Arthritis Care & Research

Aims and Scope

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

Volume 74 No. 10 October 2022

Arthritis Care & Research

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

Arthritis Care & Research

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

VOLUME 74 • OCTOBER 2022 • NO. 10

Cover image: The figure on the cover (from van Dijk et al, page 1719) shows longitudinal magnetic resonance images of decreasing intermetatarsal bursitis in a patient with early rheumatoid arthritis patient at baseline (top) and 24 months (bottom).

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1567–1574
DOI 10.1002/acr.24610

American College of Rheumatology
Empowering Rheumatology Professionals

Parent-Reported Medication Side Effects and Their Impact on Health-Related Quality of Life in Children With Juvenile Idiopathic Arthritis

Gaëlle Chédeville,[1] Katherine McGuire,[2] David A. Cabral,[3] Natalie J. Shiff,[4] Dax G. Rumsey,[5] Jean-Philippe Proulx-Gauthier,[6] Heinrike Schmeling,[7] Roberta A. Berard,[8] Michelle Batthish,[9] Gordon Soon,[10] Kerstin Gerhold,[11] Tommy Gerschman,[3] Alessandra Bruns,[12] Ciaran M. Duffy,[13] Lori B. Tucker,[3] and Jaime Guzman,[3] on behalf of the CAPRI Registry Investigators

Objective. To describe the frequency and severity of parent-reported medication side effects (SEs) in children with juvenile idiopathic arthritis (JIA) relative to physician-reported actionable adverse events (AEs), and to assess their impact on health-related quality of life (HRQoL).

Methods. Newly diagnosed JIA patients recruited between 2017 and 2019 to the Canadian Alliance of Pediatric Rheumatology Investigators (CAPRI) Registry were included. Parents reported presence and severity (0 = no problem, 10 = very severe) of medication SEs at every clinic visit. Physicians were asked to report any actionable AE. HRQoL was assessed using the Quality of My Life (QoML) questionnaire (0 = the worst, 10 = the best) and parent's global assessment (0 = very well, 10 = very poor). Analyses included proportion of visits with SEs or actionable AEs, cumulative incidence by Kaplan-Meier methods, and HRQoL impact measured with longitudinal mixed-effects models.

Results. SEs were reported at 371 of 884 (42%) visits (95% confidence interval [95% CI] 39, 45%) in 249 patients, with a median of 2 SEs per visit (interquartile range [IQR] 1–3), and median severity of 3 (IQR 1.5–5). Most SEs were gastrointestinal (32.5% of visits) or behavioral/psychiatric (22.4%). SE frequency was lowest with nonsteroidal antiinflammatory drugs alone (34.7%) and highest with prednisone and methotrexate combinations (66%). SE cumulative incidence was 67% (95% CI 59, 75) within 1 year of diagnosis, and 36% (95% CI 28, 44) for actionable AEs. Parent global and QoML scores were worse with SEs present; the impact persisted after adjusting for pain and number of active joints.

Conclusion. Parents report that two-thirds of children with JIA experience SEs impacting their HRQoL within 1 year of diagnosis. SE mitigation strategies are needed in managing JIA.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is the most common chronic rheumatic disease of childhood and affects ~1 in 1,000 children (1). One or more medications are often necessary to control disease symptoms and prevent long-term damage. Medication side effects (SEs) are a concern for parents and physicians and may impact adherence to treatment and health-related quality of life (HRQoL) (2).

While physician-reported adverse events (AEs) are commonly captured in clinical trials and drug registries, there has been little systematic study of parent perceptions of the frequency and severity of SE in JIA and their impact on quality of life (3), with the

Supported by Arthritis Society, Canada. Dr. Guzman's work was supported by the BC Children's Hospital Research Institute (Clinical Investigator award). Ms. McGuire's work was supported by a Canadian Rheumatology Association studentship.

[1]Gaëlle Chédeville, MD: McGill University, Montreal, Quebec, Canada; [2]Katherine McGuire: University of Ottawa, Ottawa, Ontario, Canada; [3]David A. Cabral, MBBS, Tommy Gerschman, MD, Lori Tucker, MD, FAAP, Jaime Guzman, MD, MSc: University of British Columbia, British Columbia, Canada; [4]Natalie J. Shiff, MD, MHSc: University of Saskatchewan, Saskatoon, Saskatchewan, Canada; [5]Dax G. Rumsey, MD, MSc: University of Alberta, Edmonton, Alberta, Canada; [6]Jean-Philippe Proulx-Gauthier, MD: Université Laval, Quebec, Quebec, Canada; [7]Heinrike Schmeling, MD: University of Calgary, Calgary, Alberta, Canada; [8]Roberta A. Berard, MD, MSc: London Health Sciences Centre and Western University, London, Ontario, Canada; [9]Michelle Batthish, MD: McMaster University, Hamilton, Ontario, Canada; [10]Gordon Soon, MD: University of Toronto, Toronto, Ontario, Canada; [11]Kerstin Gerhold, MD: University of Manitoba, Winnipeg, Manitoba, Canada; [12]Alessandra Bruns, MD: Université de Sherbrooke, Sherbrooke, Quebec, Canada; [13]Ciaran M. Duffy, MBBCh, MSc: University of Ottawa, Ottawa, Ontario, Canada.

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

Address correspondence to Gaëlle Chédeville, MD, Division of Pediatric Rheumatology, Department of Pediatrics, McGill University Health Centre, 1001 Decarie Boulevard, A4-6306, Montreal H4A 3J1, Quebec, Canada. Email: gaelle.chedeville@muhc.mcgill.ca.

Submitted for publication November 15, 2020; accepted in revised form March 25, 2021.

SIGNIFICANCE & INNOVATIONS

- Physician-reported adverse events are commonly captured in clinical trials and drug registries, but adherence and quality of life are likely directly influenced by parents' perceptions of medication side effects.
- Parent-reported medication side effects were present in two-thirds of children with juvenile idiopathic arthritis within 1 year of diagnosis.
- Most side effects were gastrointestinal or behavioral and of mild-to-moderate severity.
- Reported side effects were associated with decreased quality of life, independent of pain scores and the number of active joints.

exception of the well-known nausea and vomiting associated with methotrexate (MTX) (4). Recently, the parent/patient perspective has been emphasized with the development of patient-reported outcomes; experience with medications is included among the different domains that are assessed (5–7). Two well-known juvenile arthritis questionnaires, the Juvenile Arthritis Quality of Life Questionnaire (JAQQ) (8) and the Juvenile Arthritis Multidimensional Assessment Report (JAMAR) (9), include lists of SEs reported by patients and parents, but to date, no systematic analysis of reported SEs has been published.

The aim of this study was to describe parent-perceived SEs associated with all antirheumatic treatments prescribed in a Canadian inception cohort of children with JIA, with 2 specific objectives: 1) to describe the frequency and severity of parent-reported medication SEs relative to physician-reported actionable AEs; and 2) to assess the association of parent-reported SEs with HRQoL.

MATERIALS AND METHODS

Data from the Canadian Alliance of Pediatric Rheumatology Investigators (CAPRI) JIA Registry were used in this study (10). Recruitment into the registry began in February 2017. Children were enrolled within 3 months of JIA diagnosis at 1 of 14 participating sites, each of which obtained local ethics board approval.

Core data are collected for the Registry at every clinic visit by parents, patients, and physicians, including information on disease activity, treatments, physician-reported actionable AEs, parent-reported medication SEs, disease outcomes, and quality of life (10). There are no arbitrarily fixed study visit intervals to enter the information in the Registry. For the purpose of this study, data were extracted in May 2019 and focused on questions related to medications and HRQoL.

SEs. At every visit, parents were asked the following: 1) Is your child taking any medication for his/her arthritis? 2) If yes, is your child having any SEs from medications taken for his/her arthritis? 3) If yes, parents selected SEs from a 17-item list and could add any additional SEs that were not listed (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24610). This list was created based on the questions about SEs from 2 validated questionnaires (the JAQQ and JAMAR) (8,11) plus 2 additional items often raised by parents in clinic (infections and poor attention). Parents were also asked to rate on a 21-point horizontal numerical scale 4) How difficult or bothersome is it for your child to take their arthritis medication (by mouth or injection) (0 = no problem, 10 = very bothersome)?; and 5) Overall, what is the severity of the SEs your child has from medication taken for arthritis (0 = no problem, 10 = very severe)?

AEs. At every visit, physicians were asked if the patient had any actionable AEs since their last visit. An actionable AE was defined as any untoward medical occurrence that requires additional medical visits, investigations, treatments, or a change in arthritis medications, irrespective of its cause. If yes, AEs were selected from an 18-item list, and any not listed could be added (see Supplementary Appendix A, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24610). The AE items listed incorporated those from the German Biologics in Pediatric Rheumatology (BIKER) registry (12) and the Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry (13). The physician was also asked to report what actions were taken and if the AE was serious, defined as any AE that resulted in death, was life threatening, required hospitalization (admission for overnight stay), or resulted in a significant disability or a congenital anomaly, or if it was an AE that required medical or surgical intervention to prevent death, significant disability, or a congenital anomaly. The CAPRI Registry does not collect information on AEs that are not actionable.

Quality of life. HRQoL was assessed by the child at every visit using the Quality of My Life questionnaire (QoML) (14) if the child was old enough to answer the questions according to the parent, usually ≥6 years. The question of interest was, "Considering my HEALTH, my life is…"; the answer was rated on a 21-point horizontal numerical scale from 0 = the worst to 10 = the best. We also used the parent's global assessment (15) as a measure of the parent's perception of the child's HRQoL: "Considering all the ways that arthritis affects your child, please rate how your child is doing over the PAST WEEK"; the response was noted on a 21-point horizontal numerical rating scale from 0 = very well to 10 = very poor.

Statistical analysis. All analyses were done with Stata, version 12. Descriptive statistics included the median, interquartile range (IQR), and proportions. The frequency of SEs was calculated as the number of visits at which SEs were reported, divided by the total number of visits observed in the cohort. Using

Table 1. Characteristics of patients included in the study (n = 249)*

Characteristic	Value
Age, years	7.9 (3.4, 12.6)
Female, %	60
Weeks from diagnosis to enrollment	4 (0, 8.6)
Disease duration at baseline, weeks	22.6 (12.9, 39.4)
JIA category, no. (%)	
Oligoarthritis	108 (43.4)
Polyarthritis rheumatoid factor negative	46 (18.5)
Enthesitis-related arthritis	40 (16.1)
Psoriatic arthritis	14 (5.6)
Systemic arthritis	14 (5.6)
Undifferentiated arthritis	14 (5.6)
Polyarthritis rheumatoid factor positive	7 (2.8)
Missing	6 (2.4)
Physician global assessment at baseline	3 (1.5, 4.25)
Active joint count at baseline	2 (1, 4)
Parent global assessment at baseline	1.5 (0, 4)
QoML score at baseline (n = 160)†	7.5 (5, 9)
Medications, calculated per visit, no. (%)	
Total visits	884 (100)
No medications	132 (14.9)
Naproxen	453 (51.2)
Other NSAID	102 (11.5)
Methotrexate oral	187 (21.2)
Methotrexate subcutaneous	176 (19.9)
Other DMARD	21 (2.4)
Prednisone	94 (10.6)
Other glucocorticoid	4 (0.5)
Biologic	104 (11.8)
Ocular glucocorticoid	24 (2.7)

* Values are the median (25th, 75th percentiles) unless indicated otherwise. DMARD = disease-modifying antirheumatic drug; JIA = juvenile idiopathic arthritis; NSAID = nonsteroidal antiinflammatory drug; QoML = Quality of My Life.
† The QoML score was available for children who were old enough to complete the questionnaire at the parent's discretion, usually ≥6 years.

this global denominator provides a measure of the overall burden of SEs in the cohort and comparable metrics across all SEs. It also avoids ambiguity because in some visits, physicians and parents disagreed as to whether the child was receiving antirheumatic medications (perhaps the patient was not taking prescribed treatments, or the parents were providing treatments they considered antirheumatic but were not prescribed by their rheumatologist). For comparisons of frequency of SEs across different medication regimens, the denominator was the total number of visits at which that regimen was reported by the rheumatologist, and *P* values were calculated with the chi-square test. Calculation of SE incidence as SEs per 100 patient-years of observation was not done because SEs often persisted from one visit to the next, and some patients reported multiple SEs. Instead, Kaplan-Meier survival methods were used to estimate the cumulative incidence of parent-reported SEs and physician-reported actionable AEs. Longitudinal mixed-effects models were used to assess the impact of SE severity on QoML and parent global scores before and after adjusting for pain severity and the number of active joints. All variables were modeled as time-varying,

Table 2. Frequency of parent-reported side effects (SEs) and physician-reported actionable adverse events (AEs)*

SEs grouped by system	Parent-reported medication SEs	Physician-reported actionable AEs
At least 1 GI symptom	287 (32.5)	69 (7.8)
Abdominal pain	117 (13.2)	15 (1.7)
Loss of appetite	106 (12.0)	2 (0.2)
Nausea	93 (10.5)	45 (5.1)
Constipation	55 (6.2)	1 (0.1)
Mouth sores	46 (5.2)	2 (0.2)
Weight gain	43 (4.9)	–
Diarrhea	34 (3.8)	2 (0.1)
Heartburn	27 (3.1)	–
Blood in stool	6 (0.7)	6 (0.7)
Increase of appetite	1 (0.1)	–
Weight loss	1 (0.1)	–
IBD	–	1 (0.1)
At least 1 behavioral symptom	198 (22.4)	7 (0.8)
Mood changes	108 (12.2)	1 (0.1)
Sleep problems	73 (8.3)	–
Headaches	54 (6.1)	–
Tired, fatigue	24 (2.7)	1 (0.1)
Poor attention	16 (1.8)	–
Lightheaded or dizzy	3 (0.3)	3 (0.3)
Anxiety	2 (0.2)	–
Depression	1 (0.1)	1 (0.1)
Fidgeting	1 (0.1)	–
Irritable	1 (0.1)	1 (0.1)
At least 1 skin disorder	70 (7.9)	17 (1.9)
Rash or hives	33 (3.7)	10 (1.1)
Injection reaction	25 (2.8)	3 (0.3)
Facial edema	5 (0.6)	2 (0.2)
Dry skin	4 (0.5)	–
Hair loss	4 (0.5)	1 (0.1)
Excess hair growth	3 (0.3)	–
Stretch mark	1 (0.1)	–
Photosensitivity	1 (0.1)	–
Acne	1 (0.1)	–
Ecchymosis	1 (0.1)	–
Pseudoporphyria	–	1 (0.1)
SC atrophy after joint injection	–	1 (0.1)
Infections	14 (1.5)	6 (0.7)
Other	8 (0.9)†	21 (2.4)‡

* Values are the no. (%) of total visits. GI = gastrointestinal; IBD = inflammatory bowel disease; SC = subcutaneous.
† Other SEs included dark or blood-tinged urine in 2 visits and 1 each of the following: aches, heavy periods, nose bleeds, urinary incontinence, hand tremors, and excessive salivation.
‡ Other AEs included 12 abnormal bloodwork results, 4 infusion reactions, 2 dyspnea episodes, 1 adrenal suppression episode after glucocorticoid joint injection, 1 epistaxis, and 1 joint surgery.

and models included a quadratic term for time, as recommended by Rabe-Hesketh and Skrondal (16).

Data availability. The data underlying this article are available by contacting the authors. Access to unpublished CAPRI JIA registry data may be granted to other investigators provided that 1) they collaborate in a team that includes at least 1 CAPRI JIA registry investigator, and 2) their research protocol is approved

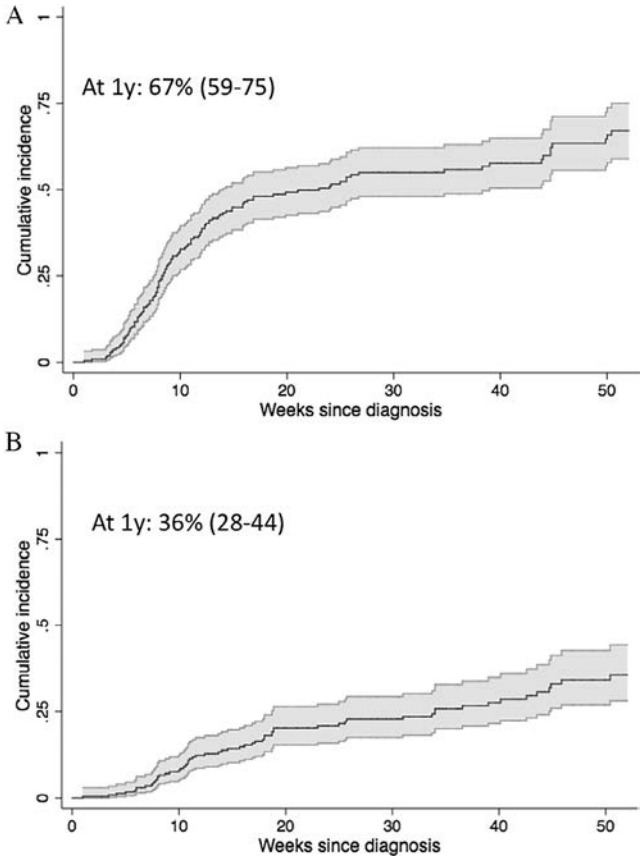

Figure 1. Cumulative incidence of parent-reported side effects (**A**) and physician-reported actionable adverse events (**B**) calculated with Kaplan-Meier methods. The shaded area is the 95% confidence interval.

by the Canadian Alliance of Pediatric Rheumatology Investigators Scientific Protocol Evaluating Committee. For more details, contact Dr. Jaime Guzman at jguzman@cw.bc.ca.

