, RhMSUS: Emory University School of Medicine and the Children's Hospital of Atlanta, Atlanta, Georgia; [3]Edward J. Oberle, MD, RhMSUS: Nationwide Children's Hospital, Columbus, Ohio; [4]Deirdre De Ranieri, MD, RhMSUS: Ann and Robert H. Lurie Children's Hospital of Chicago, Chicago, Illinois; [5]Hulya Bukulmez, MD: MetroHealth Medical Center, Cleveland, Ohio; [6]Clara Lin, MD, RhMSUS: Children's Hospital Colorado and University of Colorado Denver; [7]David Moser, DO, RhMSUS: Cigna, Chicago, Illinois; [8]Nicholas J. Barrowman, PhD: University of Ottawa, Ottawa, Ontario, Canada; [9]Yongdong Zhao, MD, PhD, RhMSUS: Seattle Children's Hospital, Seattle, Washington; [10]Heather M. Benham, DNP, RN, CPNP: Texas Scottish Rite Hospital for Children, Dallas, Texas; [11]Laura Tasan, MD: University of Pittsburgh Medical Center Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania; [12]Akaluck Thatayatikom, MD, RhMSUS: University of Florida, Gainesville; [13]Johannes Roth, MD, PhD, FRCPC, RhMSUS: Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada.

Dr. Zhao has received research support from Bristol-Myers Squibb. No other disclosures relevant to this article were reported.

Address correspondence to Tracy V. Ting, MD, MSc, RhMSUS, 3333 Burnet Avenue, MLC 4010, Cincinnati OH 45229. E-mail: tracy.ting@cchmc.org.

Submitted for publication May 16, 2018; accepted in revised form September 4, 2018.

SIGNIFICANCE & INNOVATIONS

- A comprehensive set of images incorporating views that were not previously included in the assessment of arthritis in the pediatric knee was standardized through consensus.
- A novel scoring system for the assessment of synovitis specific to the pediatric knee was developed.
- Specific scoring systems for individual joints will increase reliability and ultimately lead to time-efficient application of these tools in clinical practice and research.

processes. Magnetic resonance imaging has several limitations, including the use of nephrotoxic contrast agents to enhance synovial lining, the cost, and the frequent use of sedation for young children. Ultrasonography has significant advantages over other imaging technologies due to its cost-effectiveness, accessibility as a point-of-care tool, and ability to evaluate children without needing intravenous contrast or sedation. The use of ultrasonography to evaluate arthritis in adults and children with rheumatic diseases can guide both diagnosis and treatment (5–11). Thus, developing standards for pediatric MSUS is imperative.

Due to the specific anatomy of a growing child, adult imaging standards do not clearly apply to pediatric patients. Normal growth and bony ossification vary by age, maturity, and sex (12), and all compartments in the joint of a growing child may demonstrate vascularity by Doppler (13–15). Furthermore, both evidence-based recommendations on specific imaging protocols as well as reliable scoring systems for assessing arthritis among children are limited (11). Previously presented scoring definitions do not clearly apply to referenced illustrations (11), nor do general descriptions adequately apply to all joints and all views (16). Publications focusing specifically on sonography of the knee joint often assess only the suprapatellar recess and do not include the parapatellar recesses (1,17), which provide additional value in determining synovitis (18). The parapatellar recesses may be especially valuable in assessing synovitis, because they are generally more superficial, which may increase sensitivity compared to the deeper suprapatellar recess (8,19–22).

Therefore, the objectives of this study were to establish a comprehensive image acquisition protocol for the pediatric knee and assess the feasibility of performing these views, and to develop a reliable, standardized scoring system for the assessment of knee arthritis in B-mode and Doppler.

MATERIALS AND METHODS

Development of an image acquisition protocol. A comprehensive literature review was performed using the terms "knee, arthritis, synovitis, ultrasound, imaging, and pediatric." Six pediatric rheumatologists trained in MSUS (6–10 years) reviewed the literature summarizing key points relative to the assessment of knee arthritis, including the most common sites of effusion, synovitis, and positioning of the knee. Via 3 teleconferences and at the Childhood Arthritis and Rheumatology Research Alliance (CARRA) JIA Ultrasound subgroup meeting in Toronto, Canada (2016), several discussions occurred regarding the key findings from the literature review. From these meetings, an initial protocol of recommended views for the knee was developed following 100% consensus among the 6 pediatric rheumatologists.

Feasibility was assessed via a practical exercise carried out in Cincinnati, Ohio, with 6 pediatric rheumatologists trained in MSUS (2–10 years). The study was approved by the Cincinnati Children's Hospital Medical Center Institutional Review Board and patients recruited from the clinic provided written assent and consent to participate. During this exercise, 3 patients with JIA (ages 11–14 years) underwent the initial imaging protocol. Following this first exercise, certain views were eliminated. The second exercise involved scanning 6 additional patients with JIA (ages 8–19 years) with the revised version of the imaging protocol. Participants were seated with their examined knee flexed at 30 degrees and were either in a relaxed position or instructed to contract their quadriceps muscles. Images were viewed in both B-mode and power Doppler. Clear definitions were provided for the bony anatomic landmarks as well as for the soft tissues to ensure standardized image acquisition. In the initial exercise, both still images and videos were acquired. Video clips (in both B-mode and Doppler) were obtained by scanning across the area of interest in order to determine the area of maximal pathology. We also scanned each region prior to the acquisition of the still images to ensure the capture of both the maximum distension of the synovial recess and the maximum number of Doppler signals. B-mode settings included a frequency range of 9–15 MHz (depending on the body size of the participant), and Doppler was measured with low flow settings, including a pulse repetition frequency <1.0 (typically between 0.4 and 0.6), low wall filter, and frequency adjusted to obtain maximum sensitivity as well as gain set to just below artifact levels. All images underwent an initial quality assessment, including the visualization of a thin layer of gel to avoid extensive pressure as well as the presence and good depiction of anatomic landmarks. The images were then used to evaluate the feasibility of identifying pathology, i.e., synovitis. Consensus agreement among the 6 scanners led to the development of the final recommended image acquisition protocol. Figure 1 shows the various anatomic structures included in the scanning protocol.

Development of a scoring system. An additional literature review was performed using the terms "knee, ultrasound, arthritis, synovitis, scoring, assessment, and B-mode, Doppler" in order to understand the ultrasound models that were cur-

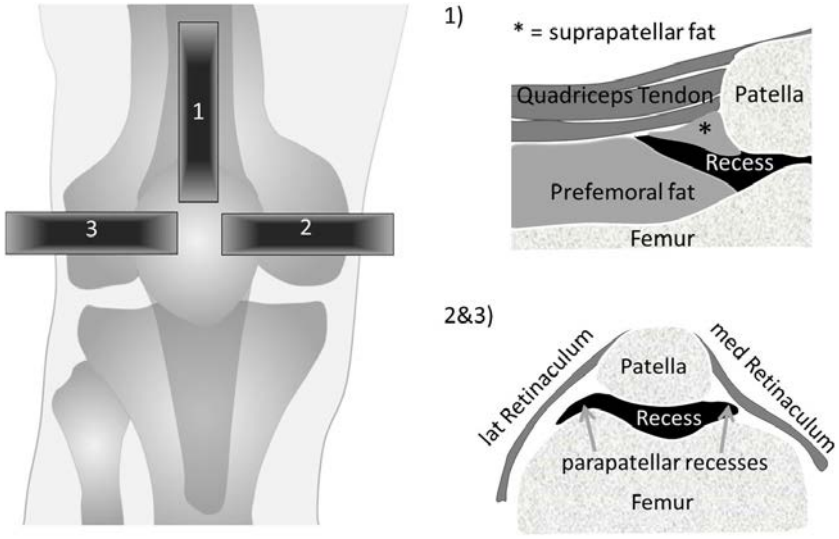


Figure 1. Schematic of the anatomic structures of the scanning protocol. The probe position for the suprapatellar (**1**), medial parapatellar (**2**), and lateral parapatellar (**3**) scan is shown on the left with the corresponding schematic depiction of the visualized anatomic structures on the right. lat = lateral; med = medial.

rently being used for scoring. Analysis of these models and applicability to pediatric arthritis were discussed by the authors, and consensus was reached regarding a preliminary system for both B-mode and Doppler, specifically for the images recommended in the final version of the imaging protocol. The B-mode scoring system was built upon a previously established scoring system for hemophilia (21) that was modified and expanded. The Doppler scoring system follows other adult and pediatric-specific scoring systems and definitions (11,23–25). However, these systems were not all developed for or assessed in the pediatric knee with inflammatory arthritis, and most do not include the parapatellar recesses, which provide an important sonographic window for the detection of pathology.

Three scoring exercises of B-mode and Doppler were performed on images of the knee, evaluating the suprapatellar, medial parapatellar, and lateral parapatellar recesses. These images were selected ensuring equal distribution across ages 2–18 years, and both sexes, as well as fulfilling the quality criteria of appropriate machine settings, clear visibility of bone contours, a layer of gel indicating the absence of compression, and the absence of Doppler artifacts. Additional calibration of the system was assessed during the CARRA JIA Ultrasound subgroup meeting in Houston, Texas (2017), where 16 members of the subgroup (of variable ultrasound experience, from <1 year to 10 years) also participated in scoring exercises. Refinement of the scoring system occurred after each exercise following discussions to clarify definitions and improve the scoring consensus. Definitions were sent for the next round of scoring after the authors agreed with all aspects.