RESULTS

As of May 2019, 975 visits in 275 newly diagnosed patients were available in the CAPRI Registry. Eleven visits (1.1%) were excluded due to missing diagnosis date, 14 visits (1.4%) due to missing physician data, and 66 visits (6.8%) due to missing parent data. The remaining 884 visits (90.5%) from 249 patients were included in the analysis. The characteristics of included patients are shown in Table 1.

SEs and actionable AEs. Parents reported at least 1 medication SE in 371 of 884 visits (42%) (95% confidence interval [95% CI] 39, 45), with a median of 2 SEs per visit (IQR 1–3). Table 2 reports the frequency of different types of SE and the frequency of physician-reported actionable AEs. The most frequent SEs were gastrointestinal (GI) (32.5% of 884 visits) and behavioral/psychiatric symptoms (22.4% of 884 visits). Physicians reported at least 1 actionable AE at 112 of 884 visits (12.7%). Eighty-nine visits had both a physician-reported actionable AE and a parent-reported SE. In total, 9 actionable AEs were considered serious: 3 infections with hospitalization, and 1 each of nausea/vomiting, adrenal suppression, inflammatory bowel disease, GI bleed, facial edema, and epistaxis.

Figure 1 contrasts the cumulative incidence of SEs and actionable AEs. Within the first year of diagnosis, the proportion

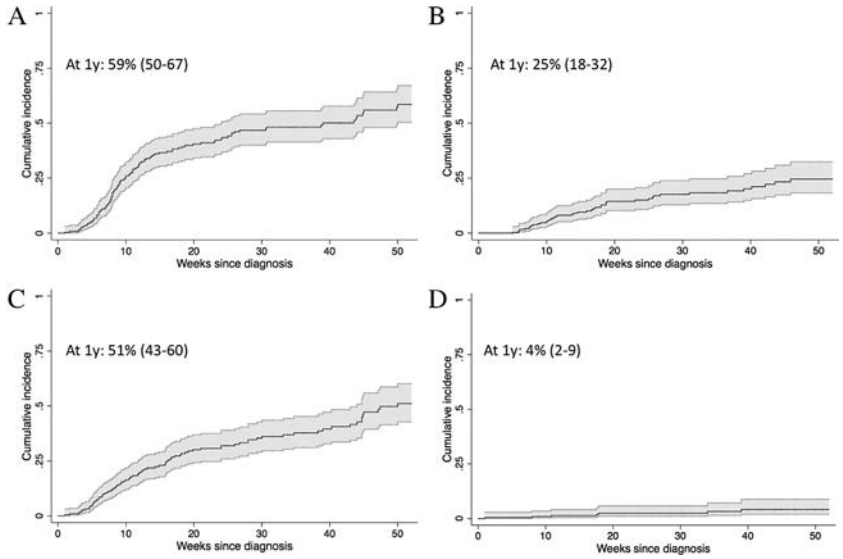

Figure 2. Cumulative incidence of parent-reported gastrointestinal (GI) side effects (**A**), physician-reported actionable GI adverse events (**B**), parent-reported behavioral side effects (**C**), and physician-reported actionable behavioral adverse events (**D**) calculated with Kaplan-Meier methods. The shaded area is the 95% confidence interval.

Table 3. Frequency and severity of parent-reported side effects (SEs) for the most common drug regimens*

Medication regimen†	No. of visits/no. of patients‡	SE frequency no. (%) of visits	SE severity	Difficulty taking medication
NSAID only	314/148	109 (34.7)	2.5 (1.5–4.0)	1 (0–5)
MTX only	98/54	58 (59.2)	2 (1–4)	2 (0–5)
NSAID plus MTX	147/75	79 (53.7)	3 (2–5)	3.75 (2.0–6.5)
Prednisone plus MTX ± other	50/27	33 (66)	3.5 (1–5)	4 (0.5–7.5)
Biologic only	10/6	3 (30)	2 (0.5–3.0)	2 (1–8)
Biologic plus NSAID	6/3	4 (66.7)	3.75 (2.0–5.25)	5.25 (1–7)
Biologic plus MTX	31/15	11 (35.5)	3 (1–5)	4.5 (2.5–6.0)
Biologic plus MTX ± other	72/26	35 (48.6)	3.5 (1–5)	6 (3.0–7.5)
MTX oral versus subcutaneous				
NSAID plus oral MTX	82/45	39 (49)	3 (2–4)	3 (0.5–5.0)
NSAID plus subcutaneous MTX	65/37	39 (60)	3.5 (1.5–5.0)	5 (3–7)
Oral MTX	50/29	28 (56)	2.75 (1.5–4.2)	1.7 (0–6)
Subcutaneous MTX	48/28	30 (63)	2 (1–4)	2 (0–5)

* Values are the median (interquartile range) unless indicated otherwise. MTX = methotrexate; NSAID = nonsteroidal antiinflammatory drug.
† Any of these regimens could be accompanied by intraarticular glucocorticoid injections.
‡ The no. of patients adds to more than the total number of patients in the cohort because patients can have one medication regimen at one visit and another at a subsequent visit.

of parents reporting at least 1 SE was 67% (95% CI 59, 75), and the proportion of physicians reporting at least 1 actionable AE was 36% (95% CI 28, 44). Cumulative incidence of GI and behavioral effects is shown in Figure 2. For behavioral/psychiatric symptoms, the cumulative incidence was 51% for SEs, and only 4% for actionable AEs.

Table 3 reports the frequency and severity of SEs and difficulty in taking medications according to the drug regimen prescribed. The median severity of all reported SEs was 3 (IQR 1.5–5). At more than one-half of the visits (58.4%), the severity of SEs was rated ≤3 and in 83.8% was rated ≤5. Of note, in 15 visits, the SE had negligible severity, rated as 0. The frequency of SE was 35% with nonsteroidal antiinflammatory drug (NSAID) monotherapy, 59% with MTX monotherapy, and 66% with prednisone and MTX combinations ($P < 0.001$ by chi-square for differences across regimens). However, the median severity was similar across regimens. There was no significant difference in frequency of SEs according to the route of MTX (oral versus subcutaneous) (Table 3).

The overall median difficulty in taking medications was 2 (IQR 0–5); 2.5 (IQR 0–6) in the presence of SE, and 1.0 (IQR 0–3.5) if no SE was present. It varied across medication regimens, from 1 in patients taking NSAIDs only to 6 in patients taking biologics and MTX. Regimens including medications given by injection had higher difficulty scores (Table 3).

Quality of life. HRQoL was evaluated by parents using the parent global assessment and by children using the QoML. The median parent global assessment was 2.0 (IQR 0.5–5) in the presence of SEs and 0.5 (IQR 0–2) if no SE was present. The parent global assessment was higher (worse) with more severe SEs, a median of 4 (IQR 1.5–6.5) if SE severity was >3 compared to 1.5

Table 4. Selected factors influencing health-related quality of life according to mixed-effects models*

	Impact on parent global assessment				Impact on Quality of My Life scale			
Variable	Unadjusted β coefficient (95% CI)	*P*	Adjusted β coefficient (95% CI) (n = 817)	*P*	Unadjusted β coefficient (95% CI)	*P*	Adjusted β coefficient (95% CI) (n = 534)	*P*
Severity of SE	0.356 (0.280, 0.431)	<0.001	0.185 (0.125, 0.245)	<0.001	−0.186 (−0.269, −0.102)	<0.001	−0.087 (−0.164, −0.010)	0.03
Pain	0.635 (0.587, 0.684)	<0.001	0.577 (0.526, 0.627)	<0.001	−0.327 (−0.387, −0.268)	<0.001	−0.284 (−0.345, −0.222)	<0.001
No. of active joints	0.141 (0.107, 0.175)	<0.001	0.051 (0.025, 0.078)	<0.001	−0.106 (−0.140, −0.073)	<0.001	−0.064 (−0.096, −0.032)	<0.001

* Unadjusted β coefficients are from models including only the variable of interest. Adjusted β coefficients are from models including all 3 variables at once. 95% CI = 95% confidence interval; SE = side effect.

(IQR 0–3) if SE severity was ≤3. The median QoML score was 7 (IQR 5–8.5) when SEs were present and 8 (IQR 6.5–9.5) if no SE was present. The QoML score was lower (worse) in the presence of more severe SEs, with a median of 6 (IQR 4.7–8.2) if SE severity was >3 compared to 7.5 (IQR 5.5–9) if severity was ≤3.

Longitudinal mixed-effects models showed that SE severity had a measurable impact on HRQoL even after adjusting for pain scores and the number of active joints (Table 4). Pain had the largest impact on HRQoL in adjusted analyses, with a β coefficient of 0.577 for the parent global assessment and of –0.284 for QoML. In other words, a 1-unit increment in the 0–10 pain scale corresponded to a 0.577-unit increase in the parent global assessment and a 0.284-unit decrease in the QoML score. SE severity had a β coefficient for the parent global assessment of 0.185, and for QoML score, a β coefficient of –0.087, numerically larger than the impact of the number of active joints (0.051 and –0.064, respectively) (Table 4).

DISCUSSION

In this study, we report the perspective of parents concerning SEs of antirheumatic medications taken by their children for the treatment of JIA. In two-thirds of children with JIA, parents reported at least 1 SE within 1 year of diagnosis. Children reported by parents to have SEs had impaired quality of life, independent of JIA-related pain scores or the number of active joints. While most pediatric rheumatologists are aware of difficulties that parents face giving medications, we believe that this is the first systematic study of parent-reported SEs associated with all antirheumatic medications used in a modern inception cohort of children with JIA.

We found a high frequency of parent-reported medication SEs (42% of visits), and the risk of developing at least 1 SE during the first year after diagnosis was 67%.These frequency and incidence estimates cannot be directly compared to the rates of physician-reported actionable AEs, but their side-by-side analysis helps put SEs in perspective and offers interesting insights. Parents and physicians were provided different lists to choose from (see Supplementary Appendix A, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24610), with the parent's list emphasizing symptoms using lay language and recording all levels of severity, while physicians were instructed to report only AEs selected from the list that required an action to be taken. Some of the SEs were mild and may have been deemed not to require action by both parents and physicians and therefore would have been recorded by the parent as an SE but not by the physician as an actionable AE. Other SEs may have been unknown to the physician, even though we encouraged sharing of the parents' list. There was a higher frequency of SEs while receiving MTX, which reflects the well-known SEs of this medication (2,4). It is important to underline that NSAID monotherapy also had a high frequency of SEs.

In our study, the most frequent GI SEs reported by parents were abdominal pain, poor appetite, and nausea. With MTX, it is well-known that parent-reported nausea, vomiting, and behavioral difficulties occur in approximately one-half of children and impact quality of life (4,17). A MTX Intolerance Severity Score to measure MTX intolerance has been validated (18), and models to predict intolerance have been published (19). A more general questionnaire, Gastrointestinal Symptom Scale for Kids (GISSK), was developed by Brunner et al in a convenience sample of children with JIA receiving second-line agents (20). It includes a visual analog scale to assess severity similar to the one used in our study, but with a different anchor (severe stomach problems). Although 58% of parents in their study reported some GI symptoms, the median severity was low (6 of 100). Several randomized trials of NSAIDs have reported GI AEs in children with JIA. Foeldvari et al reported a 36.1% frequency of GI AEs with naproxen at 7.5 mg/kg twice a day for 12 weeks (21). Ruperto et al reported 32% with naproxen at 5 mg/kg twice a day in a 1-year trial (22), and Lovell et al reported 37% with different doses of naproxen/esomeprazole combinations for up to 6 months (23). These rates are comparable to our 34.7% frequency of parent-perceived GI SEs with NSAID monotherapy (mostly naproxen).

Behavioral/psychiatric SEs in children with JIA have not been well characterized in the literature but had a remarkably high incidence in our study, with an estimated 50% of parents reporting symptoms of mood change, sleep problems, or headache in their children at least once during the first year after diagnosis. Headaches have been reported by physicians in 6–15% of children with JIA receiving naproxen in the above-mentioned trials (18–20) and were reported by parents in 6% of visits in our study. Without a control group of age- and sex-matched children as a comparison, it is difficult to know what the general background report for these symptoms might be, but parents in our study clearly attributed them to the antirheumatic medications.

Some of these GI and behavioral SEs could be the result of nocebo effects (24,25). As an example, in placebo-controlled trials of NSAIDs for the control of migraine, patients receiving placebo often reported GI and behavioral symptoms, and furthermore, studies that use structured lists to elicit SEs also report higher frequencies of SEs (26). On the other hand, we should be cautious of dismissing parent reports as nocebo effects without further evidence. Even if the nocebo effect plays a role in these reports, we can see that parent-perceived SEs in their children clearly impacted HRQoL, likely influencing parents' choices about treatment and adherence to treatment.

Strategies to mitigate SEs are needed, but it is not clear what those strategies should be. Trials with cyclooxygenase 2 inhibitors and combinations of NSAIDs with proton-pump inhibitors (PPI) have not been very encouraging to date (18–20), and these drugs (PPI) are unlikely to have much impact on behavioral/psychiatric symptoms. As for MTX, possibilities such as adding ondansetron

and cognitive behavioral therapies have been reported with mixed results (17).

HRQoL is an important outcome of the care that we provide to our patients with JIA. When looking at the presence of SEs and their severity, we can note a clear impact on HRQoL. The lowered HRQoL of JIA patients has been previously established (2,27,28), and our study suggests that SEs play a role in this lowered HRQoL. Brunner et al showed that patients with a GISSK score of ≥2 had a significantly lower HRQoL than others without GI symptoms but similar disease activity (20). In their cohort of JIA patients, Weitzman et al enrolled 180 parent–patient dyads to complete patient-reported outcomes during routine care. They reported that measures of disease and treatment burden were independently negatively associated with HRQoL (2). These data, connecting medication issues to low HRQoL, are important information for physicians treating children with JIA if we want to improve the lives of our patients. The tradeoff between pain control and improved function versus SEs is a well-known struggle for physicians and parents alike, and we should continue to study patients' perceptions to include this vital data in our decision-making. In concordance with our finding that pain has the largest effect on a patient's quality of life, a study by Burnett et al suggested that parents value a medication's effectiveness at controlling pain and improving function more highly than the possibility of negative SEs (29).

Similar to what we observed in our JIA population, Cooper et al reported a high prevalence of patient-reported SEs, contrasting with low clinician estimates of SEs in a cross-sectional study of adults with asthma (30). They also noted that a greater number of SEs were associated with nonadherence to oral steroids, which we did not specifically assess in our study. Similar findings were also reported for adults with rheumatic diseases enrolled in the German RABBIT (Rheumatoide Arthritis: Beobachtung der Biologika-Therapie) registry (31). There was good agreement for easily observable, objective medication SEs between patients and physicians, but the agreement was low for more subjective SEs that could have an impact on quality of life.

When evaluating the negative experiences of JIA patients with medications in clinical trials or registries, it is likely not sufficient to consider physician-reported AEs alone, as the frequency of negative events appears to be underestimated when compared to parent-reported SEs. Over the past years, the importance of including patient-reported outcomes has been increasingly recognized, and these measures are now part of most registries; however, SEs reported by parents are not routinely included in clinical trials (3). Our study highlights the importance of adding these measures in future trials to better understand how families cope with medication SEs and the impact of these SEs on compliance. Our CAPRI registry is ongoing, and pragmatic trials that include patient-reported outcomes to better address parent/patients' concerns and voice about their care are underway. We believe it is very important to evaluate treatment experiences with disease burden when measuring outcomes, particularly given the high frequency of potentially distressing SEs observed in our study.

Our study has several limitations. The parent perceptions of medication SEs were not verified, nor was their attribution to any specific medication. Furthermore, physicians may have known about these parental concerns but did not report them because they were not actionable AEs. We did not ask parents if they were willing to accept the described SEs without intervention (i.e., nonactionable) because the medication was perceived as being beneficial in other ways. Therefore, no direct comparison is possible between parent-reported SEs and physician-reported actionable AEs. We did not ask patients themselves to report SEs; discrepancies between patients and parents have been reported for some patient-reported outcomes (5). Additionally, there are limitations in comparing our actionable AE rates to the rates reported in pharmaceutical trials because physicians were specifically instructed to only report events that required a medical action.

In conclusion, in this modern JIA inception cohort, parents of children with JIA reported a very high frequency of medication SEs that had a measurable effect on the parent's global assessment of well-being and on the patient's assessment of HRQoL. The most common SEs were GI and behavioral/psychiatric symptoms. Addressing medication SEs reported by patients and parents may improve the HRQoL of children with JIA. Good communication with families is key to avoiding dismissal of medication SEs that they feel are important. Studies developing and testing effective strategies to mitigate these SEs are needed.

ACKNOWLEDGMENTS

We acknowledge the contribution of the following CAPRI Registry investigators: Daniah Basodan, Gilles Boire, Roxana Bolaria. Nicholas Blanchette, Sarah Campillo, Mercedes Chan, Tania Cellucci, Anne-Laure Chetaille, Paul Dancey, Erkan Demirkaya, Muhammad Dhalla, Karen Duffy, Janet Ellsworth, Liane Heale, Kristin Houghton, Andrea Human, Nicole Johnson, Roman Jurençak, Claire LeBlanc, Lillian Lim, Lily Lim, Nadia Luca, Tara McGrath, Tara McMillan, Paivi Miettunen, Kim Morishita, Johannes Roth, Evelyn Rozenblyum, and Rosie Scuccimarri.

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. Chédeville 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. Chédeville, Cabral, Shiff, Schmeling, Gerhold, Duffy, Tucker, Guzman.

Acquisition of data. Chédeville, Cabral, Rumsey, Proulx-Gauthier, Schmeling, Berard, Batthish, Soon, Gerschman, Bruns, Duffy, Tucker, Guzman.

Analysis and interpretation of data. Chédeville, McGuire, Guzman.

REFERENCES

1. Shiff NJ, Lix LM, Joseph L, Duffy C, Tucker LB, Svenson LW, et al. The prevalence of systemic autoimmune rheumatic diseases in Canadian pediatric populations: administrative database estimates. Rheumatol Int 2015;35:569–73.
2. Weitzman ER, Wisk LE, Salimian PK, Magane KM, Dedeoglu F, Hersh AO, et al. Adding patient-reported outcomes to a multisite registry to quantify quality of life and experiences of disease and treatment for youth with juvenile idiopathic arthritis. J Patient Rep Outcomes 2018;2:1.
3. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan AW, King MT, et al. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. JAMA 2018;319:483–94.
4. Falvey S, Shipman L, Ilowite N, Beukelman T. Methotrexate-induced nausea in the treatment of juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2017;15:52.
5. Brandon TG, Becker BD, Bevans KB, Weiss PF. Patient-reported outcomes measurement information system tools for collecting patient-reported outcomes in children with juvenile arthritis. Arthritis Care Res (Hoboken) 2017;69:393–402.
6. Hersh AO, Salimian PK, Weitzman ER. Using patient-reported outcome measures to capture the patient's voice in research and care of juvenile idiopathic arthritis. Rheum Dis Clin North Am 2016;42:333–46.
7. Kimman ML, Wijsenbeek MS, van Kuijk SMJ, Wijnsma KL, van de Kar N, Storm M, et al. Validity of the Patient Experiences and Satisfaction with Medications (PESaM) Questionnaire. Patient 2019;12: 149–62.
8. Duffy CM, Arsenault L, Duffy KN, Paquin JD, Strawczynski H. The Juvenile Arthritis Quality of Life Questionnaire: development of a new responsive index for juvenile rheumatoid arthritis and juvenile spondyloarthritides. J Rheumatol 1997;24:738–46.
9. Filocamo G, Consolaro A, Schiappapietra B, Dalpra S, Lattanzi B, Magni-Manzoni S, et al. A new approach to clinical care of juvenile idiopathic arthritis: the Juvenile Arthritis Multidimensional Assessment Report. J Rheumatol 2011;38:938–53.
10. Batthish M, Berard R, Cabral D, Bolaria R, Chédeville G, Duffy C, et al. A new Canadian inception cohort for juvenile idiopathic arthritis: the Canadian Alliance of Pediatric Rheumatology Investigators Registry. Rheumatology (Oxford) 2020;59:2796–805.
11. Miettunen P, Chédeville G, Mazza J, Hofer M, Montobbio C, Consolaro A, et al. The Canadian English and French versions of the Juvenile Arthritis Multidimensional Assessment Report (JAMAR). Rheumatol Int 2018;38 Suppl 1:83–90.
12. Horneff G, Schmeling H, Biedermann T, Foeldvari I, Ganser G, Girschick HJ, et al. The German etanercept registry for treatment of juvenile idiopathic arthritis. Ann Rheum Dis 2004;63:1638–44.
13. Beukelman T, Kimura Y, Ilowite NT, Mieszkalski K, Natter MD, Burrell G, et al. The new Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry: design, rationale, and characteristics of patients enrolled in the first 12 months. Pediatr Rheumatol Online J 2017;15:30.
14. Gong GW, Young NL, Dempster H, Porepa M, Feldman BM. The Quality of My Life questionnaire: the minimal clinically important difference for pediatric rheumatology patients. J Rheumatol 2007;34: 581–7.
15. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 1994;37:1761–9.
16. Rabe-Hesketh S, Skrondal A. Multilevel and longitudinal modeling using Stata. 2nd ed. College Station (TX): Stata press; 2008.
17. Khan S, Mancini J, Hopper C, Rennick JE. Perceptions of methotrexate intolerance and its impact on daily life in school-age children with juvenile idiopathic arthritis. J Pediatr Nurs 2019;48:49–54.
18. Bulatović M, Heijstek MW, Verkaaik M, van Dijkhuizen EH, Armbrust W, Hoppenreijs EP, et al. High prevalence of methotrexate intolerance in juvenile idiopathic arthritis: development and validation of a methotrexate intolerance severity score. Arthritis Rheum 2011; 63:2007–13.
19. Van Dijkhuizen EH, Bulatovic Calasan M, Pluijm SM, de Rotte MC, Vastert SJ, Kamphuis S, et al. Prediction of methotrexate intolerance in juvenile idiopathic arthritis: a prospective, observational cohort study. Pediatr Rheumatol Online J 2015;13:5.
20. Brunner HI, Johnson AL, Barron AC, Passo MH, Griffin TA, Graham TB, et al. Gastrointestinal symptoms and their association with health-related quality of life of children with juvenile rheumatoid arthritis: validation of a gastrointestinal symptom questionnaire. J Clin Rheumatol 2005;11:194–204.
21. Foeldvari I, Szer IS, Zemel LS, Lovell DJ, Giannini EH, Robbins JL, et al. A prospective study comparing celecoxib with naproxen in children with juvenile rheumatoid arthritis. J Rheumatol 2009;36:174–82.
22. Ruperto N, Nikishina I, Pachanov ED, Shachbazian Y, Prieur AM, Mouy R, et al. A randomized, double-blind clinical trial of two doses of meloxicam compared with naproxen in children with juvenile idiopathic arthritis: short- and long-term efficacy and safety results. Arthritis Rheum 2005;52:563–72.
23. Lovell DJ, Dare JA, Francis-Sedlak M, Ball J, LaMoreaux BD, von Scheven E, et al. A 6-month, multicenter, open-label study of fixed dose naproxen/esomeprazole in adolescent patients with juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2018;16:41.
24. Colloca L, Barsky AJ. Placebo and nocebo effects. N Engl J Med 2020;382:554–61.
25. Benedetti F, Thoen W, Blanchard C, Vighetti S, Arduino C. Pain as a reward: changing the meaning of pain from negative to positive co-activates opioid and cannabinoid systems. Pain 2013;154:361–7.
26. Amanzio M, Corazzini LL, Vase L, Benedetti F. A systematic review of adverse events in placebo groups of anti-migraine clinical trials. Pain 2009;146:261–9.
27. Charuvanij S, Chaiyadech C. Health-related quality of life in children with early-stage juvenile idiopathic arthritis. Musculoskeletal Care 2019;17:215–20.
28. Stevanovic D, Susic G. Health-related quality of life and emotional problems in juvenile idiopathic arthritis. Qual Life Res 2013;22:607–12.
29. Burnett HF, Regier DA, Feldman BM, Miller FA, Ungar WJ. Parents' preferences for drug treatments in juvenile idiopathic arthritis: a discrete choice experiment. Arthritis Care Res (Hoboken) 2012;64: 1382–91.
30. Cooper V, Metcalf L, Versnel J, Upton J, Walker S, Horne R. Patient-reported side effects, concerns and adherence to corticosteroid treatment for asthma, and comparison with physician estimates of side-effect prevalence: a UK-wide, cross-sectional study. NPJ Prim Care Respir Med 2015;25:15026.
31. Gawert L, Hierse F, Zink A, Strangfeld A. How well do patient reports reflect adverse drug reactions reported by rheumatologists? Agreement of physician- and patient-reported adverse events in patients with rheumatoid arthritis observed in the German biologics register. Rheumatology (Oxford) 2011;50:152–60.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1575–1584
DOI 10.1002/acr.24609

American College of Rheumatology
Empowering Rheumatology Professionals

Differences Sustained Between Diffuse and Limited Forms of Juvenile Systemic Sclerosis in an Expanded International Cohort

Ivan Foeldvari,[1] Jens Klotsche,[2] Ozgur Kasapcopur,[3] Amra Adrovic,[3] Maria Teresa Terreri,[4] Ana Paula Sakamoto,[4] Valda Stanevicha,[5] Flavio Sztajnbok,[6] Jordi Anton,[7] Brian Feldman,[8] Ekaterina Alexeeva,[9] Maria Katsicas,[10] Vanessa Smith,[11] Tadej Avcin,[12] Edoardo Marrani,[13] Mikhail Kostik,[14] Thomas Lehman,[15] Walter-Alberto Sifuentes-Giraldo,[16] Natalia Vasquez-Canizares,[17] Simone Appenzeller,[18] Mahesh Janarthanan,[19] Monika Moll,[20] Dana Nemcova,[21] Anjali Patwardhan,[22] Maria Jose Santos,[23] Sujata Sawhney,[24] Dieneke Schonenberg-Meinema,[25] Cristina Battagliotti,[26] Lillemor Berntson,[27] Blanca Bica,[28] Jürgen Brunner,[29] Patricia Costa-Reis,[30] Despina Eleftheriou,[31] Liora Harel,[32] Gerd Horneff,[33] Daniela Kaiser,[34] Tilmann Kallinich,[35] Dragana Lazarevic,[36] Kirsten Minden,[35] Susan Nielsen,[37] Farzana Nuruzzaman,[38] Yosef Uziel,[39] Nicola Helmus,[1] and Kathryn S. Torok[40]

Objective. To evaluate the baseline clinical characteristics of juvenile systemic sclerosis (SSc) patients in the international juvenile SSc inception cohort, and to compare these characteristics between the classically defined juvenile diffuse cutaneous SSc (dcSSc) and limited cutaneous SSc (lcSSc) subtypes and among those with overlap features.

Methods. A cross-sectional study was performed using baseline visit data. Information on demographic characteristics, organ system evaluation, treatment, and patient- and physician-reported outcomes was extracted and summary statistics applied. Comparisons between juvenile dcSSc and lcSSc subtypes and patients with and without overlap features were performed using chi-square and Mann-Whitney U tests.

Results. At data extraction, 150 juvenile SSc patients were enrolled across 42 centers; 83% were White, 80% were female, juvenile dcSSc predominated (72%), and 17% of the cohort had overlap features. Significant differences were found between juvenile dcSSc and juvenile lcSSc regarding modified Rodnan skin thickness score, the presence of Gottron's papules, digital tip ulceration, results of the 6-minute walk test, and composite pulmonary and cardiac involvement. All of these were more frequent in dcSSc except for cardiac involvement. Juvenile dcSSc patients had significantly worse scores for physician-rated disease activity and damage. A significantly higher occurrence of Gottron's papules and musculoskeletal and composite pulmonary involvement, and a significantly lower frequency of Raynaud's phenomenon, were seen in those with overlap features.

Conclusion. Results from a large international juvenile SSc cohort demonstrate significant differences between juvenile dcSSc and juvenile lcSSc patients, including more globally severe disease and increased frequency of interstitial lung disease in juvenile dcSSc patients, while those with lcSSc have more frequent cardiac involvement. Those with overlap features had an unexpected higher frequency of interstitial lung disease.

INTRODUCTION

Juvenile systemic sclerosis (SSc) is a rare disease with an estimated prevalence of 3 in 1,000,000 children (1). Only a few publications are available summarizing clinical variables in larger cohorts of these patients (n > 50) (2–6). Limitations of prior publications include cross-sectional data collected retrospectively with chart review across centers (5,6), or patient data collected before

The inception cohort project is supported by an unrestricted grant from the Joachim Hertz Stiftung, Hamburg, Germany.

[1]Ivan Foeldvari, MD, Nicola Helmus: Hamburg Centre for Pediatric and Adolescent Rheumatology, Schön Klinik Hamburg Eilbek, Hamburg, Germany; [2]Jens Klotsche, MD: German Rheumatism Research Center, Berlin, Germany; [3]Ozgur Kasapcopur, MD, Amra Adrovic, MD: Cerrahpasa Medical School, Istanbul University–Cerrahpasa, Istanbul, Turkey; [4]Maria Teresa Terreri, MD, Ana Paula Sakamoto, MD: Universidade Federal de São Paulo, Sao Paulo, Brazil; [5]Valda Stanevicha, MD: Riga Stradins University, University Children Hospital, Riga, Latvia; [6]Flavio Sztajnbok, MD: Universidade do Estado, Rio de Janeiro, Brazil; [7]Jordi Anton, MD: Hospital Sant Joan de Déu, Esplugues, Universitat de Barcelona, Barcelona, Spain; [8]Brian Feldman, MD: SickKids, The Hospital for Sick Children, Toronto, Ontario, Canada; [9]Ekaterina Alexeeva, MD: National Medical Research Center of Children's Health, Sechenov First Moscow State Medical University of the Ministry of Health of the Russian Federation, Moscow, Russia; [10]Maria Katsicas, MD: Hospital de

SIGNIFICANCE & INNOVATIONS

- Juvenile systemic sclerosis (SSc) patients demonstrate significant differences between juvenile diffuse cutaneous SSc (dcSSc) and juvenile limited cutaneous SSc subtypes regarding frequency of skin, vascular, pulmonary, and cardiac involvement.
- Physician global assessment of disease activity and damage is higher in the juvenile dcSSc group.
- Juvenile SSc patients with overlap features are not protected from major internal organ involvement and have a higher frequency of lung disease compared to those without overlap features.

2006 when clinical evaluation and management was different from the current practice (4–6). This includes the standardly collected data in the Scalapino cohort (n = 111), with patient data collection between 1960 and 2003 (4). More recently, there are reports from 2 prospective registries for juvenile SSc patients: the Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry (2), and the juvenile SSc inception cohort (3), both with original description of baseline characteristics of n = 64 and n = 80 juvenile SSc subjects, respectively. One limitation of the recent CARRA cohort was the lack of designation, and therefore description, of limited versus diffuse cutaneous clinical phenotypes in juvenile SSc.

To overcome this limitation in our juvenile SSc inception cohort, organ systems manifestations were extensively captured and compared between limited and diffuse cutaneous subtypes because the extent of skin involvement has been universally accepted to categorize patients with adult-onset SSc (7) and has been strongly linked to certain organ manifestations, augmenting patient care guidance (8). For example, in adult-onset SSc, with the knowledge that scleroderma renal crisis (SRC) is strongly associated with diffuse cutaneous disease, clinicians will more closely monitor blood pressure in early disease and avoid prednisone when possible because it is a risk factor in developing SRC. An additional categorization of importance in which the frequency of organ manifestations requires further clarification in juvenile SSc is overlap SSc. These overlap SSc patients meet classification criteria for SSc but also display overlap features of other connective tissue diseases, such as dermatomyositis (4,9,10), and have been reported in higher frequency in prior juvenile SSc cohorts compared to adult SSc (4).

Since the publication of the original study (3), 70 additional subjects have been enrolled in the juvenile SSc cohort registry and are reported here. With 150 subjects enrolled currently, our cohort represents the largest juvenile SSc patient cohort, affording the opportunity to make comparisons between clinical and patient-reported variables across limited cutaneous, diffuse cutaneous, and overlap SSc in juvenile-onset disease. This enables us to build upon the original study comparing diffuse and limited cutaneous clinical features, as well as to provide the first study in juvenile SSc to systematically compare juvenile SSc patients with and without overlap features. Our overall objective is to determine if there are important associations of organ involvement and patient impact among these subtypes, which may ultimately influence patient evaluation and monitoring.

PATIENTS AND METHODS

The juvenile SSc inception cohort registry, as previously described (3), is an international prospective observational cohort

Pediatria J. P. Garrahan, Buenos Aires, Argentina; [11]Vanessa Smith, MD: Ghent University and Ghent University Hospital, Ghent, Belgium; [12]Tadej Avcin, MD: University Children's Hospital University Medical Center Ljubljana, Ljubljana, Slovenia; [13]Edoardo Marrani, MD: Meyer Children's Hospital, Florence, Italy; [14]Mikhail Kostik, MD: Saint-Petersburg State Pediatric Medical University, Saint-Petersburg, Russia; [15]Thomas Lehman, MD: Hospital for Special Surgery, New York, New York; [16]Walter-Alberto Sifuentes-Giraldo, MD: University Hospital Ramón y Cajal, Madrid, Spain; [17]Natalia Vasquez-Canizares, MD: Children's Hospital at Montefiore, Bronx, New York; [18]Simone Appenzeller, MD: School of Medical Science University of Campinas, Sao Paulo, Brazil; [19]Mahesh Janarthanan, MD: Sri Ramachandra University, Chennai, India; [20]Monika Moll, MD: University Tuebingen, Tuebingen, Germany; [21]Dana Nemcova, MD: Charles University, Prague, Czech Republic; [22]Anjali Patwardhan, MD: University of Missouri, Columbia; [23]Maria Jose Santos, MD: Hospital Garcia de Orta, Almada, Portugal; [24]Sujata Sawhney, MD: Sir Ganga Ram Hospital, Delhi, India; [25]Dieneke Schonenberg-Meinema, MD: Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands; [26]Cristina Battagliotti, MD: Hospital de Niños Dr. Orlando Alassia, Santa Fe, Argentina; [27]Lillemor Berntson, MD: Uppsala University, Uppsala, Sweden; [28]Blanca Bica, MD: Hospital Universitário Clementino Fraga Filho and Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; [29]Jürgen Brunner, MD: Medical University Innsbruck, Innsbruck, Austria; [30]Patricia Costa-Reis, MD: Hospital de Santa Maria, Universidade de Lisboa, Lisbon, Portugal; [31]Despina Eleftheriou, MD: Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK; [32]Liora Harel, MD: Schneider Children's Medical Center of Israel Sackler Faculty of Medicine, Tel Aviv University, Petah-Tikva, Israel; [33]Gerd Horneff, MD: Asklepios Klnik Sankt Augustin, Sankt Augustin, Germany; [34]Daniela Kaiser, MD: Luzerner Kantonsspital, Kinderspital, Luzern, Switzerland; [35]Tilmann Kallinich, MD, Kirsten Minden, MD: Charité University Medicine and German Rheumatism Research Center Berlin, Berlin, Germany; [36]Dragana Lazarevic, MD: Clinical Center Nis, University of Niš, Nis, Serbia; [37]Susan Nielsen, MD: Rigshospitalet, Copenhagen, Denmark; [38]Farzana Nuruzzaman, MD: Stony Brook Children's Hospital, Stony Brook, New York; [39]Yosef Uziel, MD: Meir Medical Center, Kfar Saba, Tel Aviv University, Tel Aviv, Israel; [40]Kathryn S. Torok, MD: University of Pittsburgh, Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania.