Images (n = 654) were scored individually by each scanner and then analyzed following each exercise. To measure interreader reliability, or agreement of scoring between multiple raters, the 2-way single-score intraclass correlation coefficient (ICC) method with 95% confidence interval (95% CI) was used. Agreement was also assessed separately for the 6 authors with longer experience in the use of MSUS and for members of the CARRA JIA Ultrasound subgroup with less experience. ICC is a commonly used measure of interreader reliability for variables scored by multiple raters (26). An excellent ICC was considered to be 0.75–1.00, good 0.60–0.74, fair 0.40–0.59, and poor <0.4 (27).

RESULTS

Image acquisition protocol. A total of 21 views of the knee were obtained during the first iteration of the protocol, including static images as well as dynamic video clips sweeping across a predetermined area (Table 1). Standard patient positioning was included, with the knee flexed to 30 degrees, and images were obtained with and without contraction of the quadriceps muscle. Feasibility was assessed during the first practical exercise. Regardless of experience level, the entire imaging protocol was completed in approximately 10 minutes per scanner.

A consensus meeting among the authors/scanners was conducted following the initial exercise. The group decided to eliminate views requiring quadriceps muscle contraction, given the challenges in interpreting both B-mode findings and Doppler due to motion artifact and the variability of the participants (ages 8–19 years) to perform contraction adequately. Some of the participants verbally indicated confusion regarding the technical process of contraction, and others fatigued easily during contraction, rendering image acquisition unreliable. Moreover, the location and

Table 1. Initial and final proposed ultrasound views for the pediatric knee*

	B-mode	CPD
Initial protocol		
Midline suprapatellar	1	2 (v)
Midline suprapatellar, with contracture	3	4 (v)
Sweep medial to lateral	5 (v)	6 (v)
Medial parapatellar, with contracture	7	8 (v)
Medial parapatellar, no contracture	9	10 (v)
Medial parapatellar, sweep proximal to distal	11 (v)	12 (v)
Lateral parapatellar, with contracture	14	15 (v)
Lateral parapatellar, no contracture	16	17 (v)
Lateral parapatellar, sweep proximal to distal	18 (v)	19 (v)
Cartilage in maximum knee flexion	20	21 (v)
Final protocol†		
Midline suprapatellar longitudinal	1	2
Medial parapatellar transverse	3	4
Lateral parapatellar transverse	5	6

* Values are the image number. CPD = color power Doppler; v = video; sweep = scanning across a region.
† For the final protocol, all images were performed with the knee flexed at 30°. The knee was flexed and extended 3 times prior to scanning. Proximal third of the patella must be in view for suprapatellar images. Patella and femur must be in view for parapatellar images. In each location, the probe was moved to obtain a view with maximal distension of synovial recess or Doppler signals, as long as the defined bony landmarks were still visible.

geometric dimensions of effusion were not reliable from scanner 1 to scanner 6, indicating likely variability in effort and/or contraction technique (data not shown). We noted that involuntary movements or changes in the position of a limb also affected visibility of fluid in the suprapatellar recess, irrespective of active contraction. In order to avoid variations due to involuntary contractions or other movements of the joint and subsequent redistribution of the synovial fluid, we then decided that contraction (simulated by 3 full flexions/extensions of the knee) should occur prior to scanning, which was easier to understand and could be performed by the patient or scanner. Finally, videos of dynamic views were eliminated because they did not clearly demonstrate differences in synovial recess distensions or presence of Doppler signals from those already noted on static views (data not shown). Furthermore, the presence of Doppler signal artifact was significant due to dynamic motion while scanning across a recess.

Based on the evaluation of images obtained during this exercise, a final image acquisition protocol was developed (Table 1). Indeed, the combination of static midline suprapatellar and dynamic parapatellar (medial and lateral) views was sufficient to capture disease findings. Unlike previous suggestions for parapatellar assessment (16), we elected to place the probe in the transverse view in the midpatellar portion as demonstrated in prior studies (21), because the clear identification of the patella (either fully ossified or partially ossified) as one of the landmarks is more reliable in this location. To further ensure accurate assessment in the parapatellar views, the probe was moved proximally and distally to capture the maximum pathology. Figure 1 shows schematic depictions of the 3 views.

Scoring system. Overall, a semiquantitative scoring system ranging from 0 (normal) to 3 (severe) was developed for both the B-mode and Doppler systems. The B-mode system was adapted from an existing scoring system for hemophilia (21). The Doppler scoring system was applied to Doppler signals within the synovial recess and synovial hypertrophy only (25). Physiologic Doppler signals, such as feeding vessels, were excluded from the scoring system (25,28,29).

In total, 3 scoring exercises (Table 2) were performed by the authors (n = 6). The second and third scoring exercises were also completed by several other members of the CARRA JIA Ultrasound subgroup (10 members for the second and 4 for the third exercise). Exercise 1 included a total of 126 B-mode images and 83 Doppler images. Interreader reliability as determined by ICC analysis revealed fair agreement among the majority of images but excellent agreement for suprapatellar B-mode (0.78). Medial and lateral parapatellar B-mode were 0.52 and 0.60, respectively. Doppler for all 3 views had fair reliability: 0.39, 0.57, and 0.54, respectively.

For calibration of the system, a second scoring exercise was performed in Houston, Texas, on day 2 of the CARRA meeting, with a total of 16 participants (authors plus CARRA JIA Ultrasound subgroup members). Day 1 involved discussion of the process and a review of the initial scoring system/atlas. The scoring exercise (90 total images) revealed fair to good reliability with B-mode scoring for suprapatellar, medial parapatellar, and lateral parapatellar views at 0.72 for CARRA subgroup members (0.82 for authors only), 0.56 (0.53 for authors), and 0.44 (0.52 for authors), respectively. Doppler scoring, however, was variable: 0.24 (0.50 for authors), 0.38 (0.24 for authors), 0.68 (0.64 for authors), respectively.

Discussions following the 2 exercises and during the calibration exercise revealed limitations of the scoring system in a few areas: 1) the differentiation of grades 2 and 3 in B-mode in the suprapatellar recess, which led to the introduction of additional measures compared to the previously published hemophilia scoring system (21); 2) the grading according to quartile

Table 2. Intraclass correlation coefficient with 95% confidence interval from scoring exercises by authors and Childhood Arthritis and Rheumatology Research Alliance Juvenile Idiopathic Arthritis Ultrasound members*

	Exercise 1		Exercise 2		Exercise 3	
Images	B-mode (n = 136)	CPD (n = 83)	B-mode (n = 55)	CPD (n = 35)	B-mode (n = 208)	CPD (n = 137)
All views						
Authors	0.66 (0.58–0.73)	0.53 (0.37–0.67)	0.67 (0.57–0.77)	0.48 (0.33–0.64)	0.84 (0.81–0.86)	0.75 (0.69–0.81)
All	–	–	0.61 (0.52–0.71)	0.48 (0.36–0.63)	0.78 (0.75–0.81)	0.64 (0.56–0.72)
Suprapatellar						
Authors	0.78 (0.70–0.85)	0.39 (0.22–0.58)	0.82 (0.7–0.91)	0.50 (0.25–0.78)	0.89 (0.86–0.92)	0.55 (0.41–0.69)
All	–	–	0.72 (0.58–0.85)	0.24 (0.09–0.52)	0.83 (0.78–0.83)	0.55 (0.43–0.69)
Medial parapatellar						
Authors	0.52 (0.37–0.67)	0.57 (0.38–0.75)	0.53 (0.32–0.75)	0.24 (0.07–0.52)	0.76 (0.68–0.83)	0.75 (0.66–0.83)
All	–	–	0.56 (0.39–0.77)	0.38 (0.20–0.64)	0.67 (0.58–0.76)	0.59 (0.46–0.71)
Lateral parapatellar						
Authors	0.60 (0.47–0.73)	0.54 (0.32–0.73)	0.52 (0.34–0.72)	0.64 (0.37–0.87)	0.82 (0.75–0.88)	0.76 (0.66–0.84)
All	–	–	0.44 (0.29–0.64)	0.68 (0.48–0.88)	0.79 (0.72–0.85)	0.66 (0.54–0.77)

* Authors were experienced sonographers; All includes group members with variable levels of expertise. CPD = color power Doppler.

percentages (<25% or more, etc.) for the parapatellar recess proved to be less reliable, and therefore a process to divide the recess into thirds was introduced; 3) the need for a clear definition of the outline of the normal parapatellar recesses; 4) reviewing the area in which Doppler signals are relevant for scoring, particularly parapatellar recesses; 5) clarifying Doppler artifacts; and 6) avoidance of overgrading (i.e., giving a higher score due to artifact or misinterpretation).

Finally, a third scoring exercise among the authors and 4 participants of the CARRA JIA Ultrasound subgroup was performed, with improvements to the atlas as noted above, including additional clarification of the normal synovial recess in the parapatellar views (see Supplementary Figure 1 [description in Supplementary Appendix 1], available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23746/abstract). Additionally in this third exercise, participants were asked to also score using thirds in lieu of percentages, to see whether this practice would significantly change scores or improve the ease of scoring.

Interreader reliability results from the third exercise (total images n = 345) were much improved (Table 2): suprapatellar view B-mode ICC 0.89 (95% CI 0.86–0.92) and Doppler 0.55 (95% CI 0.41–0.69), medial parapatellar view B-mode 0.76 (95% CI 0.68–0.83) and Doppler 0.75 (95% CI 0.66–0.83), and lateral parapatellar view B-mode 0.82 (95% CI 0.75–0.88) and Doppler 0.76 (95% CI 0.66–0.84). Agreement did not significantly change between percentages versus thirds for the B-mode assessment of parapatellar images; however, we felt that thirds provided a lower risk of overgrading and was conceptually easier to apply. Therefore, the consensus decision was to use thirds as the parameter to differentiate the various grades.

B-mode suprapatellar images. Scoring of B-mode images for the suprapatellar view was based on grading scores of 0 (normal) to 3 (severe). The knee should be flexed at 30 degrees and images collected after the patient completes flexion and extension 3 times. Schematic illustrations and ultrasound images are shown in Figure 2 and in Supplementary Figure 2 (description in Supplementary Appendix 1), available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23746/abstract. Longitudinal images of the suprapatellar joint space should include the proximal third of the patella and a clearly visualized quadriceps tendon. Key variables of assessment include fat pad elevation, presence and degree of extension of effusion (E), and/or synovial hypertrophy (SH). Due to physiologic amounts of fluid, a normal knee or grade 0 allows for a slit of fluid/synovium without elevation of the prepatellar fat pad but with only minimal extension beyond the prepatellar fat pad. Mild or grade 1 findings include minimal E/SH with elevation of the prepatellar fat pad and extension proximally <50% of the visualized portion of the quadriceps tendon. Grade 2 findings reveal a moderate E/SH elevating the prepatellar fat pad with extension proximally >50% of the visualized portion of the quadriceps tendon. Last, a marked E/SH is considered grade 3 if there is significant distension of the suprapatellar recess between the undersurface of the quadriceps tendon and the prefemoral fat pad, extending throughout the image, and with

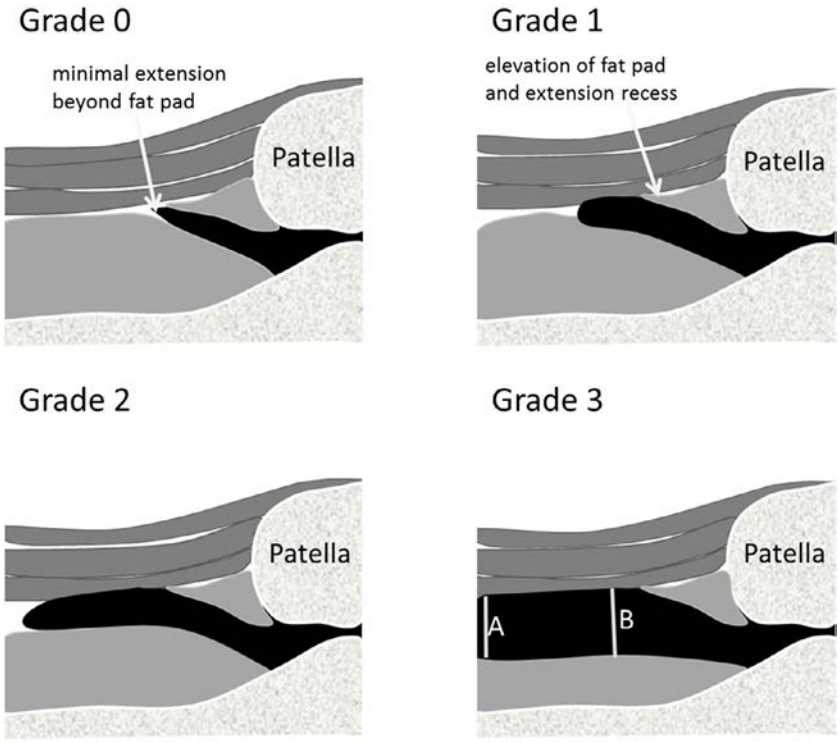


Figure 2. Sonographic scoring system for the suprapatellar recess. The scoring system in B-mode for the suprapatellar recess is shown, further detailed in the article. For Grade 3 the length of line A has to be at least 50% of the length of line B. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23746/abstract.

the most proximal portion of the synovial recess being >50% of the maximum distension of the recess.

B-mode parapatellar images. Images of the parapatellar gutters were obtained with the probe in transverse position over the midpatella (or the area of greatest distension) with both the patella and femur in view. Schematic illustrations as well as corresponding ultrasound images are shown in Figure 3 and in Supplementary Figures 3 and 4 (description in Supplementary Appendix 1), available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23746/abstract. Parapatellar normal knee (grade 0) indicates an empty parapatellar recess, but a minimal bulge of E/SH may be found extending to the patellofemoral joint line. Mild findings (grade 1) include the presence of E/SH filling less than one-third of the full area of the parapatellar recess. Moderate findings (grade 2) have E/SH filling between one- and two-thirds of the full area of the parapatellar recess. Finally, severe findings (grade 3) show E/SH that fills greater than two-thirds of the full area of the parapatellar recess.

Doppler images. The Doppler box should include the full recess area and extend to the top of the screen. Signals should only be considered if located within the area of SH in the recess. One should also be aware of normal feeding vessels in developing children. Normal Doppler (grade 0) shows the presence of no signal. Grade 1 includes 1–3 signals within the area of SH only. Grade 2 should show >3 signals or confluent signals present in <50% of the area of SH. Finally, significant Doppler signal or grade 3 is scored when confluent signals are present in >50% of the area of SH. Notably, the area in relation to which this percentage is calculated is strictly determined only within the area of synovial proliferation. The overall pathology of the synovial recess might be larger; however, if Doppler signals were to be calculated relative to the entire synovial recess (i.e., including a synovial effusion) a lower grade might result. An illustration of the definitions is given in Supplementary Figures 5, 6, and 7 (description in Supplementary Appendix 1), available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23746/abstract.

DISCUSSION

We propose a comprehensive, reliable, and quick MSUS scanning protocol with a well-defined scoring system of the knee in JIA. To our knowledge, this is the first study to present a scanning and scoring system specifically for the assessment of arthritis in the pediatric knee including suprapatellar and parapatellar views. Previous studies have been focused on adult patients (28–31), have only assessed the suprapatellar recess, or have adopted a single standard scoring system for all joints (11,16). However, these previous publications have shown the challenges of a sin-

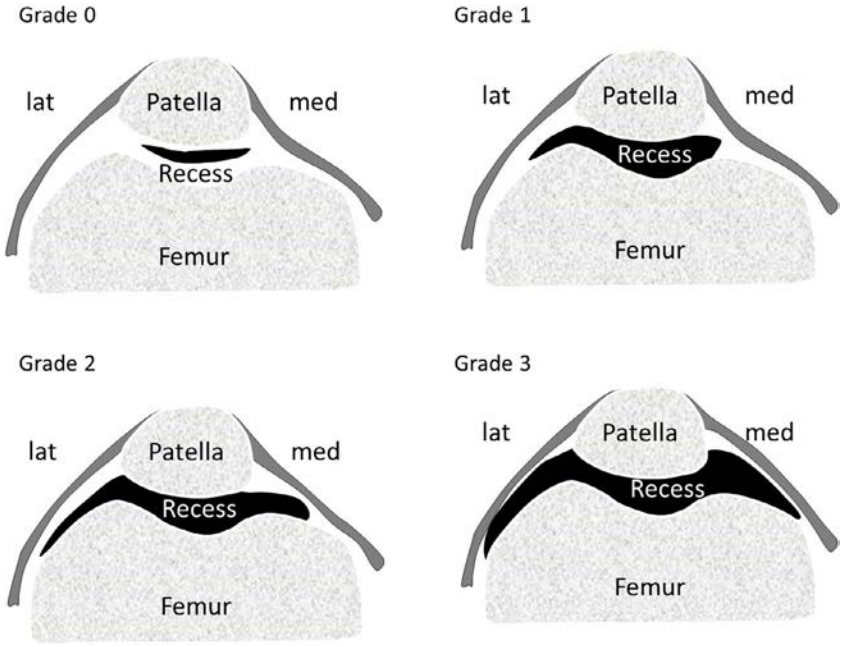


Figure 3. Sonographic scoring system for the parapatellar recesses. The scoring system in B-node for the parapatellar recesses is shown, further detailed in the article. The structure in dark grey attaching at each side of the patella is the medial and lateral retinaculum. lat = lateral; med = medial. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23746/abstract.

gle standard scoring system to adequately capture the various views of a given joint (11). Our detailed descriptions of each view will likely provide a more comprehensive understanding of pathology and disease activity important for both clinical and research use. Despite the need for multiple views of the knee, this imaging protocol can be completed at the bedside within 10 minutes. By electing to reference visible portions of the patella and quadriceps tendon, we obviated the need for a specific probe size when acquiring images. Although videos have been used in some adult scoring studies (28,32), we were able to capture a wide area of synovitis and maximal Doppler signal using still images alone. This restriction also eliminated the challenge of interpreting movement artifact, and it increased scanning efficiency, because less time was required for image acquisition. We evaluated the protocol with power Doppler images, but its use with color Doppler is not precluded, because both techniques are equally sensitive when settings are optimized (33). Furthermore, the choice among the 2 modalities may depend on the machine and individual preferences of the examiner.