Dr. Santos has received speaking fees from AbbVie, Novartis, and Astra Zeneca (less than $10,000 each). Dr. Brunner has received speaking fees from Novartis (less than $10,000). Dr. Minden has received speaking fees from AbbVie, GSK, Pfizer, Roche, Novartis, and Sanofi (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Ivan Foeldvari, MD, Hamburg Centre for Pediatric and Adolescence Rheumatology, Centre for Treatment of Scleroderma and Uveitis in Childhood and Adolescence, Teaching Unit of the Asklepios Campus of the Semmelweis Medical School, Budapest, An der Schön Klinik Hamburg Eilbek, Dehnhaide 120, 22081 Hamburg, Germany (email: sperchstunde@kinderrheumatologie.de); or to Kathryn S. Torok, MD, Division of Pediatric Rheumatology, UPMC and University of Pittsburgh Scleroderma Center, Pediatric Scleroderma Clinic, University of Pittsburgh, UPMC Children's Hospital of Pittsburgh, 4401 Penn Avenue, Pittsburgh, PA 15224 (email: kathryn.torok@chp.edu).

Submitted for publication July 1, 2020; accepted in revised form March 25, 2021.

study, including 25 centers from Europe, 5 from Asia, 6 from North America, and 6 from South America, representing 42 academic institutions. All participating centers had the research protocol approved by their local ethics committee. We are presenting a cross-sectional analysis of the data obtained at the patients' baseline cohort visit.

The inclusion criteria of the juvenile SSc inception cohort registry required fulfilling classification criteria of SSc using the stricter pediatric provisional 2007 classification criteria (11) from January 2008 to September 2017 and, after an amendment from October 2017 modifying these criteria, the more inclusive 2013 American College of Rheumatology/European Alliance of Associations for Rheumatology adult classification criteria for SSc (12), which allow for earlier detection of disease, with gaining points for sclerodactyly, and do not require the progression of skin thickness beyond the metacarpophalangeal joints, which was a limitation of the preliminary pediatric classification criteria, in the authors' opinion (IF and KST). The other criteria were unchanged throughout the study and include the following: age of <16 years at the time of the first non–Raynaud's phenomenon sign of disease, and <18 years at the time of the enrollment.

Data collection from juvenile SSc patients included demographic characteristics, results from physical examination, clinical testing variables, and physician- and patient-related outcome measures, as described in the original study of the first 80 subjects (3) (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609, which includes the clinical research form (CRF) obtained at the visits). Patients were scored for the presence and degree of skin thickness using the modified Rodnan skin thickness score (13), and cutaneous involvement was classified into diffuse and limited subtypes, with diffuse cutaneous defined by widespread and rapidly progressive skin thickening (starting at fingers and toes and spreading proximally beyond elbows and knees), and limited cutaneous characterized by restricted and nonprogressive skin thickening (starting at the fingertips and toes but limited to distal extremities, not crossing antecubital or popliteal fossa) (7). Data on the overlap subset of juvenile SSc were not collected independently, but overlap features were collected among juvenile diffuse cutaneous SSc (dcSSc) and juvenile limited cutaneous SSc (lcSSc) patients, including variables such as Gottron's papules, myositis, arthritis, and sicca symptoms.

In addition to the variables listed in our prior study (3), we created a "composite pulmonary involvement" variable, defined as meeting at least 1 of the following criteria: a forced vital capacity (FVC) of <80% of the predicted value; a diffusion capacity for carbon monoxide (DLco) of <80%; or high-resolution computed tomography (HRCT) findings consistent with interstitial lung disease (ILD). Moreover, digital ulcers were quite common in our initial cohort assessment (3) and have an impact on daily life in our juvenile SSc patients (2,14); therefore, we have incorporated the Digital Ulcer Clinical Assessment Score (DUCAS) (15) as an outcome variable collected prospectively. Data were collected prospectively according to a standardized assessment protocol every 6 months (see Supplementary Appendix A, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609).

Statistical analysis. Data were extracted for patients enrolled from January 2008 to December 15, 2019. Only baseline visit enrollment data were analyzed for this report. Statistical analyses were conducted using SAS software, version 9.4. Categorical variables were reported by absolute and relative frequencies and continuously distributed variables by median and interquartile range (25th and 75th percentile). Comparisons between patients with diffuse and limited cutaneous involvement and those with and without overlap features were performed using chi-square test and Fisher's exact when appropriate for categorical variables and Mann-Whitney U test for continuously distributed variables. *P* values less than 0.05 were considered statistically significant.

RESULTS

Patient demographic, autoantibody, and laboratory findings. At the time of data query, 150 patients were enrolled in the juvenile SSc inception cohort across the 42 academic institutions, with the majority being White (83%) and female (80%) (Table 1). All patients who fulfilled the pediatric SSc classification criteria (11) fulfilled the adult SSc criteria (12), which were applied for the inclusion since October 2017. Ninety-seven patients in this cohort were included before the amendment. The diffuse cutaneous subtype was predominant (72%) compared to juvenile lcSSc (28%). Overlap features were present in 17% of the cohort, with higher frequency in juvenile lcSSc compared to juvenile dcSSc (n = 12, 28% versus n = 14, 13%, respectively; $P = 0.023$). Although slightly younger in the juvenile dcSSc group compared to the juvenile lcSSc group, the median age at onset of Raynaud's phenomenon and first non–Raynaud's phenomenon symptom was not significantly different between the cutaneous subtypes (10.3 versus 11.9 years, and 10.7 versus 13.1 years, respectively). Median disease duration at the time of enrollment was 2.6 years in the juvenile dcSSc group and 1.8 years in the juvenile lcSSc group ($P = 0.038$). The majority (81%) of juvenile SSc patients were being treated with disease-modifying agents regardless of subtype (Table 1). Evaluation of autoantibodies supported antinuclear antibody (ANA) positivity in 91% of the cohort, with a similar distribution of antibodies against extractable nuclear antigens (Scl-70 and centromere) between the 2 cutaneous subtypes, reflecting the findings from the original 80 patients described (3). Specifically, anti–Scl-70 positivity was found in approximately one-third of the cohort (35% in juvenile dcSSc versus 36% in juvenile lcSSc), and anticentromere positivity was found at a very low rate in both subsets (3% versus 7%, respectively) (Table 1). Anti-PM/Scl antibody, reflecting overlap disease,

Table 1. Demographic information, disease characteristics, and autoantibody and laboratory measures of the 150 juvenile systemic sclerosis patients in the cohort compared by cutaneous subtype*

	Whole group (n = 150)	Diffuse subtype (n = 108)	Limited subtype (n = 42)	P†
Female to male ratio‡	4.2:1 (121/29)	4.1:1 (87/21)	4.2:1 (34/8)	0.571
Race, no. (%)				0.871
White	124 (83)	90 (83)	34 (81)	
African	9 (6)	8 (7)	1 (2)	
Indian	9 (6)	3 (3)	6 (14)	
Other	8 (5)	7 (6)	1 (2)	
Disease duration, median (IQR) years	2.4 (0.9–4.4)	2.6 (1.3–4.8)	1.8 (0.6–4.1)	0.038
Age at onset of RP, median (IQR) years	10.8 (6.9–13.1)	10.3 (7.0–12.8)	11.9 (6.3–13.9)	0.139
Age at onset of non-RP, median (IQR) years	11.1 (6.9–13.5)	10.7 (7.0–12.7)	13.1 (6.8–14.5)	0.091
Disease-modifying drugs, no. (%)	122 (81)	86 (80)	36 (86)	0.390
Autoantibody positivity				
ANA	133/146 (91)	95/104 (91)	38 (90)	0.867
Anti–Scl-70	51/145 (35)	36/103 (35)	15 (36)	0.930
Anticentromere	4/97 (4)	2/68 (3)	2/29 (7)	0.370
Anti-PM/Scl	9/57 (16)	5/37 (14)	4/20 (20)	0.522
Laboratory values				
ESR elevated (>20 mm/hour)	36/141 (25)	30/103 (29)	6/38 (16)	0.107
CRP elevated (>5 mg/liter)	17/127 (13)	14/94 (15)	3/33 (9)	0.400
Elevated CK	23/102 (22)	19/72 (26)	4/30 (13)	0.151
Elevated CK in overlap patients	2/23 (9)	1/13 (8)	1/10 (10)	0.846
Pro-BNP increased	4/17 (23)	3/13 (23)	1/4 (15)	0.937

* Values are the no./total no. (%) unless indicated otherwise. ANA = antinuclear antibody; BNP = B-type natriuretic peptide; CK = creatine kinase; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; IQR = interquartile range (25th, 75th percentiles); RP = Raynaud's phenomenon.
† Comparison between diffuse and limited subtypes.
‡ Female n = 121, male n = 29.

was similarly present in juvenile dcSSc and juvenile lcSSc (14% and 20%, respectively), with higher frequency in those with overlap compared to those without (31% versus 10%, respectively; $P = 0.046$) (Table 2). Additional comparison of laboratory evaluation included inflammatory markers, such as erythrocyte sedimentation rate, which was elevated in 29% of the juvenile dcSSc group and 16% of the juvenile lcSSc group ($P = 0.107$). C-reactive protein elevation was less frequently encountered (15% in juvenile dcSSc, and 9% in juvenile lcSSc; $P = 0.40$) (Table 1). The patients with overlap features had similar frequencies as the juvenile dcSSc group for these variables (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609).

Clinical features. The summary of clinical features in the total cohort and between the diffuse and limited cutaneous subtypes is presented in Table 3. Several organ system outcomes had a more frequent occurrence in the juvenile dcSSc subset compared to juvenile lcSSc patients, with significant differences in the median modified Rodnan skin thickness score (17.0 in juvenile dcSSc versus 4.5 in juvenile lcSSc; $P < 0.001$), sclerodactyly (83% versus 66%; $P = 0.029$), and Gottron's papules (30% versus 13%; $P = 0.043$) for cutaneous organ involvement.

Regarding microvascular involvement, Raynaud's phenomenon was similar in both subgroups (~90%), but presence of telangiectasia was more frequent in juvenile dcSSc (42% versus 18%; $P = 0.01$), as well as history of ulceration (56% versus 32%; $P = 0.008$). The DUCAS score (15) reflecting ulceration severity did not statistically significant differ ($P = 0.147$) between juvenile dcSSc patients (median 0 [interquartile range (IQR) 0–0.25] compared to 0 [IQR 0–0], respectively) (Table 3). However, ~80% (46 of 58) of patients had a DUCAS score of 0 (juvenile lcSSc 93% versus juvenile dcSSc 75%).

Cardiopulmonary assessment demonstrated some differences between cutaneous subtypes, with more pulmonary morbidity in the juvenile dcSSc group and more cardiac morbidity in the juvenile lcSSc group (Table 3). The pulmonary parameters were more frequently abnormal in the juvenile dcSSc subtype, including FVC <80%, DLco <80%, and ILD findings on HRCT. Although not statistically significant individually, combining these factors in the composite pulmonary involvement variable, juvenile dcSSc demonstrated more pulmonary involvement than juvenile lcSSc in a statistically and clinically significant manner (49% versus 31%; $P = 0.045$). In accordance with this finding, results from the 6-minute walk test were more frequently below the 10th percentile in juvenile dcSSc (85% versus 54%; $P = 0.044$). Cardiac involvement overall was relatively infrequent in the cohort (6%), but when it occurred, it was more frequent in the juvenile lcSSc group (17% versus 2%; $P = 0.002$). The majority of patients did have cardiac screening, with electrocardiography conducted in 80% of the patients (78% in juvenile dcSSc, and 86% in juvenile lcSSc) and transthoracic echocardiography in 64% of the patients

Table 2. Main differences between clinical manifestations of the 150 juvenile systemic sclerosis patients in the cohort compared by overlap features*

	Patients without overlap (n = 124)	Patients with overlap (n = 26)	P†
Autoantibody positivity			
Anti–Scl-70	48/120 (40)	3/25 (12)	0.008‡
Anticentromere	4/81 (5)	0/18 (0)	0.336
Anti-PM/Scl	4/41 (10)	5/16 (31)	0.046‡
Cutaneous			
Gottron's papules, no. (%)	26 (21)	11/23 (48)	0.022‡
Pulmonary			
Composite pulmonary involvement, no. (%)	50 (40)	16 (61)	0.048‡
Renal			
Renal involvement assessed by urine test, no. (%)	5 (4)	2 (8)	0.421
Musculoskeletal			
Presence of swollen joints, no. (%)	18 (14)	11 (42)	0.001‡
Muscle weakness	17/102 (17)	10/22 (45)	0.003‡
Tendon friction tub	7/114 (6)	4/25 (16)	0.098
Patient reported			
Patient RP activity, median (IQR)	30 (10–55) (n = 108)	2.5 (0–40) (n = 22)	0.045‡
C-HAQ score, median (IQR)	0.25 (0–0.63) (n = 75)	0.5 (0–1) (n = 19)	0.097
C-HAQ score, mean (range)§	0.5 (0–2.6)	0.7 (0–2.5)	0.097

* Values are the no./total no. (%) unless indicated otherwise. C-HAQ = Childhood Health Assessment Questionnaire; IQR = interquartile range (25th, 75th percentiles); RP = Raynaud's phenomenon.
† Comparison between with/without overlap.
‡ Significant.
§ The mean is also presented for comparison to other published pediatric rheumatic disease group data.

(62% in juvenile dcSSc, and 69% in juvenile lcSSc). Cardiac involvement was described in 5 patients with arrhythmia: one with tricuspid insufficiency and one with mitral regurgitation in the juvenile lcSSc group, and both patients in the juvenile dcSSc group with arrhythmia. Pulmonary hypertension, screened by transthoracic echocardiography, according to the pediatric guidelines (16,17), was uncommon and similar in both groups (juvenile dcSSc: n = 7, 6%; juvenile lcSSc: n = 2, 5%) (Table 3). Primary versus secondary pulmonary arterial hypertension (PAH) was not designated by the treating physician in the CRF, but the status of ILD was recorded. Of the 9 patients with PAH, 3 of the 7 in the dcSSc group and 1 of the 2 in the lcSSc group had no associated signs of ILD; therefore, 44% (4 of 9) of those with PAH would likely be designated as primary.

No history of renal crisis was detected at the time of enrollment in the cohort, and only 1 patient had arterial hypertension in the juvenile lcSSc group. In the juvenile dcSSc group, 5 patients had proteinuria, 4 of them with <500 mg/day, and the fifth with 1.1 gm/day. A renal biopsy was performed on the juvenile dcSSc patient with significant proteinuria, and class V lupus nephritis was identified; we considered this as an overlap feature. In the juvenile lcSSc group, 1 patient had microscopic hematuria and a proteinuria level of <500 mg/day, and the other patient had isolated microscopic hematuria.

Gastrointestinal involvement occurred in 42% of the juvenile dcSSc patients and 29% of the juvenile lcSSc patients ($P = 0.138$). Esophageal involvement was the most frequent manifestation in both groups, which occurred in 39% of the juvenile dcSSc group and 29% in the juvenile lcSSc group ($P = 0.898$) (Table 3).

Muscle weakness occurred in 18% of juvenile dcSSc patients and 31% in juvenile lcSSc patients ($P = 0.132$). In patients with overlap features in both subsets, 45% had muscle weakness. Tendon friction rub was infrequent and in the same range in both groups (9% in juvenile dcSSc, and 6% in juvenile lcSSc; $P = 0.54$). Joint contractures were observed in 48% of juvenile dcSSc patients and 43% of juvenile lcSSc patients ($P = 0.630$), and swollen joints were observed in 21% of juvenile dcSSc patients and 17% of juvenile lcSSc patients ($P = 0.630$) (Table 3). Neurologic involvement was seldom (3% of the cohort) and was most commonly associated with musculoskeletal entrapment, with all 3 dcSSc patients with neurologic involvement having Carpal tunnel syndrome, while the 2 lcSSc patients with neurologic involvement were more divergent, with one having demyelinating sensorimotor axonal polyneuropathy, and the other with headache.