Our scanning and scoring protocols proved to be highly feasible, even among pediatric providers with variable levels of experience in MSUS. Reliability increased with very clear, detailed, and specific definitions for each suprapatellar and parapatellar view, further underlining the need for joint-specific scoring systems. These specific definitions are particularly necessary in pediatrics, given normal variation in anatomy (variable degrees of ossification) and physiologic ultrasound findings (Doppler signals in particular). Our views and scoring systems can be applied to all ages across the pediatric spectrum, because the soft tissue characteristics are the same, the bony landmarks, whether fully or partially ossified, can be clearly identified, and pathologic Doppler signals can be differentiated from physiologic signals independent of age (as illustrated in the supplementary figures of the various scores).

Certain components of the knee JIA–specific scoring system bear further discussion. First, the B-mode scoring system was adapted from a system designed for hemophilia, a unique disease with potentially different findings than those seen in JIA. When assessing the various options in our practical exercises, we nevertheless concluded that the basic principle behind the scoring system was relevant to our JIA patient population as well. The system appears to distinguish findings in patients with JIA, although modifications were needed, especially for grades 2 and 3. Second, one aspect that is not being addressed consistently in clinical practice and research is the interpretation of Doppler signals in an area of synovial hypertrophy only and the calculation of the percentage of involvement relative to this area. In particular, for the suprapatellar view, a grade 3 Doppler signal would be impossible to obtain in most cases when referencing the entire synovial recess, because the synovial proliferation often covers only a portion of it. Thus, inaccurate quantification of Doppler signals could result in difficulty documenting change over time. This problem has led some authors to suggest alternative scoring systems (34). Thus, for the knee, our scoring system also captures Doppler signals in the more superficial parapatellar recesses. Last, though physiologic blood flow may be more prominent in younger children, our detailed scoring system carefully differentiates normal from pathologic flow signals.

One limitation in this study is that the investigators participated in several of the scoring exercises, continuously interacted with each other, and had discussions on making improvements to the scoring protocol. The improvements in overall interrater reliability may have occurred due to increased attention, interaction of the investigators, and continuous learning throughout the study. However, different sets of images were used for each round of scoring. Furthermore, participants from the CARRA JIA Ultrasound subgroup also showed improvements from exercise 2 to 3 with a more detailed scoring atlas.

The scanning protocol and scoring atlas presented here offer a reliable, simple, and quick tool to complement our clinical examination of the pediatric knee with arthritis. Developing a scanning protocol and scoring atlas is one of the first steps to systematically using MSUS as a clinical and research tool. The scoring system proposed here is proven to have a good-to-excellent interrater reliability. Further work will need to focus on the prospective clinical application, including correlation with clinical findings, responsiveness to change, and prediction of outcomes. Furthermore, developing specific scoring systems for the remainder of the joints will be an important next step in the process of developing a comprehensive framework for objectively evaluating disease activity in JIA patients.

ACKNOWLEDGMENTS

The authors thank Gurjit S. Kaeley for his help and advice with the development of the scoring system. All figures in this article were designed by Johannes Roth according to original figures by Carlo Martinoli. The authors also thank specific members of the CARRA JIA Ultrasound subground for their meaningful contributions: Meghan Aabo, William Bernal, Kathleen Jo Corbin, Austin Dalrymple, Sarah Hoffman, M. Jamie Holland, Ginger Janow, Monica Marcus, Anjali Patwardhan, Susan Shenoi, Chelsey Smith, Sheetal Vora, Pamela Weiss, and Leandra Woolnough.

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. Ting 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. Ting, Vega-Fernandez, Oberle, De Ranieri, Bukulmez, Roth.

Acquisition of data. Ting, Vega-Fernandez, Oberle, De Ranieri, Bukulmez, Lin, Moser, Zhao, Benham, Tasan, Thatayatikom, Roth.

Analysis and interpretation of data. Ting, Vega-Fernandez, Oberle, De Ranieri, Bukulmez, Barrowman, Roth.

REFERENCES

1. Möller I, Janta I, Backhaus M, Ohrndorf S, Bong DA, Martinoli C, et al. The 2017 EULAR standardised procedures for ultrasound imaging in rheumatology. Ann Rheum Dis 2017;76:1974–9.
2. Terslev L, Iagnocco A, Bruyn GA, Naredo E, Vojinovic J, Collado P, et al. The OMERACT Ultrasound Group: a report from the OMERACT 2016 meeting and perspectives. J Rheumatol 2017;44:1740–3.
3. Plagou A, Teh J, Grainger AJ, Schueller-Weidekamm C, Sudoł-Szopińska I, Rennie W, et al. Recommendations of the ESSR Arthritis Subcommittee on ultrasonography in inflammatory joint disease. Semin Musculoskelet Radiol 2016;20:496–506.
4. Collado P, Jousse-Joulin S, Alcalde M, Naredo E, D'Agostino MA. Is ultrasound a validated imaging tool for the diagnosis and management of synovitis in juvenile idiopathic arthritis? A systematic literature review. Arthritis Care Res (Hoboken) 2012;64:1011–9.
5. Lanni S, Wood M, Ravelli A, Magni-Manzoni S, Emery P, Wakefield RJ. Towards a role of ultrasound in children with juvenile idiopathic arthritis. Rheumatology (Oxford) 2013;52:413–20.
6. Collado P, Malattia C. Imaging in paediatric rheumatology: is it time for imaging? Best Pract Res Clin Rheumatol 2016;30:720–35.
7. Naredo E, Iagnocco A. One year in review: ultrasound in arthritis. Clin Exp Rheumatol 2016;34:1–10.
8. Lanni S, Bovis F, Ravelli A, Viola S, Magnaguagno F, Pistorio A, et al. Delineating the application of ultrasound in detecting synovial abnormalities of the subtalar joint in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2016;68:1346–53.
9. Rebollo-Polo M, Koujok K, Weisser C, Jurencak R, Bruns A, Roth J. Ultrasound findings on patients with juvenile idiopathic arthritis in clinical remission. Arthritis Care Res (Hoboken) 2011;63:1013–9.
10. Nusman CM, de Horatio LT, Hemke R, van Gulik EC, Ording Müller LS, Malattia C, et al. Imaging in juvenile idiopathic arthritis: international initiatives and ongoing work. Pediatr Radiol 2018;48:828–34.
11. Collado P, Naredo E, Calvo C, Gamir ML, Calvo I, García ML, et al. Reduced joint assessment vs comprehensive assessment for ultrasound detection of synovitis in juvenile idiopathic arthritis. Rheumatology (Oxford) 2013;52:1477–84.
12. Staheli LT. Fundamentals of pediatric orthopedics. 5th ed. Seattle: Lippincott Williams & Wilkins; 2015.
13. Collado P, Vojinovic J, Nieto JC, Windschall D, Magni-Manzoni S, Bruyn GA, et al. Toward standardized musculoskeletal ultrasound in pediatric rheumatology: normal age-related ultrasound findings. Arthritis Care Res (Hoboken) 2016;68:348–56.
14. Windschall D, Collado P, Vojinovic J, Magni-Manzoni S, Balint P, Bruyn GA, et al. Age-related vascularization and ossification of joints in children: an international pilot study to test multi-observer ultrasound reliability. Arthritis Care Res (Hoboken) 2017. E-pub ahead of print.
15. Windschall D, Trauzeddel R, Haller M, Krumrey-Langkammerer M, Nimtz-Talaska A, Berendes R, et al. Pediatric musculoskeletal ultrasound: age- and sex-related normal B-mode findings of the knee. Rheumatol Int 2016;36:1569–77.
16. Vojinovic J, Magni-Manzoni S, Collado P, Windschall D, Ravagnani V, Hernandez-Diaz C, et al. Ultrasonography definitions for synovitis grading in children: the OMERACT pediatric ultrasound task force [abstract]. Ann Rheum Dis 2017;76 Suppl 2:1015.
17. Algergawy S, Haliem T, Al-Shaer O. Clinical, laboratory, and ultrasound assessment of the knee in juvenile rheumatoid arthritis. Clin Med Insights Arthritis Musculoskelet Disord 2011;4:21–7.
18. Hirsch G, O'Neill T, Kitas G, Klocke R. Distribution of effusion in knee arthritis as measured by high-resolution ultrasound. Clin Rheumatol 2012;31:1243–6.
19. Beitinger N, Ehrenstein B, Schreiner B, Fleck M, Grifka J, Lüring C, et al. The value of colour Doppler sonography of the knee joint: a useful tool to discriminate inflammatory from non-inflammatory disease? Rheumatology (Oxford) 2013;52:1425–8.

20. Hong BY, Lim SH, Cho YR, Kim HW, Ko YJ, Han SH, et al. Detection of knee effusion by ultrasonography. Am J Phys Med Rehabil 2010;89:715–21.
21. Martinoli C, Della Casa Alberighi O, Di Minno G, Graziano E, Molinari AC, Pasta G, et al. Development and definition of a simplified scanning procedure and scoring method for haemophilia early arthropathy detection with ultraound (HEAD-US). Thromb Haemost 2013;109:1170–9.
22. Terslev L, D'Agostino MA, Brossard M, Aegerter P, Balint P, Backhaus M, et al. Which knee and probe position determines the final diagnosis of knee inflammation by ultrasound? Results from a European multicenter study. Ultraschall Med 2012;33:E173–8.
23. Luz KR, Pinheiro MM, Petterle GS, Dos Santos MF, Fernandes AR, Natour J, et al. A new musculoskeletal ultrasound scoring system (US10) of the hands and wrist joints for evaluation of early rheumatoid arthritis patients. Rev Bras Reumatol Engl Ed 2016;56:421–31.
24. Mandl P, Naredo E, Wakefield RJ, Conaghan PG, D'Agostino MA, for the OMERACT Ultrasound Task Force. A systematic literature review analysis of ultrasound joint count and scoring systems to assess synovitis in rheumatoid arthritis according to the OMERACT filter. J Rheumatol 2011;38:2055–62.
25. Roth J, Ravagnani V, Backhaus M, Balint P, Bruns A, Bruyn GA, et al. Preliminary definitions for the sonographic features of synovitis in children. Arthritis Care Res (Hoboken) 2017;69:1217–23.
26. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8: 23–34.
27. Cicchetti DV. Multiple comparison methods: establishing guidelines for their valid application in neuropsychological research. J Clin Exp Neuropsychol 1994;16:155–61.
28. D'Agostino MA, Terslev L, Aegerter P, Backhaus M, Balint P, Bruyn GA, et al. Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce. Part 1: definition and development of a standardised, consensus-based scoring system. RMD Open 2017;3:e000428.
29. Collado P, Windschall D, Vojinovic J, Magni-Manzoni S, Balint P, Bruyn GA, et al. Amendment of the OMERACT ultrasound definitions of joints' features in healthy children when using the Doppler technique. Pediatr Rheumatol Online J 2018;16:23.
30. Terslev L, Naredo E, Aegerter P, Wakefield RJ, Backhaus M, Balint P, et al. Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce. Part 2: reliability and application to multiple joints of a standardised consensus-based scoring system. RMD Open 2017;3:e000427.
31. Naredo E, Rodríguez M, Campos C, Rodríguez-Heredia JM, Medina JA, Giner E, et al. Validity, reproducibility, and responsiveness of a twelve-joint simplified power Doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. Arthritis Rheum 2008;59:515–22.
32. Ventura-Ríos L, Hernández-Díaz C, Ferrusquia-Toríz D, Cruz-Arenas E, Rodríguez-Henríquez P, Alvarez Del Castillo AL, et al. Reliability of ultrasound grading traditional score and new Global OMERACT-EULAR Score System (GLOESS): results from an inter- and intra-reading exercise by rheumatologists. Clin Rheumatol 2017;36:2799–804.
33. Torp-Pedersen ST, Terslev L. Settings and artefacts relevant in colour/power Doppler ultrasound in rheumatology. Ann Rheum Dis 2008;67:143–9.
34. Schmidt WA, Schönau V, Reiche BE, Oberdorfer PD, Ohrndorf S, Backhaus M. Grading of ultrasound Doppler signals in synovitis: does it need an update? Rheumatology (Oxford) 2015;54: 1897–903.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 7, July 2019, pp 986–992
DOI 10.1002/acr.23706

Patients With Early-Onset Gout and Development of Earlier Severe Joint Involvement and Metabolic Comorbid Conditions: Results From a Cross-Sectional Epidemiologic Survey

Tristan Pascart,[1] Laurène Norberciak,[1] Hang-Korng Ea,[2] Pascal Guggenbuhl,[3] and Frédéric Lioté[2]

Objective. Little is known of the clinical features and comorbidity profile of patients presenting with early-onset gout (EOG), although international guidelines recommend rapid treatment after diagnosis. The objective of this study was to assess specific characteristics and comorbidities of patients with gout who had an early onset.

Methods. Patients from a cross-sectional French national cohort who experienced their first gout flare before age 40 years were included in the EOG group and compared to patients with an onset after age 40 years, the common gout group.

Results. A total of 120 patients were included in the EOG group (mean ± SD age 49.5 ± 11.9 years) and 865 patients in the common gout group (mean ± SD age 64.4 ± 10.1 years). Patients with EOG more often presented with a history of polyarticular flares ($P < 0.01$), but had similar frequency of flares ($P = 0.16$), gout arthropathy ($P = 0.79$), and tophi ($P = 0.53$). Prevalence of each item comprising metabolic syndrome did not differ between groups. In patients with EOG, all cardiovascular comorbidities were diagnosed after gout onset. Greater age, low high-density lipoprotein, and excessive alcohol intake were associated in multivariate analysis with the common gout group, while a familial history of gout, longer duration of urate-lowering treatment, higher serum uric acid levels, and metabolic syndrome were associated with the EOG group.

Conclusion. Patients with EOG developed slightly more severe joint involvement and earlier metabolic disorders than patients with common gout.

INTRODUCTION

Gout is the most common inflammatory arthritis, with a recent prevalence estimated at 0.9% in France and 3.9% in the US (1,2). The disease is triggered by monosodium urate deposition after longstanding hyperuricemia (3). Unsurprisingly, given the natural history of the disease, results of epidemiologic studies agree that gout incidence increases with age until 70 years and that onset before age 40 years is unusual (4,5). Nonetheless, this observation does not apply to a significant proportion of patients; the prevalence of gout onset in adults between the ages of 30 and 39 years reaches 1.3% in the US (2).

Patients presenting with early-onset gout (EOG) have been given specific attention by the recent European League Against Rheumatism (EULAR) and British Society of Rheumatology guidelines that recommend a rapid initiation of urate-lowering therapy in patients diagnosed with gout before age 40 years (6,7). Apart from a few studies of Asian patients (8–10), little is known of the clinical features of patients with EOG in other populations, particularly their comorbidity profile. Profiling these patients is a prerequisite to

Supported by the Association pour la Recherche en Pathologie Synoviale. The Gospel research project was supported by Laboratoires Galéniques Vernin, a former Mayoly-Spindler company, and Ipsen Pharma.

[1]Tristan Pascart, MD, PhD, Laurène Norberciak, BSc: Université de Lille, EA440, and Hôpital Saint-Philibert, F-59160, Lomme, France; [2]Hang-Korng Ea, MD, PhD, Frédéric Lioté, MD, PhD: Université Paris Diderot, Sorbonne Paris Cité, F-75205, AP-HP, Hôpital Lariboisière, Viggo Petersen, F-75010, and INSERM, UMR 1132, Paris, France; [3]Pascal Guggenbuhl, MD, PhD: CHU de Rennes, F-35000, Institut Numecan, INSERM U 1241, INRA U 1341, F-35000, and Université de Rennes 1, F-35000 Rennes, France.

Dr. Pascart has received consulting fees from Ipsen Pharma and Mayoly-Spindler (less than $10,000 each). Dr. Guggenbuhl has received consulting frees from Ipsen Pharma and Menarini France (less than $10,000 each). Dr. Lioté has received consulting fees from AstraZeneca, Grunenthal, Ipsen Pharma, Menarini France, Mayoly-Spindler, Novartis France, and Novartis Global (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Frédéric Lioté, MD, PhD, Centre Viggo Petersen, Hôpital Lariboisière, 2 rue Ambroise Paré, F-75010, Paris, France. E-mail: frederic.liote@aphp.fr.

Submitted for publication April 16, 2018; accepted in revised form July 17, 2018.

SIGNIFICANCE & INNOVATIONS

- Patients presenting with gout before age 40 years develop earlier metabolic comorbidities.
- Patients presenting with gout before age 40 years develop earlier severe joint involvement.
- Patients with early-onset gout have preserved their renal function compared to patients developing gout later on.

identifying patients and confirming the need for tailored management of gout in this population.

The GOSPEL cohort (Goutte–observation des stratégies de prise en charge en médecine ambuLatoire) included a nationwide representative population of patients treated for gout in outpatient practice in France (11). The objective of this GOSPEL 4 study was to compare the clinical presentation, evolution, disease characteristics, and comorbidities of patients with EOG to the general population of patients with gout.

PATIENTS AND METHODS

Study population. This study is part of the GOSPEL survey, completed in 2009, whose design and patient characteristics have been published elsewhere (11). This national cross-sectional epidemiologic survey included 1,003 outpatients ages >18 years diagnosed by their own physician (private practice only) as having gout. Patients who experienced their first gout flare prior to age 40 years were included in the EOG group and compared to patients with onset after age 40 years (common gout group).