Global assessments (physician and patient reported). Patients with juvenile dcSSc had significantly worse scores for physician global assessment of disease activity compared to juvenile lcSSc patients (visual analog scale [VAS] median score 37.5 versus 20 [range 0–100]; $P = 0.002$) and for physician global assessment of disease damage (VAS median score 30 versus 10 [range 0–100]; $P < 0.001$) (Table 4). Physician-rated ulceration activity was in the similar range (VAS median score 5 versus 0 [range 0–100]; $P = 0.113$). There was no statistically significant

Table 3. Clinical manifestations of the 150 juvenile systemic sclerosis patients in the cohort compared by cutaneous subtype*

	Whole group (n = 150)	Diffuse subtype (n = 108)	Limited subtype (n = 42)	P†
Overlap features, no. (%)	26 (17)	14 (13)	12 (28)	0.023‡
Cutaneous				
MRSS, median (IQR)	12.5 (5–22.5)	17 (9–27)	4.5 (0–10)	<0.001‡
Gottron's papules	37/142 (26)	32/105 (30)	5/37 (13)	0.043‡
Gottron's papules in overlap patients	11/23 (48)	9/11 (82)	2/12 (17)	0.002‡
Puffy fingers	39/126 (31)	29/90 (32)	10/36 (28)	0.626
Sclerodactyly	108/138 (78)	83/100 (83)	25/38 (66)	0.029‡
Vascular				
Raynaud's phenomenon, no. (%)	135 (90)	98 (91)	37 (88)	0.628
Nailfold capillary changes	101/141 (72)	70/99 (71)	31 (74)	0.709
Telangiectasia	56/128 (44)	38/90 (42)	7/38 (18)	0.010‡
History of ulceration	73/148 (49)	60/107 (56)	13/41 (32)	0.008‡
Active ulceration	22/148 (15)	17/107 (16)	5/41 (12)	0.572
DUCAS score, median (IQR)	0 (0–0)	0 (0–0.25)	0 (0–0)	0.147
Calcinosis	11/64 (17)	10/48 (21)	1/16 (6)	0.181
Pulmonary				
FVC <80%	33/106 (31)	27/78 (35)	6/28 (21)	0.196
DLco <80%	31/71 (44)	22/50 (44)	9/21 (43)	0.929
Abnormal findings on HRCT	46/110 (42)	37/82 (45)	9/28 (32)	0.229
6-minute walk test under the normal range§	29/38 (76)	23/27 (85)	6/11 (54)	0.044‡
Composite pulmonary involvement, no. (%)	66 (44)	53 (49)	13 (31)	0.045‡
Cardiac, no. (%)				
Cardiac involvement	9 (6)	2 (2)	7 (17)	0.002‡
Pulmonary hypertension assessed by US	9 (6)	7 (6)	2 (5)	0.691
Renal, no. (%)				
Renal involvement assessed by urinalysis	7 (5)	5 (5)	2 (5)	0.972
Hypertension assessed by RR	1 (1)	0 (0)	1 (2)	0.108
Renal crisis	0 (0)	0 (0)	0 (0)	–
Gastroenterology, no. (%)				
Total gastrointestinal involvement	57 (38)	45 (42)	12 (29)	0.138
Total esophageal involvement	54 (36)	42 (39)	12 (29)	0.898
Musculoskeletal				
Overall	92/149 (62)	66/107 (62)	26 (62)	0.929
Presence of swollen joints	29/149 (19)	22/107 (21)	7 (17)	0.606
Presence of joints with decreased range	81/149 (54)	60/107 (56)	21 (50)	0.540
Presence of joints with pain on motion	35/149 (23)	22/107 (21)	13 (31)	0.169
Contractures	69/148 (47)	51/106 (48)	18 (43)	0.630
Muscle weakness	27/124 (22)	17/92 (18)	10/32 (31)	0.132
Muscle weakness in overlap patients	10/22 (45)	5/11 (45)	5/11 (45)	–
Tendon friction rub	11/139 (8)	9/103 (9)	2/36 (6)	0.543
Neurologic involvement, no. (%)				
Overall neurologic involvement	5 (3)	3 (3)	2 (5)	0.543

* Values are the no./total no. (%) unless indicated otherwise. DUCAS = Digital Ulcer Clinical Assessment Score; FVC = functional vital capacity; HRCT = high-resolution computed tomography; IQR = interquartile range (25th, 75th percentiles); MRSS = modified Rodnan skin thickness score; DLco: diffusion capacity for carbon monoxide; RR = Riva Rocci (method); US = ultrasound.
† Comparison between diffuse and limited subtypes.
‡ Significant.
§ Less than 10 percentile of normal range.

difference in patient-rated global disease activity, global disease damage, Raynaud's phenomenon activity, and ulceration activity on a VAS (range 0–100) between diffuse and limited cutaneous subtypes. The mean score in the Childhood Health Assessment Questionnaire (C-HAQ) was 0.5 in juvenile dcSSc subjects, 0.4 in juvenile lcSSc subjects (P = 0.707; Table 4), and 0.7 in those with overlap features (Table 2 and Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609).

Patients with overlap features. Overlap features occurred in 17% (26 of 150) of all juvenile SSc patients, 13% in the diffuse cutaneous subtype group, and 28% in the limited cutaneous subtype group (Table 3). Those with overlap features had similar demographic characteristics (sex, race, and disease onset and duration; see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609) as those without but did have some notable clinical differences. Overlap patients showed characteristics of dermatomyositis in 23 cases,

Table 4. Patient- and physician-related outcomes of the 150 juvenile systemic sclerosis patients in the cohort compared by cutaneous subtype*

	Whole group (n = 150)	Diffuse subtype (n = 108)	Limited subtype (n = 42)	P†
Physician reported				
Physician global disease activity	30 (20–50) (n = 116)	37.5 (25–50) (n = 88)	20 (10–32.5) (n = 28)	0.002
Physician global disease damage	30 (15–45) (n = 115)	30 (20–50) (n = 88)	10 (5–25) (n = 27)	<0.001
Physician ulceration activity	0 (0–20) (n = 136)	5 (0–20) (n = 104)	0 (0–12.5) (n = 32)	0.113
Patient reported				
Patient global disease activity	40 (30–60) (n = 106)	40 (30–55) (n = 86)	50 (22.5–60) (n = 20)	0.964
Patient global disease damage	40 (20–60) (n = 105)	40 (20–60) (n = 85)	47.5 (5–60) (n = 20)	0.424
Patient RP activity	25 (5–55) (n = 130)	30 (10–55) (n = 102)	15 (0–50) (n = 28)	0.159
Patient ulceration activity	5 (0–30) (n = 131)	7.5 (0–30) (n = 102)	0 (0–25) (n = 29)	0.242
C-HAQ score	0.25 (0–0.75) (n = 94)	0.25 (0–0.63) (n = 68)	0.25 (0–0.75) (n = 26)	0.707
C-HAQ score, mean (range)‡	0.5 (0–2.6)	0.5 (0–2.6)	0.4 (0–2)	0.707

* Values are the median (IQR) unless indicated otherwise. All physician- and patient-reported measures are from a visual analog scale (0–100 mm, minimum to maximum). C-HAQ = Childhood Health Assessment Questionnaire; IQR = interquartile range (25th, 75th percentiles); RP = Raynaud's phenomenon.
† Comparison between diffuse and limited subtypes.
‡ The mean is also presented for comparison to other published pediatric rheumatic disease group data.

1 combined with Sjögren's syndrome, and 3 had juvenile arthritis characteristics. More frequent cutaneous and musculoskeletal manifestations in patients with overlap features include Gottron's papules, number of joints with swelling, decreased range of motion, joint contractures, and muscle weakness (Table 2 and Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609). Vascular features such as Raynaud's phenomenon occurred more commonly in the nonoverlap group (93 versus 77%; $P = 0.015$) (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609). Digital ulcer frequency was similar between those with and without overlap features (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609). ILD appeared more prevalent in those with overlap features, with the composite pulmonary involvement variables: FVC >80% and/or DLco of <80%, and/or abnormal findings on HRCT significantly more common in this group (61 versus 40%; $P = 0.048$) (Table 2 and Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609). The uncommon organ systems involved in juvenile SSc, including cardiac, renal, and neurologic, were similar in those with and without overlap features (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609). The overlap patients most commonly had positive ANA results without a specific extractable nuclear antigen, followed by positive PM/Scl (31%) and Scl-70 (12%) results, with no patients having positive anticentromere antibody results (Table 2 and Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609). Physician- and patient-reported outcomes were not significantly different between those with and without overlap characteristics besides patient rating of the Raynaud's phenomenon activity, which was significantly higher in the nonoverlap patients (30 versus 2.5; $P = 0.044$) (Table 2 and Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609). Although not statistically significant, C-HAQ score was more impacted in the overlap patients compared to nonoverlap patients (0.7 versus 0.5; $P = 0.097$) (Table 2 and Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24609).

DISCUSSION

We present the largest cohort of juvenile SSc patients with prospectively collected standardized clinical assessment. It is reassuring that the unique findings that we described in our previous study of this cohort (3) regarding the dominance of the juvenile dcSSc subtype and the unique distribution of the antibody pattern are further confirmed. The additional 70 patients enrolled since the prior publication (n = 150 versus n = 80) allow for the identification of additional cutaneous and vascular differences between juvenile dcSSc and juvenile lcSSc patients, in addition to enabling the characterization of overlap SSc patients. Patients with juvenile dcSSc have, as expected by definition, higher mean modified Rodnan skin thickness scores, but they also have a significantly higher rate of cutaneous and vascular features: sclerodactyly, Gottron's papules, history of ulceration, and presence of

telangiectasia. It was surprising that telangiectasias were not predominant in the limited cutaneous subtype, as one might expect in the classic teaching of CREST syndrome (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasias) in adult-onset SSc, the clinical phenotype of which is consistent with lcSSc (18). One possible explanation is that the disease duration on average was 1 year longer in the juvenile dcSSc subjects, allowing more time for telangiectasias to accumulate, and that the juvenile lcSSc subjects may approach a similar frequency with longer follow-up analyses. As demonstrated in our earlier study (3), juvenile dcSSc patients have the significantly higher rate of pulmonary involvement, evaluated using the composite pulmonary items or the 6-minute walk distance test. Cardiac involvement, although rare, is a major cause of morbidity in juvenile SSc and, confirming our earlier study, was significantly higher in the juvenile lcSSc group and therefore deserves particular attention in this cutaneous subtype. Overall disease severity, gauged by the physician global assessment of disease activity and damage, supports more impact on patients with juvenile dcSSc, who tend to have cumulative higher total organ morbidity. Efforts to decrease this cumulative burden are underway with the more liberal use of disease-modifying agents in SSc earlier on in the disease process in both adult- and pediatric-onset SSc (19–21).

In contrast to adult-onset SSc cohorts comparing large numbers of diffuse and limited cutaneous patient subsets, such as the European Scleroderma Trials and Research (EUSTAR) database and the Patient-Centered Intervention Network Cohort (SPIN), we did not find the increased frequency of the following variables in juvenile dcSSc that were demonstrated in adult dcSSc: male patients; positive Scl-70; renal crisis; joint contractures; tendon friction rub; and functional impairment (8,22,23). Similar to the frequency of clinical manifestations between the adult cohorts and our juvenile SSc cohort was the finding of more frequent pulmonary involvement in the diffuse subset. A main overall difference between our juvenile SSc cohort and these large adult SSc cohorts is the overall percentage of limited compared to diffuse cutaneous, in which lcSSc predominates in adults (60%) and dcSSc predominates in our juvenile SSc cohort (72%). One explanation for this difference may be the significantly longer disease duration upon cohort entry, with 11.7 years from the first non–Raynaud's phenomenon symptom in the SPIN cohort, and 6.4 years for lcSSc and 4.2 years for dcSSc in the EUSTAR cohort, allowing the capture of more lcSSc patients. This is in comparison to the median disease duration of 2.6 years in the diffuse subtype and 1.8 years in the limited subtype in our juvenile SSc cohort, which are relatively short in disease duration contrasted to adult-onset SSc and are similar in timing. The juvenile SSc inception cohort has been enrolling over the past 10 years, and this diffuse cutaneous predominance persists, suggesting that we are not necessarily missing the late-bloomer lcSSc patients; but indeed, pediatric-onset patients have a unique subset distribution at the beginning of the disease and a unique organ pattern presentation.

Overlap features occurred in 17% of the patients in our pediatric cohort, with the vast majority of the patients (88%) with dermatomyositis overlap, and the few others with Sjögren's syndrome (n = 1) and juvenile arthritis (n = 3) overlap. The general percentage of overlap SSc subtype in adult-onset SSc cohorts ranges approximately from 5% to 20% and includes overlap with the following connective tissue diseases (CTDs) (24): Sjögren's syndrome, polymyositis/dermatomyositis, rheumatoid arthritis, and systemic lupus erythematous (SLE), with varying dominant CTDs among published cohorts, although dermatomyositis and SLE were noted in younger adult onset (16–40 years) compared to older adult onset (≥40 years) (4,25–27). Another juvenile SSc cohort that categorized overlap patients (n = 32 of 110; 29%) also found juvenile dermatomyositis to predominate heavily, with 72% (23 of 32) (4) having results similar to our juvenile SSc inception cohort findings.

The most notable finding in our overlap patients is the higher risk for ILD compared to those without overlap features, with abnormal DLco and HRCT results being more common in this subgroup. This is important, as it contradicts more traditional teaching (27) that those with overlap disease possess a less severe phenotype, and it should instead prompt clinicians to be on higher alert for ILD and internal organ manifestations and not only focus on musculoskeletal, vascular, and cutaneous involvement. A recent study of a large German cohort of 3,240 adult-onset SSc patients specifically examined their registry patients with SSc overlap syndrome (10%; n = 325) and evaluated their organ frequency as well as trajectory and found that the patients with overlap syndromes had a higher risk of developing lung fibrosis and heart involvement compared to those with lcSSc, although less than those with dcSSc, and harbor an intermediate rate of cardiopulmonary progression between lcSSc and dcSSc (26). We are collecting longitudinal data to further study juvenile SSc overlap subtype trajectory compared to nonoverlap dcSSc and lcSSc. Those with overlap features indeed may be at risk for poorer outcomes and overall well-being. There does appear to be a significant impact on physical functioning in our patients with juvenile SSc overlap defined by the C-HAQ. The mean C-HAQ score of 0.7 in those with overlap features is higher than the mean C-HAQ score reported in the CARRA legacy registry cohort for juvenile-onset SLE (0.26), dermatomyositis (0.41), and juvenile arthritis (0.38) (2), which is most likely clinically relevant given the general floor effect of the C-HAQ, with low total score of 0–3.

Limitations of our study include missing data. Despite the use of a standardized assessment protocol, this is an observational cohort in which participating clinicians report according to their standard of care in juvenile SSc. Assessment of antibodies was at the physician's discretion and the capacity of the health system to assess them in routine care, and the lack of testing for all subjects could have influenced our interpretation. Performance of additional organ evaluation, such as esophageal manometry, was not mandatory due to the observational study design and

ethical reasons. In consequence, the results of specific organ manifestation screenings included a proportion of missing data and may be slightly biased toward patients with more severe organ involvement, but the stability of the observed organ involvement pattern between the 80 patients in our earlier study and the 150 patients of the present study is reassuring. Another limitation is the cross-sectional analysis of our cohort at cohort entry. Therefore, all results have to be interpreted with caution, and no causal inference should be drawn from our results.

In conclusion, we present the largest juvenile SSc patient population with a prospectively collected standardized assessment. The unique findings that we had previously published summarizing 80 patients of the cohort (3) persist for the increased cohort size (n = 150) and are similar to those of the other large, published cohorts (2,4,5). A few differences exist between juvenile dcSSc and juvenile lcSSc in children, such as increased frequency of ILD in juvenile dcSSc, and cardiac involvement in juvenile lcSSc, which should be noted for clinical screening and monitoring evaluation. Additionally, analyses of those with overlap features demonstrated expected cutaneous and musculoskeletal involvement but unexpected increased frequency of ILD. Future, longitudinal study of this cohort will determine if the juvenile dcSSc and juvenile lcSSc subtypes and patients with overlap features retain these organ manifestations or follow a different trajectory. Data on medications are also captured at every visit and will be documented in the longitudinal evaluation to query relationships between medication regimen and organ systems outcomes while we await traditional clinical trials in juvenile SSc, which are difficult due to the rarity of the disease. In addition to clinical phenotype, future collection of molecular markers in tandem may assist in further immunophenotype classification as being evaluated in adult-onset SSc (28,29).

AUTHOR CONTRIBUTIONS

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

Study conception and design. Foeldvari, Klotsche, Torok.

Acquisition of data. Foeldvari, Klotsche, Kasapcopur, Adrovic, Terreri, Sakamoto, Stanevicha, Sztajnbok, Anton, Feldman, Alexeeva, Katsicas, Smith, Avcin, Marrani, Kostik, Lehman, Sifuentes-Giraldo, Vasquez-Canizares, Appenzeller, Janarthanan, Moll, Nemcova, Patwardhan, Santos, Sawhney, Schonenberg-Meinema, Battagliotti, Berntson, Bica, Brunner, Costa-Reis, Eleftheriou, Harel, Horneff, Kaiser, Kallinich, Lazarevic, Minden, Nielsen, Nuruzzaman, Uziel, Helmus, Torok.