Patients' clinical features, gout history, comorbidities, and treatments prescribed were recorded by physicians (general practitioners and private practice rheumatologists) at the end of the baseline visit. In particular, time from the first manifestations of gout was noted, as well as time from and to the diagnosis of comorbid conditions. Items of metabolic syndrome were defined using the latest accepted definition: obesity (increased waist circumference >94 cm for men and >80 cm for women), elevated blood pressure (systolic ≥130 mm Hg or diastolic ≥85 mm Hg or ongoing antihypertensive therapy), elevated triglycerides (≥150 mg/dl or treatment), low high-density lipoprotein (HDL) cholesterol (≤40 mg/dl in men and ≤50 mg/dl in women or treatment), and hyperglycemia (≥100 mg/dl or drug treatment for elevated glucose) (12). The prevalence of metabolic syndrome was secondarily calculated by the investigators and defined by the presence of ≥3 items of metabolic syndrome, whether or not those items had been searched for by physicians (12).

The creatinine clearance level estimated by the Cockcroft-Gault formula gave an estimated glomerular filtration rate (eGFR). Only significant moderate or worse chronic kidney diseases (CKDs) were considered (13,14). Stage 2 CKD was not determined, given that there was no data collected for proteinuria, renal imaging, or kidney histologic findings. Stage 3 CKD was defined as a moderate alteration in eGFR between 30 and 60 ml/minute, stage 4 CKD was defined as a severe decrease in eGFR to between 15 and 30 ml/minute, and stage 5 CKD related to kidney failure, with eGFR below 15 ml/minute (13).

Statistical analysis. All statistical analyses were performed using R software, version 3.4. Qualitative variables were described as number (%) of each response modality; the number of missing data was recorded. Quantitative variables were described as mean ± SD, semiquantitative variables as median (interquartile interval) and the number of missing data. The 2 groups were compared on all variables. For quantitative and semiquantitative variables, Student's *t*-test was used for normal data, and the nonparametric Mann-Whitney test was used otherwise. Qualitative variables were assessed using the chi-square or Fisher's exact test as appropriate.

Multivariate analysis was then applied to identify significant associations of the variables with the EOG and common gout groups. Using patients who had all data available (no missing values), a binary logistic regression model was fitted with the variables exhibiting *P* values less than 0.2 in the bivariate analysis (group comparison). Selection of variables by an automatic step-by-step method based on the Akaike information criterion was used. Adjusted odds ratios (ORs) and 95% confidence intervals (95% CIs) were presented.

Since the sample of patients without any missing values (379 of 985) was insufficiently representative of the whole study population, a multivariate analysis including a multiple imputation strategy using chained equations was implemented. We considered the hypothesis that the process generating missing data is missing at random. The Mice R software package was used, with 5 imputations for all missing data. Five imputed samples were obtained. The binary logistic model with automatic variable selection was fitted on each sample and pooled ORs and 95% CIs were computed.

In order to observe the influence of missing data imputation on the estimation model, the results obtained with the complete cases and those obtained using multiple imputations were compared. All statistical tests were 2-sided, and *P* values less than or equal to 0.05 were considered significant.

RESULTS

Bivariate analysis. The age of the first gout flare was known for 985 of the 1,003 patients (98.2%) comprising the GOSPEL cohort. Of the 985 patients, 120 (12.2%) were included in the EOG group and were mean ± SD age 49.5 ± 11.9 years at the time of the study, whereas the 865 patients in the common gout group were mean ± SD age 64.4 ± 10.1 years ($P < 0.0001$) (Table 1). The age of gout symptom onset is shown in Figure 1.

Clinical presentation suggested disease to be more severe in the EOG group as compared to patients with common gout. There was a significantly greater proportion of patients who had experi-

Table 1. Characteristics of the early onset and common gout groups*

Characteristics†	Early-onset (n = 120)	Common (n = 865)	*P*
Demographics			
Age, years (n = 985)	49.5 ± 11.9	64.4 ± 10.1	<0.0001
Age at gout onset, years (n = 985)	32.8 ± 5.7	57.2 ± 10.9	<0.0001
Gout duration, years (n = 985)	16.2 ± 13.1	6.9 ± 6.7	<0.0001
Men, %	96.7	86.6	0.08
Known family history of gout, % (n = 980)	38.1	16.7	<0.0001
Renal stone, % (n = 974)	8.5	3.6	<0.05
Excessive alcohol consumption, % (n = 793)	41.8	50.8	0.12
Daily alcohol consumption, grams/day (n = 793)	27.7 ± 27.2	31.8 ± 31.1	0.1
Daily consumption of sugar-sweetened beverages, % (n = 799)	35.4	26.6	0.09
Body mass index, kg/m^2 (n = 979)	28.7 ± 4.0	28.3 ± 4.1	0.16
Clinical tophi, % (n = 985)	17.5	19.9	0.54
Treatment			
Ongoing ULT, % (n = 973)	68.9	67.9	0.91
ULT (allopurinol) duration, years (n = 970)	11.3 ± 10.2	6.6 ± 6.0	<0.001
Serum UA level at ULT initiation, mg/dl (n = 943)	8.86 ± 1.52	8.54 ± 1.19	<0.05
Last serum UA level, mg/dl (n = 802)	6.97 ± 1.70	6.77 ± 1.97	0.14
Last serum UA level below 6.0 mg/dl, % (n = 802)	19.1	29.8	<0.05
Joint disease			
Gout arthropathy, % (n = 985)	22.5	23.6	0.79
Number of flares per year (n = 985)	2.14 ± 1.75	1.93 ± 1.49	0.06
≥2 flares per year, % (n = 985)	96.7	92.7	0.16
≥1 polyarticular attack, % (n = 956)	49.6	34.8	<0.01
Arthritis other than 1st MTP joint, % (n = 954)	53.8	40.5	<0.01
Comorbidities, %			
eGFR below 60 ml/minute (n = 690)	6.3	21.3	<0.01
Ischemic heart disease (n = 973)	3.4	9.4	<0.05
Physician-identified dyslipidemia (n = 975)	36.4	48.9	<0.01
Physician-identified hypertension (n = 979)	30.8	57.8	<0.0001
Cerebrovascular accident (n = 970)	1.7	3.3	0.57
Physician-identified diabetes mellitus (n = 973)	12	15.5	0.38
Diuretics use (n = 985)	9.2	23.5	<0.0001
Metabolic syndrome (n = 985)	53.3	64.0	<0.05
High blood pressure (n = 984)	85	91	0.06
Hyperglycemia/type 2 diabetes mellitus (n = 691)	61.8	60.5	0.93
Abdominal obesity (n = 899)	82.1	82.7	0.99
Low HDL cholesterol (n = 662)	57.6	67 .4	0.14
Hypertriglyceridemia (n = 696)	77 .6	81.6	0.53

* Values are the mean ± SD unless indicated otherwise. Metabolic syndrome was defined as the presence of ≥3 of the following (all items with missing data were considered negative): abdominal obesity (elevated waist circumference >94 cm for men, >80 cm for women), high blood pressure (systolic ≥130 mm Hg, diastolic ≥85 mm Hg, or ongoing antihypertensive therapy), hypertriglyceridemia (triglycerides ≥150 mg/dl or treatment), low high-density lipoprotein (HDL) cholesterol (≤40 mg/dl in men, ≤50 mg/dl in women or treatment), or hyperglycemia (fasting glucose ≥100 mg/dl or drug treatment for elevated glucose). Statistical significance was defined as $P < 0.05$. ULT = urate-lowering therapy; UA = uric acid; MTP = metatarsophalangeal; eGFR = estimated glomerular filtration rate (Cockcroft-Gault formula).
† n = number of patients with available data.

Figure 1. Age at gout onset and distribution of study participants.

enced arthritis other than the first metatarsophalangeal joint in the EOG group (53.8%) than in the common gout group (40.5%) ($P < 0.01$). Furthermore, significantly more patients of the EOG group versus the common gout group had experienced polyarticular flare. Disease activity was similar between groups regarding the past year's number of flares, although a lower proportion of patients reached the serum uric acid (UA) target of <6.0 mg/dl in the EOG group (19.1% versus 29.8%; $P < 0.05$) (15). Severity of the disease was also similar for both arthropathy and clinically palpable tophi (Table 1).

Patients from the EOG group reported their general health as being better (75.8% good, 23.3% fair, and 0.8% poor) compared to the health of the common gout group (63.4% good, 32.6% fair, and 4.0% poor; $P = 0.02$). On a scale of 0 (none) to 100 (worst), gout tended to have a greater impact on the mood of patients in the EOG group, with a mean ± SD score of 41.3 ± 33.0 compared to mean ± SD score of 33.6 ± 27.7 in the common gout group ($P = 0.09$). Using a similar scale, all patients considered that gout had a negative impact on their social life with a mean ± SD score of 29.3 ± 30.9 in the EOG group versus a mean ± SD score of 27.9 ± 25.5 in the common gout group ($P = 0.52$).

Regarding comorbidities, a greater proportion of patients in the common gout group had moderate to severe CKD (eGFR below 60 ml/minute using the Cockcroft-Gault formula), and 130 of 611 patients (21.3%) with available eGFR information had CKD 3 or 4 in the common gout group, compared to 5 of 79 patients (6.3%) in the EOG group ($P < 0.01$). Metabolic syndrome was significantly more prevalent in the common gout group (554 of 865 patients [64.0%]) than in the EOG group (64 of 120 patients [53.3%]; $P < 0.05$). When considering individual items of metabolic syndrome separately, the prevalence of each item was not significantly different between groups when they were measured (Table 1).

The time from the physician diagnosis of cardiovascular complications to the first symptoms of gout was significantly different between groups. Gout preceded all cardiovascular events in the EOG group in contrast with the common gout group, where all events were diagnosed at approximately the same time as gout.

Multivariate analysis. Multivariate analysis was performed using 18 variables exhibiting *P* values less than 0.2 in the bivariate analysis (group comparison). Only 12 variables were retained in ≥1 of the reduced models (a no missing values model and models on the 5 imputed samples). ORs for these 12 variables (the

Table 2. Variables associated with early-onset gout group, using the common gout group as reference*

Variables	No missing values (n = 379)		Multiple imputations (n = 985)	
	Adjusted OR (95% CI)	*P*	Adjusted OR (95% CI)	*P*
Age	0.83 (0.78–0.88)†	<0.0001†	0.8 (0.77–0.83)†	<0.0001†
Known familial history of gout	3.12 (1.28–7.69)†	0.01†	2.33 (1.31–4.13)†	0.004†
Renal stone	1.25 (0.09–8.22)	0.84	2.82 (0.86–9.18)	0.08
Excessive alcohol consumption	0.41 (0.16–1.06)	0.068	0.51 (0.29–0.9)†	0.02†
ULT (allopurinol) duration, years	1.2 (1.13–1.29)†	<0.0001†	1.23 (1.17–1.29)†	<0.0001†
Last serum UA level, mg/dl	1.006 (1.002–1.01)†	0.009†	1.006 (1.003–1.01)†	0.0006†
≥1 polyarticular attack	1.12 (0.39–3.5)	0.84	1.59 (0.86–2.95)	0.14
Arthritis other than 1st MTP joint	1.34 (0.55–3.29)	0.51	1.64 (0.93–2.9)	0.08
Physician-diagnosed dyslipidemia	3.17 (0.8–14.43)	0.12	2.15 (0.81–5.68)	0.12
Diuretics use	0.41 (0.09–1.39)	0.19	0.82 (0.35–1.89)	0.64
Metabolic syndrome	7.04 (1.4–47.03)†	0.03†	1.87 (0.91–3.83)	0.09
Low high-density lipoprotein	0.2 (0.04–0.87)†	0.04†	0.26 (0.09–0.75)†	0.01†

* Multivariate analyses of explanatory variables using a model integrating patients with all available values only (no missing values model) and a model using multiple imputations for missing values. OR = odds ratio; 95% CI = 95% confidence interval; ULT = urate-lowering therapy; UA = uric acid; MTP = metatarsophalangeal.
† Statistical significance defined as $P < 0.05$.

no missing values only model and pooled ORs for the multiple imputations model) are shown in Table 2. Overall, 40 patients in the EOG group and 339 patients in the common gout group had all data available. In the EOG group, the available data were excessive alcohol consumption (117 of 120), metabolic syndrome (70 of 120), and serum uric acid (UA) level (113 of 120). Greater age, low HDL level, and excessive alcohol intake were associated with the common gout group, while a familial history of gout, a longer duration of urate-lowering therapy treatment, higher serum UA levels, and metabolic syndrome were associated with the EOG group. Association of the EOG group with metabolic syndrome was highly significant in the model including only patients having all data available, with an odds ratio (OR) of 7.04 (95% confidence interval [95% CI] 1.4–47.03), but significance was lost in the model using multiple imputations (OR 1.87 [95% CI 0.91–3.83], $P = 0.09$). Conversely, excessive alcohol intake was not significantly associated with the common gout group in the model including only patients with all data available ($P = 0.07$), but was significant in the multiple imputation model ($P = 0.02$). Prevalence of metabolic syndrome was strongly associated with excessive alcohol consumption ($P = 0.009$). Excessive alcohol consumption was frequently a missing value (19% of all missing data), and its imputed values had therefore a high influence on the significance of the association of metabolic syndrome.

DISCUSSION

This study provides an assessment of the profile of patients experiencing evolved EOG in France. Despite longer disease duration at the time of the study, these patients with EOG presented with fewer renal and physician-identified cardiovascular comorbidities than in the common profile of patients with gout followed in clinical practice. Yet notwithstanding their younger age, patients with EOG presented with joint involvement as severe as that of patients who were 15 years older with the so-called classical profile. Even more concerning, the patients age 50 years shared the same prevalence of diabetes mellitus and individual items of metabolic syndrome as patients age 65 years with classical gout. Whereas the first signs of gout usually appear around the time of diagnosis of other metabolic comorbidities, in our patients with early onset, gout preceded most other comorbidities. These results further suggest the existence of a window of opportunity for the rapid treatment of patients developing gout before age 40 years, advocated by EULAR relying on expert opinion (6).

Knowledge of gout genetics is growing, and we now know that gene polymorphisms participate in the disease progression. Genetics can account for the development of EOG even in the absence of associated risk factors such as metabolic syndrome, excessive alcohol intake, drugs, or CKD (16). The probability is high that such patients with EOG developed gout largely because of genetic polymorphisms, given the fact that they were not particularly heavy drinkers, had on average better renal function, took fewer diuretics, and had less prevalent metabolic syndrome features. Genetic mutations, such as partial hypoxanthine guanine phosphoribosyltransferase deficiency (17), or mostly UA transportosome mutations (18), are not routinely tested for in clinical practice for gout management, and the weight of genetics in the development of gout in younger patients cannot be thoroughly addressed by this study. High frequencies of ABCG2 proteins have been found in a recent retrospective cohort study from China, without a difference between EOG and common gout (10). Findings in the study by Matsuo et al (19), demonstrate that common dysfunction of ABCG2 is a major cause of EOG, and its detection might serve to improve earlier prevention and therapy for high-risk individuals. However, the higher prevalence of a known familial history of gout in the EOG group further supports the hypothesis of a strong underlying genetic basis.

Missing data account for discrepancies in the performance of metabolic syndrome and excessive alcohol consumption in the 2 multivariate models. Multiple imputations models consider multiple scenarios in their construction, which widens the range of the OR, providing a possible explanation of why metabolic syndrome performs differently in the 2 models, because 42% of values had to be imputed (versus 29% in the common gout group). In contrast, given the very small number of missing data for alcohol consumption, the range of the OR was reduced in the multiple imputation model because the sample was increased and few imputations needed to be performed.

The results of our study suggest that patients with EOG develop earlier metabolic conditions, and despite longer disease durations, tend to preserve their kidney function. A Taiwanese case–control study of very early gout onset (before age 20 years) showed that despite their higher body mass index, patients with very early-onset tophaceous gout had on average lower lipid and fasting glucose levels when compared to middle-aged patients with common onset gout. Data from the Chinese cohort studied by Zhang et al (10) showed that patients in that cohort had overall a better cardiovascular profile and particularly had a much lower prevalence of metabolic syndrome compared to patients with late-onset gout. Furthermore, Asian patients with a very early onset of gout also had preserved renal functions, despite a longer disease duration in both Asian populations (8,20). Our multivariate analysis has also shown that the lipid profile of EOG was better than that of the common gout group, because a low HDL level was significantly associated with the latter in all models tested. Nevertheless, in bivariate analysis, the smaller than expected difference between the groups concerning the prevalence of metabolic syndrome was surprising, given the age difference (21). Multivariate analysis has confirmed that suspicion; after the model was adjusted for age, metabolic syndrome was significantly associated with the EOG group in the model using only patients with all available data. This result was only a trend in the model using multiple imputations, due to the association of excessive alcohol consumption with the common gout group. Indeed, given the strong association of metabolic syndrome

and excessive alcohol intake, some of the crude higher prevalence of metabolic syndrome in the common gout group was related to higher alcohol consumption (22–24). Furthermore, bivariate analysis and the model using multiple imputations provided worst case scenarios of the link between EOG and metabolic syndrome, because missing data, which in proportion were more important in the EOG group, were considered an absence of the item of metabolic syndrome, with a higher risk to underestimate its prevalence in the EOG group. The model including only patients with all available data certainly provided the best case scenario, and the strength of the link between EOG and metabolic syndrome probably lies in between.

Although gout has been shown to be implicated in the progression of CKD, our results support the hypothesis that renal function is less impacted by early than late gout (10,25). Taking this knowledge into account, specific attention should be given to metabolic conditions in patients with EOG, moreso than the preservation of kidney function that seems less at stake.

Overall, results of our study suggest that disease activity and severity of the joint involvement of EOG are comparable to those of later-onset gout, as confirmed by multivariate analysis. This finding is surprising, given the results of prior data found in the literature. Yu and Luo (9) performed a retrospective analysis of 1,079 Taiwanese patients with gout showing a younger age of onset than usual (average age 41.6 years) and far shorter disease duration than in our EOG group (4.2 years). Patients presented with more severe gout and recurrent yearly flares (75.9% of patients presented with ≥3 flares per year), higher serum UA levels (10.3 mg/dl), and almost as many tophi (16.8%) compared to our cohort (9). Abhishek et al (26) tried to identify factors associated with recurrent gout flares among patients with untreated gout and found that high serum UA levels and long disease duration predict recurrent flares, but with poor performance. Higher serum UA levels in the Taiwanese patients probably explain the more recurrent flares and early tophi than in our GOSPEL cohort EOG group. The 3-fold longer disease duration in the EOG group led to similar disease severity in these patients, who on average had not yet reached age 50 years, than in patients with gout with the classical profile, who are on average 15 years older. This finding further supports the recent recommendation made by EULAR and the British Society of Rheumatology to treat patients presenting with early gout, not only to prevent aggravation of comorbidities but also the outcome of severe chronic joint lesions (6,7).