Analysis and interpretation of data. Foeldvari, Klotsche, Torok.

REFERENCES

1. Beukelman T, Xie F, Foeldvari I. Assessing the prevalence of juvenile systemic sclerosis in childhood using administrative claims data from the United States. J Scleroderma Relat Disord 2018;3:189–90.
2. Stevens BE, Torok KS, Li SC, Hershey N, Curran M, Higgins GC, et al. Clinical characteristics and factors associated with disability and impaired quality of life in children with juvenile systemic sclerosis: results from the Childhood Arthritis and Rheumatology Research Alliance Legacy Registry. Arthritis Care Res (Hoboken) 2018;70:1806–13.
3. Foeldvari I, Klotsche J, Torok KS, Kasapcopur O, Adrovic A, Stanevica V, et al. Characteristics of the first 80 patients at timepoint of first assessment included in the Juvenile Systemic Sclerosis Inception Cohort. J Scleroderma Relat Disord 2019;4:49–61.
4. Scalapino K, Arkachaisri T, Lucas M, Fertig N, Helfrich DJ, Londino AV Jr, et al. Childhood onset systemic sclerosis: classification, clinical and serologic features, and survival in comparison with adult onset disease. J Rheumatol 2006;33:1004–13.
5. Martini G, Foeldvari I, Russo R, Cuttica R, Eberhard A, Ravelli A, et al. Systemic sclerosis in childhood: clinical and immunologic features of 153 patients in an international database. Arthritis Rheum 2006;54:3971–8.
6. Foeldvari I, Zhavania M, Birdi N, Cuttica RJ, de Oliveira SH, Dent PB, et al. Favourable outcome in 135 children with juvenile systemic sclerosis: results of a multi-national survey. Rheumatology (Oxford) 2000;39:556–9.
7. LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. J Rheumatol 2001;28:1573–6.
8. Allanore Y. Limited cutaneous systemic sclerosis: the unfairly neglected subset. J Scleroderma Relat Disord 2016;1:241–6.
9. Foeldvari I, Nihtyanova SI, Wierk A, Denton CP. Characteristics of patients with juvenile onset systemic sclerosis in an adult single-center cohort. J Rheumatol 2010;37:2422–6.
10. Li SC. Scleroderma in children and adolescents: localized scleroderma and systemic sclerosis. Pediatr Clin North Am 2018;65:757–81.
11. Zulian F, Woo P, Athreya BH, Laxer RM, Medsger TA Jr, Lehman TJ, et al. The Pediatric Rheumatology European Society/American College of Rheumatology/European League Against Rheumatism provisional classification criteria for juvenile systemic sclerosis. Arthritis Rheum 2007;57:203–12.
12. Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
13. Rodnan GP, Lipinski E, Luksick J. Skin thickness and collagen content in progressive systemic sclerosis and localized scleroderma. Arthritis Rheum 1979;22:130–40.
14. Guedes M, Zilhao C, Almeida I, Silva I. Raynaud and digital ulcers in patients with juvenile systemic sclerosis: ambulatory iloprost protocol. A single center experience. Pediatr Rheumatol Online J 2014;12 Suppl 1:P308.
15. Bruni C, Ngcozana T, Braschi F, Pucci T, Piemonte G, Benelli L, et al. Preliminary validation of the digital ulcer clinical assessment score in systemic sclerosis. J Rheumatol 2019;46:603–8.
16. Hansmann G, Koestenberger M, Alastalo TP, Apitz C, Austin ED, Bonnet D, et al. 2019 updated consensus statement on the diagnosis and treatment of pediatric pulmonary hypertension: the European Pediatric Pulmonary Vascular Disease Network (EPPVDN), endorsed by AEPC, ESPR and ISHLT. J Heart Lung Transplant 2019;38:879–901.
17. Abman SH, Hansmann G, Archer SL, Ivy DD, Adatia I, Chung WK, et al. Pediatric pulmonary hypertension: guidelines from the American Heart Association and American Thoracic Society. Circulation 2015;132:2037–99.
18. Wollheim FA. Classification of systemic sclerosis: visions and reality. Rheumatology (Oxford) 2005;44:1212–6.

19. Kowal-Bielecka O, Fransen J, Avouac J, Becker M, Kulak A, Allanore Y, et al. Update of EULAR recommendations for the treatment of systemic sclerosis. Ann Rheum Dis 2017;76:1327–39.
20. Fernández-Codina A, Walker KM, Pope JE, Scleroderma Algorithm Group. Treatment algorithms for systemic sclerosis according to experts. Arthritis Rheumatol 2018;70:1820–8.
21. Roofeh D, Khanna D. Management of systemic sclerosis: the first five years. Curr Opin Rheumatol 2020;32:228–37.
22. Frantz C, Huscher D, Avouac J, Hachulla E, Balbir-Gurman A, Riemekasten G, et al. Outcomes of limited cutaneous systemic sclerosis patients: results on more than 12,000 patients from the EUSTAR database. Autoimmun Rev 2019:102452.
23. Dougherty DH, Kwakkenbos L, Carrier ME, Salazar G, Assassi S, Baron M, et al. The Scleroderma Patient-Centered Intervention Network cohort: baseline clinical features and comparison with other large scleroderma cohorts. Rheumatology (Oxford) 2018;57:1623–31.
24. Foeldvari I, Tyndall A, Zulian F, Muller-Ladner U, Czirjak L, Denton C, et al. Juvenile and young adult-onset systemic sclerosis share the same organ involvement in adulthood: data from the EUSTAR database. Rheumatology (Oxford) 2012;51:1832–7.
25. Pakozdi A, Nihtyanova S, Moinzadeh P, Ong VH, Black CM, Denton CP. Clinical and serological hallmarks of systemic sclerosis overlap syndromes. J Rheumatol 2011;38:2406–9.
26. Moinzadeh P, Aberer E, Ahmadi-Simab K, Blank N, Distler JH, Fierlbeck G, et al. Disease progression in systemic sclerosis-overlap syndrome is significantly different from limited and diffuse cutaneous systemic sclerosis. Ann Rheum Dis 2015;74:730–7.
27. Fairley JL, Hansen D, Proudman S, Sahhar J, Ngian GS, Walker J, et al. Clinical features of systemic sclerosis–mixed connective tissue disease and systemic sclerosis overlap syndromes. Arthritis Care Res (Hoboken) 2021;73:732–41.
28. Johnson S, Hinchcliff M, Asano Y. Controversies: molecular vs. clinical systemic sclerosis classification. J Scleroderma Relat Disord 2016;1:277–85.
29. Skaug B, Khanna D, Swindell WR, Hinchcliff ME, Frech TM, Steen VD, et al. Global skin gene expression analysis of early diffuse cutaneous systemic sclerosis shows a prominent innate and adaptive inflammatory profile. Ann Rheum Dis 2020;79:379–86.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1585–1592
DOI 10.1002/acr.24621

American College of Rheumatology
Empowering Rheumatology Professionals

Costs of Hospital-Associated Care for Patients With Juvenile Idiopathic Arthritis in the Dutch Health Care System

Michelle M. A. Kip,[1] Sytze de Roock,[2] Inge van den Berg,[3] Gillian Currie,[4] Deborah A. Marshall,[4] Luiza R. Grazziotin,[4] Marinka Twilt,[4] Rae S. M. Yeung,[5] Susanne M. Benseler,[4] Sebastiaan J. Vastert,[2] Nico Wulffraat,[2] Joost F. Swart,[2] and Maarten J. IJzerman[6]

Objective. The aim of this study was to quantify costs of hospital-associated care for juvenile idiopathic arthritis (JIA), provide insights in patient-level variation in costs, and investigate costs over time from the moment of JIA diagnosis. Results were reported for all JIA patients in general and by subtype.

Methods. This study was a single-center, retrospective analysis of prospective data from electronic medical records of children with JIA, ages 0–18 years, between April 1, 2011 and March 31, 2019. Patient characteristics (age, sex, JIA subtype) and hospital-based resource use (consultations, medication, radiology procedures, laboratory testing, surgeries, emergency department [ED] visits, hospital stays) were extracted and analyzed. Unit prices were obtained from Dutch reimbursement lists and pharmaceutical and hospital list prices.

Results. The analysis included 691 patients. The mean total cost of hospital care was €3,784/patient/year, of which €2,103 (55.6%) was attributable to medication. Other costs involved pediatric rheumatologist visits (€633/patient/year [16.7%]), hospital stays (€439/patient/year [11.6%]), other within-hospital specialist visits (€324/patient/year [8.6%]), radiology procedures (€119/patient/year [3.1%]), laboratory tests (€114/patient/year [3.0%]), surgeries (€46/patient/year [1.2%]), and ED visits (€6/patient/year [0.2%]). Mean annual total costs varied between JIA subtypes and between individuals and were the highest for systemic JIA (€7,772/patient/year). Over the treatment course, costs were the highest in the first month after JIA diagnosis.

Conclusion. Hospital care costs of JIA vary substantially between individuals, between subtypes, and over the treatment course. The highest annual costs were for systemic JIA, primarily attributable to medication (i.e., biologics). Costs of other hospital-associated care were comparable regardless of subtype.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is the most common chronic rheumatic disease in childhood, affecting ~1 in 1,000 children (1,2). The International League of Associations for Rheumatology classification distinguishes 7 categories of JIA, including systemic arthritis, oligoarthritis (which can be subdivided into persistent and extended oligoarthritis), rheumatoid factor (RF) negative polyarthritis, RF positive polyarthritis, psoriatic arthritis, enthesitis-related arthritis, and undifferentiated arthritis (3).

Early recognition and adequate clinical management of JIA is crucial to control inflammation, reduce pain, and prevent irreversible joint damage (4). Treatment of JIA is multifaceted, combining pharmacologic, physical, and occupational therapy with lifestyle modifications and psychosocial support (5). As a consequence, treatment costs are high (6–9). JIA is also associated with

Supported by the Canadian Institutes for Health Research (grant 381280), Genome Canada, ZonMw (The Netherlands), and the Reumafonds (The Netherlands).

[1]Michelle M. Kip, PhD: Wilhelmina Children's Hospital, Utrecht, The Netherlands, and University of Twente, Enschede, The Netherlands; [2]Sytze de Roock, PhD, Sebastiaan J. Vastert, MD, PhD, Nico Wulffraat, MD, PhD, Joost F. Swart, MD, PhD: Wilhelmina Children's Hospital and Utrecht University, Utrecht, The Netherlands; [3]Inge van den Berg, MSc: University of Twente, Enschede, The Netherlands; [4]Gillian Currie, PhD, Deborah A. Marshall, PhD, Luiza R. Grazziotin, MSc, Marinka Twilt, MD, MSCE, PhD, Susanne M. Benseler, MD, PhD: University of Calgary, Calgary, Alberta, Canada; [5]Rae S. M. Yeung, MD, PhD, FRCPC: University of Toronto, Toronto, Ontario, Canada; [6]Maarten J. IJzerman, PhD: University of Twente, Enschede, The Netherlands, and University of Melbourne, Melbourne, Australia.

Dr. Vastert has received research grants and/or consulting fees from Sobi and Novartis (less than $10,000 each). Dr. IJzerman has received consulting fees and unrestricted research grants from Illumina and consulting fees from RTI Health Solutions (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Maarten J. IJzerman, PhD, University of Twente, Technical Medical Centre, department HTSR, P.O. Box 217, 7500 AE Enschede. Email: m.j.ijzerman@utwente.nl.

Submitted for publication July 28, 2020; accepted in revised form April 8, 2021.

SIGNIFICANCE & INNOVATIONS

- To the best of our knowledge, this is the first study in the world to quantify costs of hospital-associated care in juvenile idiopathic arthritis (JIA) subtypes while simultaneously providing insights into patient-level variations in costs and trends in costs over the treatment course.
- This study provides high-level evidence that, when implementing personalized treatments, the costs of early, intensive treatment strategies in patients with severe JIA should be offset against its benefits and costs over the long term.

significant long-term issues, including the risk of long-term functional impairment, lower educational attainment (10), higher unemployment rates (10), and a lower quality of life (11–13). Consequently, JIA results in a high burden to the affected individual and to society.

To determine the burden of JIA to society, a first and critical step is to quantify JIA-related hospital-care resource use and associated costs, referred to as "hospital costs" in the remainder of this article. Although a body of evidence presenting hospital costs is available (14), the majority of these studies either do not distinguish between JIA subtypes, focus on 1 specific subtype, or do not consider costs at the individual patient level. Reporting hospital costs separately by JIA subtype is important because these subtypes differ in clinical and laboratory features, disease severity, and in the efficacy, type, and accompanying costs of pharmacologic treatments prescribed (3,15,16). In addition, substantial variation in disease severity and treatment response is observed even between patients with the same subtype (17,18). Thus, JIA is known for its personalized treatment and for its huge variation in treatment lines and sequences with different impact on health outcomes and costs. Therefore, the current study aims to quantify the impact of JIA on hospital costs, provide insights in patient-level variation in costs, and investigate costs over time from the moment of JIA diagnosis. Results were reported for all JIA patients and by JIA subtype.

MATERIALS AND METHODS

Data sources and extraction. This study was a retrospective analysis of prospective data extracted from electronic medical records from the Wilhelmina Children's Hospital (Utrecht, The Netherlands), using a previously developed research data platform (19). This resulted in a comprehensive data set enabled by linkage of several databases within the hospital through a unique, deidentified patient number. For the current study, data on medication use, radiology procedures, laboratory tests, hospital stays, surgeries, consultations with pediatric rheumatologists and other within-hospital specialists, and emergency department (ED) visits were extracted for all patients with a diagnosis of JIA between April 1, 2011 and March 31, 2019. As treatment strategies in JIA change rapidly, and because the electronic data was available after April 1, 2011, this date was set as the starting point of the analyses. In addition, as this study focuses on children, only data up until the patient's 18th birthday were included when they turned 18 before March 31, 2019. Data on within-hospital physician visits (other than pediatric rheumatologist visits) were, however, only available up to December 12, 2018 and thus were included until that point in time.

The use of data from the above-mentioned research data platform was classified by the Institutional Review Board as exempt from the Medical Research Involving Human Subjects Act (14/684). The study was conducted according to Good Clinical Practice guidelines and the Declaration of Helsinki (20). Further, the ethical committee of the faculty of Behavioural, Management and Social Sciences of the University of Twente approved this study (no. 190215).

Data selection. Patients were excluded if they reached the age of 18 years before April 1, 2011, were diagnosed with idiopathic uveitis, were not primarily treated in the Wilhelmina Children's Hospital (because they, for example, only came for a second opinion), had major comorbidities (such as inflammatory bowel disease) alongside JIA, received treatment as part of a pharmaceutical trial that they would not have received outside the trial setting (regardless of whether this occurred between April 1, 2011 and March 31, 2019), or had a follow-up in <1 year. Resource use and costs were included up to 10 years after JIA diagnosis.

Resource use and costs. Within-hospital resource use and costs were quantified from a payer's perspective. Resource use was measured from the extracted data. Unit prices were based on 2019 tariffs when available (regardless of the year in which they occurred) or converted to 2019 euros using Dutch consumer price indices (21). Medication costs were derived from Dutch pharmaceutical list prices (https://www.farmacotherapeutischkompas.nl/) and multiplied with the frequency of use, accounting for the dose used in each individual patient. All other costs were derived from Dutch reimbursement lists where possible (https://zorgproducten.nza.nl/) or, alternatively, from hospital list prices. All within-hospital costs that occurred during the inclusion period and that were assumed to be JIA-related, as decided in consultation with a pediatric rheumatologist, were included. To illustrate this, the costs of hospital stays related to (for example) a sports injury were excluded, whereas the costs of hospital stays related to JIA (for example, for treatment of disease flares or for complications related to JIA treatment, such as infections), were included. A detailed overview of all assumptions made is provided (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24621).

Analysis. Results were presented for all JIA patients in general and by JIA subtype. Patients with persistent oligoarthritis were further subdivided into antinuclear antibody (ANA) negative oligoarthritis and ANA positive oligoarthritis cohorts. Costs over the period of follow-up were reported as costs/patient/year in the years following JIA diagnosis. In other words, for a patient diagnosed with JIA on April 6, 2012, the first year of follow-up spans the time between April 6, 2012 and April 6, 2013. As a consequence, hospital-related costs immediately after JIA diagnosis were unavailable for patients diagnosed before April 1, 2011. These patients were, however, included in the calculation of hospital-related costs up to 10 years after JIA diagnosis. The analysis was performed in R (version 3.5.3) using the packages dplyr, ggplot2, lubridate, and plotrix (22–26). Patients and/or the public were not involved in the design, conduct, reporting, or disseminating of the results of this study.

RESULTS

A total of 691 patients were included in the study, including 447 girls (65%) and 244 (35%) boys, with a median age at diagnosis of 8 years and a median duration of follow-up of 4.9 years. The study excluded 278 patients (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24621). Table 1 shows a detailed overview of patient characteristics.