We acknowledge that this study has some limitations. Patients included in the GOSPEL cohort were considered by their physician to have gout, but the diagnosis had not been confirmed by crystal analysis. A large majority of patients, however, presented with at least 6 items of the 1977 American College of Rheumatology (ACR) criteria and their proportion was similar between groups (27). The new 2015 ACR/EULAR criteria could not be applied retrospectively (28). Second, not all patients had recent biologic tests, which may have impaired to some extent the assessment of serum UA levels, dyslipidemia, and hyperglycemia, but this discrepancy could be corrected by multivariate analysis using the multiple imputation model, which is a robust and stringent statistical analysis that fully takes into account the missing data. In this model, variables that remain significant are reliable. Third, recollection of the date of the first symptoms of gout and diagnosis of comorbidities is subject to imprecision.

In routine practice, patients with EOG present during their evolution with slightly different joint involvement and similar disease severity than the more common middle-aged patients with common onset gout. Despite a younger age on average, they present with a similar prevalence of diabetes mellitus and metabolic conditions as their older counterparts, but they benefit from generally preserved renal function. Given these early joint and metabolic complications, results from this study support advocacy for an early management of patients with EOG.

ACKNOWLEDGMENTS

The authors thank Dr. Marie-Christine Andro-Delestrain (Mayoly-Spindler), Prof. Alain Saraux (Brest), Dr. Sabine Lanz (Laboratoires Mayoly-Spindler), and Dr. Charles Lambert (Ipsen Pharma) for their support and contribution to the project; Dr. Pierre Chiarelli and Samy Sahbane (Vivactis Etudes Cliniques, Courbevoie) for preparing, implementing, and conducting the study's quality controls; Catherine Delva and Sylvie Lancrenon for preparing preliminary statistical analyses; the general practitioners and rheumatologists who contributed to data collection; and the patients who agreed to participate 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. Lioté 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. Pascart, Lioté.
Acquisition of data. Pascart, Norberciak, Lioté.
Analysis and interpretation of data. Ea, Guggenbuhl, Lioté.

ROLE OF THE STUDY SPONSOR

Laboratoires Galéniques Vernin, Mayoly-Spindler, and Ipsen had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Laboratoires Galéniques Vernin, Mayoly-Spindler, or Ipsen.

REFERENCES

1. Bardin T, Bouée S, Clerson P, Chalès G, Flipo RM, Lioté F, et al. Prevalence of gout in the adult population of France. Arthritis Care Res (Hoboken) 2016;68:261–6.
2. Zhu Y, Pandya BJ, Choi HK. Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutri-

tion Examination Survey 2007–2008. Arthritis Rheum 2011;63: 3136–41.
3. Pascart T, Liote F. Gout: state of the art after a decade of developments. Rheumatology (Oxford) 2010;58:27–44.
4. Doherty M. New insights into the epidemiology of gout. Rheumatology (Oxford) 2009;48 Suppl 2:ii2–8.
5. Kuo CF, Grainge MJ, Zhang W, Doherty M. Global epidemiology of gout: prevalence, incidence and risk factors. Nat Rev Rheumatol 2015;11:649–62.
6. Richette P, Doherty M, Pascual E, Barskova V, Becce F, Castaneda-Sanabria J, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis 2017;76:29–42.
7. 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 [letter]. Rheumatology (Oxford) 2017;56:1246.
8. Lu CC, Wu SK, Chen HY, Chung WS, Lee MC, Yeh CJ. Clinical characteristics of and relationship between metabolic components and renal function among patients with early-onset juvenile tophaceous gout. J Rheumatol 2014;41:1878–83.
9. Yu KH, Luo SF. Younger age of onset of gout in Taiwan. Rheumatology (Oxford) 2003;42:166–70.
10. Zhang B, Fang W, Zeng X, Zhang Y, Ma Y, Sheng F, et al. Clinical characteristics of early- and late-onset gout: a cross-sectional observational study from a Chinese gout clinic. Medicine (Baltimore) 2016;95:e5425.
11. Liote F, Lancrenon S, Lanz S, Guggenbuhl P, Lambert C, Saraux A, et al. GOSPEL: prospective survey of gout in France. Part I: design and patient characteristics (n = 1003). Joint Bone Spine 2012;79:464–70.
12. Alberti KG, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009;120:1640–5.
13. Levey AS, Eckardt KU, Tsukamoto Y, Levin A, Coresh J, Rossert J, et al. Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney Int 2005;67:2089–100.
14. Levey AS, de Jong PE, Coresh J, El Nahas M, Astor BC, Matsushita K, et al. The definition, classification, and prognosis of chronic kidney disease: a KDIGO Controversies Conference report. Kidney Int 2011;80:17–28.
15. Bursill D, Taylor WJ, Terkeltaub R, Kuwabara M, Merriman TR, Grainger R, et al. Gout, Hyperuricemia and Crystal-Associated Disease Network consensus statement regarding labels and definitions for disease elements in gout. Arthritis Care Res (Hoboken) 2019;71:427–34.
16. Matsuo H, Yamamoto K, Nakaoka H, Nakayama A, Sakiyama M, Chiba T, et al. Genome-wide association study of clinically defined gout identifies multiple risk loci and its association with clinical subtypes. Ann Rheum Dis 2016;75:652–9.
17. Lahaye C, Auge F, Soubrier M, Ceballos-Picot I. New mutation affecting hypoxanthine phosphoribosyltransferase responsible for severe tophaceous gout. J Rheumatol 2014;41:1252–4.
18. Merriman TR. An update on the genetic architecture of hyperuricemia and gout. Arthritis Res Ther 2015;17:98.
19. Matsuo H, Ichida K, Takada T, Nakayama A, Nakashima H, Nakamura T, et al. Common dysfunctional variants in ABCG2 are a major cause of early-onset gout. Sci Rep 2013;3:2014.
20. Yamanaka H. Gout and hyperuricemia in young people. Curr Opin Rheumatol 2011;23:156–60.
21. Ford ES, Giles WH, Dietz WH. Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. JAMA 2002;287:356–9.
22. Wakabayashi I. Association between alcohol drinking and metabolic syndrome in Japanese male workers with diabetes mellitus. J Atheroscler Thromb 2011;18:684–92.
23. Jin L, Huang Y, Bi Y, Zhao L, Xu M, Xu Y, et al. Association between alcohol consumption and metabolic syndrome in 19,215 middle-aged and elderly Chinese. Diabetes Res Clin Pract 2011;92:386–92.
24. Athyros VG, Liberopoulos EN, Mikhailidis DP, Papageorgiou AA, Ganotakis ES, Tziomalos K, et al. Association of drinking pattern and alcohol beverage type with the prevalence of metabolic syndrome, diabetes, coronary heart disease, stroke, and peripheral arterial disease in a Mediterranean cohort. Angiology 2007;58:689–97.
25. Richette P, Perez-Ruiz F, Doherty M, Jansen TL, Nuki G, Pascual E, et al. Improving cardiovascular and renal outcomes in gout: what should we target? Nat Rev Rheumatol 2014;10:654–61.
26. Abhishek A, Valdes AM, Zhang W, Doherty M. Association of serum uric acid and disease duration with frequent gout attacks: a case–control study. Arthritis Care Res (Hoboken) 2016;68:1573–7.
27. Wallace SL, Robinson H, Masi AT, Decker JL, McCarty DJ, Yü TF. Preliminary criteria for the classification of the acute arthritis of primary gout. Arthritis Rheum 1977;20:895–900.
28. Neogi T, Jansen TL, Dalbeth N, Fransen J, Schumacher HR, Berendsen D, et al. 2015 gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2015;67:2557–68.

ARP Announcements

Association of Rheumatology Professionals
2200 Lake Boulevard NE, Atlanta, Georgia 30319
www.rheumatology.org

ACR/ARP Annual Meeting

November 8–13, 2019, Atlanta

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*.

New Division Name

Rheumatology is truly a people specialty; We often develop lifelong relationships with our patients as well as our colleagues. We increasingly recognize that providing the best rheumatologic care requires a team effort. The collegial nature of our specialty is reflected in the ACR's mission statement: To empower rheumatology professionals to excel in their specialty.

In keeping with this mission, we are pleased to announce that our health professionals' membership division is changing its name to Association of Rheumatology Professionals (ARP). This name change highlights the dedication of the ACR to serve the entire rheumatology community. It also reflects our broadened base of interprofessional members (administrators, advanced practice nurses, health educators, nurses, occupational therapists, pharmacists, physical therapists, physician assistants, research teams, and more).

The name is new, but our commitment and promise remain the same: We are here for you, so you can be there for your patients.

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.