The impact of JIA on hospital costs. The overall mean hospital cost of JIA was €3,784/patient/year, of which €2,103 (55.6%) was attributable to medication costs and €1,681 (44.4%) to costs of other hospital-based services. Hospital costs varied substantially between subtypes, with the highest mean costs in systemic JIA (€7,772/patient/year), followed by RF+ polyarticular JIA (€6,906/patient/year). When multiplying the costs per patient with the number of patients in each JIA subgroup, patients with polyarticular RF– JIA (n = 144) contributed most to the hospital costs (i.e., 25.4%). A detailed overview, including the distribution of costs of other hospital-based services into subcategories, is shown in Table 2.

Table 1. Characteristics of patients included in the analysis*

Total number	691 (100)
Age at JIA diagnosis, median (IQR) years	8.0 (4.0–12.6)
Duration of follow-up, median (IQR) years	4.9 (2.8–7.0)
Male sex	244 (35%)
JIA subtype	
Oligoarticular persistent JIA	294 (42.5)
ANA–	147 (21.3)
ANA+	147 (21.3)
Polyarticular JIA	175 (25.3)
RF–	144 (20.8)
RF+	31 (4.5)
Extended oligoarticular JIA	70 (10.1)
Enthesitis-related JIA	59 (8.5)
Systemic JIA	57 (8.2)
Psoriatic arthritis	29 (4.2)
JIA undifferentiated	7 (1.0)

* Values are the number (%) unless indicated otherwise. ANA = antinuclear antibody; IQR = interquartile range; JIA = juvenile idiopathic arthritis; RF = rheumatoid factor.

Variation in costs between individual patients. As the mean annual costs presented in Table 2 differed substantially on an individual patient level, the mean annual costs per patient (over their entire follow-up period) were visualized in a histogram, resulting in a strongly right skewed distribution (Figure 1). More specifically, 471 (68.2%) of 691 patients had mean annual costs ranging between €0/patient/year and €2,500/patient/year, and only 11 patients had mean annual costs of €25,300 or higher. Eight of these patients with mean annual costs between ~€31,000 and ~€119,000 were not shown as these were out of range in Figure 1. These costs involved 8 patients with systemic JIA or polyarticular RF+ JIA in which high costs were mainly attributable to medication use (i.e., canakinumab [n < 5; the exact number is not provided in order to prevent traceability of the study's results to individual patients] and/or intravenous tocilizumab [n = 6]) and to hospital stays.

When considering the histograms for the different types of hospital-based services (including medication), similar right-skewed distributions were observed (see Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24621). The only category in which the distribution was more evenly distributed involved costs of consultations with pediatric rheumatologists. As a fixed cost of €159.94 per consultation was applied in the current analysis (regardless of the duration of the appointment), this figure also represents the distribution of how frequently JIA patients visited a pediatric rheumatologist. Therefore, this data indicates that consultations with pediatric rheumatologists also occurred in patients who, on average, had low JIA-related hospital costs.

Variation of costs over the course of JIA treatment. Figure 2 shows the mean monthly total hospital costs for JIA treatment over 10 years of follow-up. Each point in the graph represents the mean monthly hospital costs when taking the average over the patients for whom data was available for each of the time periods during the 120 months of follow-up (with 0 representing the moment of JIA diagnosis).

This figure demonstrates that the mean monthly total hospital costs peaked in the month following JIA diagnosis (i.e., €913) and tended to decrease over the course of follow-up. When excluding costs of medication use, this decrease was more pronounced (see Supplementary Figure 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24621),

Table 2. Overview of mean annual hospital costs per patient*

						Mean annual costs of other hospital-based services/patient (% of costs of other hospital-based services)						
	No.	Mean annual hospital costs/patient (% of population costs)	Minimum–maximum	Mean annual medication costs/patient (%)	Mean annual costs of other hospital-based services/patient (%)	Pediatric rheumatologist visits	Hospital stay	Other within-hospital specialist visits	Radiology procedures	Laboratory testing	Surgeries	ED visits
All patients	691	€3,784 (100.0)	€0–166,789	€2,103 (55.6)	€1,681 (44.4)	€633 (37.7)	€439 (26.1)	€324 (19.3)	€119 (7.1)	€114 (6.8)	€46 (2.7)	€6 (0.4)
Systemic JIA	57	€7,772 (17.0)	€0–166,789	€4,790 (61.6)	€2,981 (38.4)	€691 (23.2)	€1,685 (56.5)	€191 (6.4)	€124 (4.2)	€233 (7.8)	€48 (1.6)	€11 (0.4)
Polyarticular RF+ JIA	31	€6,906 (8.2)	€0–51,144	€5,020 (72.7)	€1,886 (27.3)	€811 (43)	€458 (24.3)	€252 (13.4)	€187 (9.9)	€168 (8.9)	€2 (0.1)	€8 (0.4)
Psoriatic arthritis	29	€4,945 (5.5)	€0–40,264	€3,300 (66.7)	€1,644 (33.3)	€709 (43.1)	€353 (21.5)	€360 (21.9)	€87 (5.3)	€124 (7.5)	€7 (0.4)	€5 (0.3)
Polyarticular RF– JIA	144	€4,592 (25.4)	€0–39,687	€2,589 (56.4)	€2,003 (43.6)	€695 (34.7)	€574 (28.7)	€375 (18.7)	€151 (7.5)	€132 (6.6)	€69 (3.5)	€6 (0.3)
Extended oligoarticular JIA	70	€4,477 (12.0)	€0–26,827	€2,723 (60.8)	€1,754 (39.2)	€797 (45.5)	€240 (13.7)	€391 (22.3)	€154 (8.8)	€129 (7.4)	€33 (1.9)	€11 (0.6)
Enthesitis-related JIA	59	€4,100 (9.3)	€0–24,914	€2,606 (63.6)	€1,494 (36.4)	€710 (47.5)	€199 (13.3)	€226 (15.2)	€205 (13.7)	€121 (8.1)	€16 (1.1)	€17 (1.1)
Oligoarticular ANA+ JIA	147	€2,340 (13.2)	€0–21,349	€893 (38.2)	€1,447 (61.8)	€587 (40.6)	€231 (15.9)	€409 (28.2)	€78 (5.4)	€82 (5.7)	€59 (4.1)	€1 (0.1)
Oligoarticular ANA– JIA	147	€1,653 (9.3)	€0–29,315	€588 (35.6)	€1,065 (64.4)	€460 (43.2)	€160 (15.1)	€259 (24.3)	€83 (7.8)	€59 (5.5)	€40 (3.7)	€4 (0.4)
Undifferentiated JIA	7	€432 (0.1)	€0–2,602	€25 (5.8)	€398 (92.1)	€271 (68.0)	€0 (0.0)	€41 (10.2)	€46 (11.5)	€41 (10.3)	€0 (0.0)	€0 (0.0)

* ANA = antinuclear antibody; ED = emergency department; JIA = juvenile idiopathic arthritis; RF = rheumatoid factor.

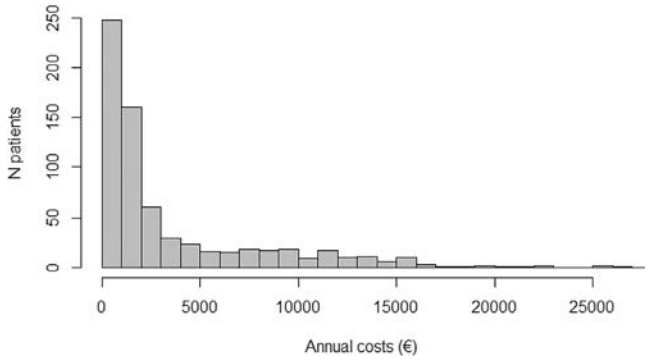

Figure 1. Histogram of the distribution of mean annual total hospital costs (including medication) per patient over the period of follow-up for each individual patient in the database, regardless of juvenile idiopathic arthritis (JIA) subtype. Eight patients (with systemic JIA or polyarticular rheumatoid factor–positive JIA) with average annual costs ranging from ~€31,000 to ~€119,000 are not shown as associated data were out of range of this figure. N = number.

attributable to the fact that costs of medication use peaked after ~25 months of follow-up. When plotting the costs for the other types of hospital-based services over time, results show a peak in costs at the time of JIA diagnosis for costs of hospital stay, consultations with pediatric rheumatologists and other within-hospital specialists, as well as for radiology and laboratory testing (see Supplementary Figure 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24621). In addition, regardless of the duration of follow-up, the majority of costs were attributable to hospital stays and consultations with pediatric rheumatologists and other within-hospital specialists, whereas costs of laboratory testing, radiology, surgeries, and ED visits only contributed to a minority of these costs.

Finally, our study demonstrates that, although costs of hospital-associated care for JIA treatment may decrease over time, costs of systemic JIA tended to peak after ~25 months of follow-up (attributable to the <5 patients who received canakinumab), which explains the peak in medication costs at this time point (see Supplementary Figure 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24621). The mean monthly total hospital costs specified according to JIA subtype are also shown. A detailed analysis of medication use and their accompanying costs, however, falls outside the scope of the current analysis but has been described in another study (27).

DISCUSSION

The overall mean hospital costs of JIA were €3,784/patient/year, of which 55.6% was attributable to medication use. These costs varied considerably between patients. Systemic JIA patients incurred (on average) the highest annual costs, which were primarily attributable to medication use and secondarily to hospital stays. The majority of the costs for hospital stays for systemic JIA patients occurred within the first month after diagnosis. Costs of other hospital-based services, like specialist consultations, laboratory testing, radiology procedures, ED visits, and surgeries were comparable between JIA subtypes (except for undifferentiated JIA). In contrast to medication costs, costs of other hospital-based services peaked in the first month after JIA diagnosis and decreased over time.

The annual hospital costs reported in this study were relatively low because, in contrast to other cost studies in JIA (6,8,9,28), this study included all patients regardless of their disease state. In line with our results, Minden et al reported that costs

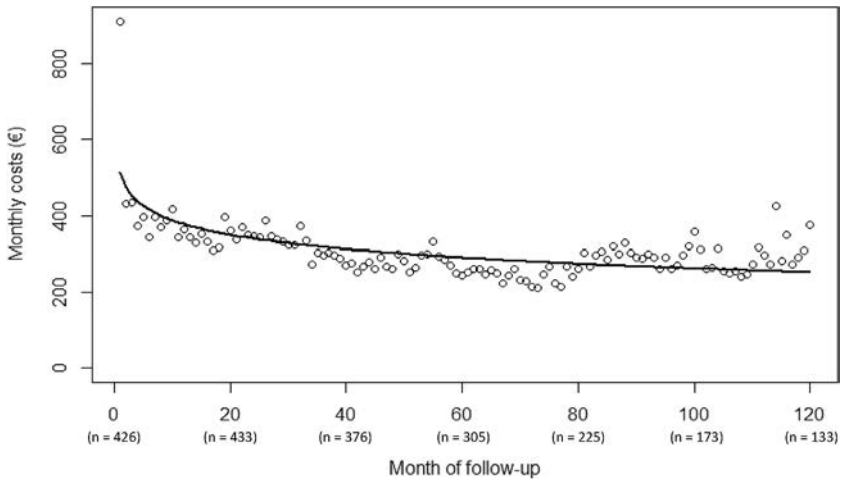

Figure 2. Mean monthly total hospital costs per patient (including medication) over the course of follow-up. Circles represent the mean costs for the set of patients for which data was available during the different months of follow-up during the 120 months. Logistic regression was used to fit a line through the data points. 0 = the moment of juvenile idiopathic arthritis diagnosis.

vary strongly depending on patients' disease state, with patients with active disease having mean annual costs of €5,681 compared with €782 for patients whose disease was in remission (29). When considering medication costs, previous studies found that (if biologics were used) medication costs contributed to almost half of the health care costs of JIA patients (7,30,31), which is in line with the 55.6% found in the current study. However, the rise in use of biologics over time (i.e., 31% in the current study versus 6% in the study by Minden et al [30]) as well as the fluctuations in prices of biologics makes these numbers hard to compare.

In the current study, the maximum duration of follow-up was 8 years (i.e., from 2011 to 2019) but differed between patients. Therefore, this study generally did not capture the entire patient's disease course (i.e., from JIA diagnosis until reaching the age of 18). As a consequence, total costs of JIA treatment on an individual patient level could not be calculated. Costs were therefore expressed as mean costs per patient per year or per month of follow-up. This approach allowed for inclusion of most of the available data in the analysis. In addition, it allowed inclusion of patients who were recently diagnosed with JIA. Despite the relatively short duration of follow-up, the inclusion of these recently diagnosed patients was nevertheless highly desirable and necessary as treatment strategies in JIA and costs for medication are continuously evolving.

The current analysis used fixed cost prices, indicating that price fluctuations over time (e.g., for biologics) were not incorporated. Such an approach was taken because the moment of JIA diagnosis was used as starting point of the analysis. To illustrate this approach, patients were analyzed as being in their first year of follow-up (i.e., the first year following JIA diagnosis), regardless of whether this diagnosis was established in, for example, 2011 or 2018.

Although differences in annual hospital costs between individual patients are (inevitably) caused by differences in disease severity, they are also attributable to the part of the treatment course captured for each patient between April 1, 2011 and March 31, 2019. In other words, for some patients, data may have been available for the first 2 years after JIA diagnosis, whereas for other patients, only a period of inactive disease was captured. In addition, fluctuations in treatment intensity on an individual patient level further increase the variability in annual hospital costs. Consequently, including uncertainty boundaries with regard to patient-level outcomes would have led to extremely large confidence intervals. Furthermore, as the annual hospital costs were strongly right skewed (which is common with cost data), reporting medians would disregard this skewness and thus underestimate the effect of rare cost-intensive cases. Therefore, histograms are preferred in health economic decision-making to visualize patient-level variations in costs (32).

Treatment options for JIA continue to develop, indicating that the costs and health impact of JIA have changed significantly over the last years, which is especially attributable to the rise in the availability and use of biologics (9). In order to increase the likelihood that patients were comparable at each year of follow-up, the duration of follow-up was limited to 10 years after JIA diagnosis. Also, as the maximum duration of follow-up a patient could reach before his/her 18th birthday depended on the age at JIA diagnosis (e.g., a patient diagnosed at the age of 12 could reach a maximum follow-up of 6 years), the number of patients decreased over time (i.e., from 426 in year 1 to 133 by the end of year 10). A duration of follow-up longer than 10 years would have decreased the reliability of the mean annual costs, as fluctuations in costs over time would then primarily be attributable to the large variation in costs between individual patients.

One strength of the present study is that it is the first patient-level analysis of hospital costs in a large database for different JIA subtypes and over the course of JIA treatment. More specifically, the number of studies that have investigated health care–related resource use and costs in JIA is limited (14), and many of these studies did not distinguish costs between subtypes of JIA, focused on a specific subtype, or did not investigate changes in health care–related costs over the course of JIA treatment.

Another strength of our study is that it is expected to provide an accurate representation of the average costs of all patients with a diagnosis of JIA, regardless of disease or medication state. To illustrate this, this study also included patients that have not received treatment or visited their pediatric rheumatologists for a substantial amount of time. Disease in these patients was most likely in remission, which was associated with considerably lower treatment costs (29).

This study also has some limitations. One limitation is that data on physician visits within the hospital (other than pediatric rheumatologist visits) were only available up to December 12, 2018, indicating that visits during the last 3.5 months of the 96 months database were missing. This is expected to represent an underestimate of costs of other physician visits with €12/patient (i.e., €336/patient instead of €324/patient). However, as the frequency as well as the type of physician visits differed considerably between JIA subtype, between patients, and over the course of follow-up, extrapolating these costs was considered to incur more uncertainty compared to the current underestimation.

Another limitation of the current study is that costs occurring outside the hospital (e.g., including costs of visits to a regional physiotherapist or ophthalmologist) as well as out-of-pocket costs and productivity losses for patients, parents, and caregivers are expected to substantially impact the societal costs of JIA (8,9,14,33), but this was beyond the scope of this retrospective analysis of hospital costs. Nevertheless, it is critical to evaluate these costs. Therefore, the impact of JIA on the overall costs to society is currently investigated in a large multicenter, international prospective collaborative study into management strategies for JIA, conducted in Canada and The Netherlands, named UCAN CAN-DU (https://www.ucancandu.com/). The findings of the

current study will be used to optimize the methodology of UCAN CAN-DU.

Our study yields implications for practice and generalizability. We found major differences in hospital-related resource use between patients, which emphasizes that JIA-related treatment costs also need to be analyzed at the individual level. More specifically, future studies should investigate the impact of early, intensive treatment in patients with severe JIA on resource use and costs for the short- and long-term and offset these against health outcomes like the Juvenile Disease Activity Score or the EuroQoL 5-Dimension 5-Level (EQ-5D-5L) measure.

The present study was conducted as a single-center study that is known to be the largest JIA treatment center in The Netherlands. Because patients participating in pharmaceutical-sponsored studies were excluded, this ensures results are highly representative of current practice. The extent to which the results are generalizable to other countries will however depend on differences in costs as well as access to hospital resources. An example of such differences is that in The Netherlands, anakinra is recommended as first-line treatment in systemic JIA patients (34), a medication that is not reimbursed in all countries. Therefore, this generalizability will largely depend on similarities and differences between treatment protocols.

In conclusion, hospital-care associated costs of JIA vary substantially between individual patients and between JIA subtypes. Mean annual costs were the highest for systemic JIA patients and were primarily attributable to medication costs. Costs of other hospital-based services were comparable regardless of JIA subtype. Except for medication use, costs of other hospital-based services decrease after JIA diagnosis. Future studies are required to capture the full impact of JIA to society, including costs associated with JIA-related care as well as productivity losses and out-of-pocket costs.

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. IJzerman 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. Kip, Roock, van den Berg, Currie, Marshall, Grazziotin, Twilt, Swart, IJzerman.

Acquisition of data. Kip, Roock, van den Berg, Vastert, Wulffraat, Swart.

Analysis and interpretation of data. Kip, Roock, van den Berg, Currie, Marshall, Grazziotin, Twilt, Yeung, Benseler, Vastert, Wulffraat, Swart, IJzerman.

REFERENCES

1. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. Lancet 2011;377:2138–49.
2. Shiff NJ, Oen K, Kroeker K, Lix LM. Trends in population-based incidence and prevalence of juvenile idiopathic arthritis in Manitoba, Canada. Arthritis Care Res (Hoboken) 2019;71:413–8.
3. 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.
4. Albers HM, Wessels JA, van der Straaten RJ, Brinkman DM, Suijlekom-Smit LW, Kamphuis SS, et al. Time to treatment as an important factor for the response to methotrexate in juvenile idiopathic arthritis. Arthritis Rheum 2009;61:46–51.
5. Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007;369: 767–78.
6. Luca NJ, Burnett HF, Ungar WJ, Moretti ME, Beukelman T, Feldman BM, et al. Cost-effectiveness analysis of first-line treatment with biologic agents in polyarticular juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2016;68:1803–11.
7. Bernatsky S, Duffy C, Malleson P, Feldman DE, St Pierre Y, Clarke AE. Economic impact of juvenile idiopathic arthritis. Arthritis Rheum 2007; 57:44–8.
8. Angelis A, Kanavos P, Lopez-Bastida J, Linertova R, Serrano-Aguilar P, Network BURQOL-RD Research Network. Socioeconomic costs and health-related quality of life in juvenile idiopathic arthritis: a cost-of-illness study in the United Kingdom. BMC Musculoskelet Disord 2016;17:321.
9. Kuhlmann A, Schmidt T, Treskova M, Lopez-Bastida J, Linertova R, Oliva-Moreno J, et al. Social/economic costs and health-related quality of life in patients with juvenile idiopathic arthritis in Europe. Eur J Health Econ 2016; Suppl 1:79–87.
10. Schlichtiger J, Haas JP, Barth S, Bisdorff B, Hager L, Michels H, et al. Education and employment in patients with juvenile idiopathic arthritis: a standardized comparison to the German general population. Pediatr Rheumatol Online J 2017;15:45.
11. Barth S, Haas JP, Schlichtiger J, Molz J, Bisdorff B, Michels H, et al. Long-term health-related quality of life in German patients with juvenile idiopathic arthritis in comparison to German general population. PLoS One 2016;11:e0153267.
12. 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 follow-up study. Arthritis Care Res (Hoboken) 2018;70:741–9.
13. Muller-Godeffroy E, Lehmann H, Kuster RM, Thyen U. Quality of life and psychosocial adaptation in children and adolescents with juvenile idiopathic arthritis and reactive arthritis. Z Rheumatol 2005;64: 177–87.
14. Kip MM, Currie G, Marshall DA, Grazziotin Lago L, Twilt M, Vastert SJ, et al. Seeking the state of the art in standardized measurement of health care resource use and costs in juvenile idiopathic arthritis: a scoping review. Pediatr Rheumatol Online J 2019;17:20.
15. Lee JJ, Schneider R. Systemic juvenile idiopathic arthritis. Pediatr Clin North Am 2018;65:691–709.
16. Davies R, Gaynor D, Hyrich KL, Pain CE. Efficacy of biologic therapy across individual juvenile idiopathic arthritis subtypes: a systematic review. Semin Arthritis Rheum 2017;46:584–93.
17. Vastert SJ, Nigrovic PA. Toward personalized treatment for systemic juvenile idiopathic arthritis [editorial]. Arthritis Rheumatol 2018;70: 1172–4.
18. Funk RS, Becker ML. Disease modifying anti-rheumatic drugs in juvenile idiopathic arthritis: striving for individualized therapy. Expert Rev Precis Med Drug Dev 2016;1:53–68.
19. Swart JF, van Dijkhuizen EH, Wulffraat NM, de Roock S. Clinical Juvenile Arthritis Disease Activity Score proves to be a useful tool in treat-to-target therapy in juvenile idiopathic arthritis. Ann Rheum Dis 2018; 77:336–42.

20. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310:2191–4.
21. Central Bureau of Statistics. Jaarmutatie consumentenprijsindex; vanaf 1963. 2019. URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70936NED/table?fromstatweb
22. R Core Team. R: A languange and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2019.
23. Wickham H, François R, Henry L, Müller K. A grammar of data manipulation. R package version 0.8.3.; 2019.
24. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
25. Grolemund G, Wickham H. Dates and times made easy with lubridate. J Stat Softw 2011;40:1–25.
26. Lemon J. Plotrix: a package in the red light district of R. R-News 2006; 6:8–12.
27. Kip MM, de Roock S, Currie G, Marshall DA, Grazziotion LR, Twilt M, et al. Costs of medication use among patients with juvenile idiopathic arthritis in the Dutch healthcare system. Expert Rev Pharmacoecon Outcomes Res 2021;21:975–84.
28. Allaire SH, DeNardo BS, Szer IS, Meenan RF, Schaller JG. The economic impacts of juvenile rheumatoid arthritis. J Rheumatol 1992;19: 952–5.
29. Minden K, Niewerth M, Listing J, Biedermann T, Schontube M, Zink A. Burden and cost of illness in patients with juvenile idiopathic arthritis. Ann Rheum Dis 2004;63:836–42.
30. Minden K, Niewerth M, Listing J, Mobius D, Thon A, Ganser G, et al. The economic burden of juvenile idiopathic arthritis-results from the German paediatric rheumatologic database. Clin Exp Rheumatol 2009;27:863–9.
31. Haapasaari J, Kautiainen HJ, Isomaki HA, Hakala M. Etanercept does not essentially increase the total costs of the treatment of refractory juvenile idiopathic arthritis. J Rheumatol 2004;31:2286–9.
32. Mani K, Lundkvist J, Holmberg L, Wanhainen A. Challenges in analysis and interpretation of cost data in vascular surgery. J Vasc Surg 2010;51:148–54.
33. Rasu RS, Cline SK, Shaw JW, Hayes O, Agbor Bawa W, Cifaldi MA. Impact of JIA on parents' work absences. Rheumatology (Oxford) 2015;54:1177–85.
34. Vastert SJ, de Jager W, Noordman BJ, Holzinger D, Kuis W, Prakken BJ, et al. Effectiveness of first-line treatment with recombinant interleukin-1 receptor antagonist in steroid-naive patients with new-onset systemic juvenile idiopathic arthritis: results of a prospective cohort study. Arthritis Rheumatol 2014; 66:1034–43.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1593–1601
DOI 10.1002/acr.24618

American College of Rheumatology
Empowering Rheumatology Professionals

Patient Acceptance of Nurse Practitioners and Physician Assistants in Rheumatology Care

Karen Frazier,[1] Kathryn A. Paez,[1] Emily Creek,[2] Arlene Vinci,[2] Andrew Amolegbe,[1] and Arifah Hasanbasri[1]

Objective. To assess whether patients with autoimmune disease would accept advanced practice providers (APPs) as an option to fill the growing shortage of rheumatologists.

Methods. We administered a cross-sectional survey to 500 patients or parents of children who reported having been diagnosed with qualifying autoimmune conditions and who had seen their primary rheumatology providers in the past 6 months. Respondents self-reported whether their primary providers were rheumatologists or APPs. Our analysis compared the attitude and experience of the patients whose primary rheumatology providers were APPs with those of patients whose primary providers were rheumatologists.

Results. Of respondents, 36.8% reported having APPs as primary rheumatology providers. Patients of APPs were significantly more likely to arrive at their provider's office in 15 minutes or less ($P < 0.01$) and to be able to schedule routine and urgent appointments sooner ($P = 0.02$ and 0.05, respectively). There were no significant differences in overall patient experience of care between provider types. Most patients rated their providers highly, but those who saw rheumatologists rated their providers significantly higher ($P < 0.01$). Patients of APPs were significantly more likely than patients of rheumatologists to prefer to see APPs over rheumatologists ($P < 0.01$) and to recommend APPs ($P < 0.01$).

Conclusion. APPs may improve access to care and, regardless of provider type, patients rated their overall experience of care similarly. Overall, patient attitudes toward APPs were positive regardless of provider type, although APP patients held more positive overall attitudes toward APPs than did rheumatologist patients.

INTRODUCTION

There is a gap between the need for rheumatology care and the supply of specialists, and this gap is projected to grow exponentially in upcoming years, with the shortage projected to increase from 1,769 full-time specialists in 2020 to nearly 5,000 in 2030 (1). Meeting the projected shortage will require increasing the 2015 American College of Rheumatology (ACR) estimates of 6,050 full-time rheumatology providers by 83%. The growing need for rheumatology care is fueled by the aging patient population and rheumatology workforce and the growing prevalence of autoimmune conditions. The number of people in the US found to have antinuclear antibodies (the most common biomarker of autoimmunity) grew by more than 19 million in 25 years, an increase in prevalence from 11% to 16% (2). Meanwhile, the supply of rheumatologists, projected in 2015 to be less than 5,600 full-time providers (1), is declining as more providers retire and are replaced by new medical graduates who choose to work fewer hours per week and see fewer patients (1). Current access to rheumatology care can be challenging, with the existing workforce unevenly distributed across the country (1). In 2015, 5 states had <15 practicing adult rheumatologists, while half of states had ≤3 practicing pediatric providers (1). Some rural residents travel ≥200 miles for an appointment with a rheumatology specialist (1).

To improve access to care in the face of this shortfall, the 2015 ACR Workforce Study recommended the ACR and the Association of Rheumatology Professionals adopt strategies to increase the inclusion of nurse practitioners (NPs) and physician assistants (PAs) in rheumatology care and provide appropriate training to prepare them. Advanced practice providers (APPs), which comprise NPs and PAs (3), specializing in rheumatology work independently to evaluate and treat patients, interpret and deliver test results, perform procedures, and prescribe medications (4,5). In addition, rheumatology practices with and without APPs see patients with similar characteristics, including disease activity, and have comparable prescribing patterns. However, those rheumatology practices that employ APPs report more patient visits than those without APPs (6).

Supported by the Arthritis Foundation.
[1]Karen Frazier, PhD, Kathryn A. Paez, PhD, MBA, RN, Andrew Amolegbe, MPH, Arifah Hasanbasri, BS: American Institutes for Research, Arlington, Virginia; [2]Emily Creek, MBA, Arlene Vinci, BBA: Arthritis Foundation, Atlanta, Georgia.
No potential conflicts of interest relevant to this article were reported.

Address correspondence to Karen Frazier, PhD, American Institutes for Research, 100 Europa Drive, Suite 315, Chapel Hill, NC 27517. Email: Kfrazier@air.org.
Submitted for publication August 24, 2020; accepted in revised form April 6, 2021.

SIGNIFICANCE & INNOVATIONS

- Regardless of whether their primary providers were rheumatologists or advanced practice providers (APPs), respondents were favorable about their overall experience of care and rated the 2 care experiences similarly.
- Respondent attitudes toward APPs were positive regardless of provider type, although patients of APPs held more positive overall attitudes toward APPs than did patients of rheumatologists.
- For some, particularly for parents of children with autoimmune conditions, travel to a rheumatology provider may be burdensome and incorporating APPs into rheumatology care may improve access to care.

Most research on the quality of care provided by APPs has been conducted in primary care or emergency departments rather than specialty care settings. Studies have found, with 1 exception (7), that care provided by APPs or physician-led teams that included APPs was comparable to, if not better than, care provide by physicians alone (8–18). In the exceptional case, researchers found that, among primary care clinicians treating diabetic patients with multiple chronic conditions, APPs were less likely than physicians to make a change in a patient's treatment when that patient showed elevated blood pressure during a visit (7). The few studies within medical subspecialties of physician-only teams or practices versus those including APPs found that teams and practices with APPs provided care that was better than or equivalent to those without APPs. These conclusions were based on indicators such as readmission rates, medication adherence, or quality metric ratings from patients with cirrhosis or cardiovascular disease (19–21).

Despite studies showing equivalence in quality of care provided by APPs and physicians, only about one-quarter of rheumatologists employed APPs in their practices in 2015, with APPs making up just 7.5% of the adult and 8% of the pediatric rheumatology workforce (1). In addition, the research is silent on the extent to which rheumatology patients accept APPs for their care. To fill in the gap, the present study sought to assess whether patients with autoimmune disease would accept APPs as their rheumatology providers, thereby helping fill the growing shortage of rheumatologists. To answer this question, we surveyed patients and parents of children with autoimmune disease and compared the attitudes and experience of those patients whose primary rheumatology providers were APPs with those of patients whose primary providers were rheumatologists.

MATERIALS AND METHODS

Survey sample and recruitment. We surveyed 500 patients who reported having been diagnosed with a qualifying autoimmune condition or parents of children who met the same criteria. Inclusion criteria required that all respondents report having seen their primary rheumatology providers in the past 6 months. Adult respondents qualified if they reported having been diagnosed with at least 1 of the following diseases: rheumatoid arthritis (RA), psoriatic arthritis, ankylosing spondylitis, juvenile arthritis persistent to adulthood, or lupus. Parents of children who had been diagnosed with at least 1 of the following illnesses also qualified: juvenile idiopathic arthritis, juvenile myositis, juvenile lupus, juvenile scleroderma, vasculitis, or fibromyalgia. The American Institutes for Research's Institutional Review Board reviewed this study and determined it to be exempt.

A survey research firm recruited respondents using Dynata's US-based nonprobability web panel of 31 million people (22) because of the low prevalence of those with autoimmune disease and the limited number of APPs employed at rheumatology practices across the US (1). The firm invited participants from the opt-in panel through email and ended data collection when 500 respondents completed the survey. To ensure the successful recruitment of respondents whose primary providers were APPs, the firm oversampled in metropolitan areas known by the Arthritis Foundation to have APPs employed in rheumatology practices and in more than 250 areas with large- to mid-size NP and PA credentialing programs. The firm administered the survey online in English and Spanish from March 24 through April 16, 2020.

Survey data and measures. The research team developed the survey to gather data on respondent demographic characteristics, disease activity and severity, provider access and visit characteristics, respondent attitudes toward APPs, and experience of care. Respondents were asked to self-report whether their primary rheumatology provider was a rheumatologist or an APP.

Global disease activity. The survey used 2 patient global assessments to assess disease activity, including 1 for adults and 1 for children. These assessments asked respondents to rate the activity of their autoimmune condition in the last month on a 10-point scale and rate their health in the last month, considering only their autoimmune condition (23).

Physical and mental health. To measure adult health, the survey included the Short Form 12 health survey version 2 (SF-12v2), which is a shortened version of the SF-36 health survey (a health-related quality-of-life measure) (24–27). The SF-12v2 assesses 8 domains, including general physical and mental health, physical and social functioning, physical and social limitations, bodily pain, and vitality, to establish a physical component summary (PCS) and mental component summary (MCS) score. The nationally reported mean for each component is 50 (31). To assess physical functioning in children, the questionnaire included the Juvenile Arthritis Functional Assessment Report for Parents (JAFAR-P), a functional assessment of juvenile idiopathic arthritis patients completed by parent proxy (28). The JAFAR-P includes 23 questions

about physical functioning during the past week with 3 response options, including "all the time," "sometimes," and "almost never." Total scores for the JAFAR-P range from 0 to 46, with lower scores representing higher functioning (28). We also included 4 questions from the Patient-Reported Outcomes Measurement Information System parent proxy fatigue SF-10, rated on a Likert scale from "never" to "often."

Provider visit activities and access to care. We created de novo questions about activities the respondents' providers engaged in during an appointment, such as a physical exam, prescribing medicine, or recommending additional treatments (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24618). We also created de novo questions asking participants how many miles and minutes they had to travel for an appointment, how long they spent at the provider's office for an appointment, and how often they accepted appointments with APPs to get earlier appointments in the last 6 months (for respondents whose primary providers were rheumatologists) (Supplementary Table 1).

Attitudes toward APPs. To assess respondent attitudes toward APPs, the survey included 6 statements (adapted from the Nurse Practitioner Satisfaction Survey [NPSS]) about their perspectives on APPs (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24618). Respondents were asked to rate their agreement with each statement from "strongly disagree" to "strongly agree" (29).

Experience of care. Finally, to assess patients' experience of care, the survey incorporated core questions tailored from the Adult and Child Consumer Assessment of Healthcare Providers and Systems (CAHPS) Clinician & Group (CG) surveys, including select supplemental items from the CG CAHPS and CAHPS Cancer Care surveys. CAHPS assesses patient experiences with their

Table 1. Patient characteristics*

	All	Rheumatologist patients	APP patients	*P*
All respondents	100.0 (500)	63.2 (316)	36.8 (184)	
Adult	89.0 (445)	85.8 (271)	94.6 (174)	<0.01
Child	11.0 (55)	14.2 (45)	5.4 (10)	<0.01
Age, mean ± SD years				
Adult	57.6 ± 15.8	58.7 ± 15.4	55.8 ± 16.2	0.06
Children	9.6 ± 3.6	9.4 ± 3.6	10.7 ± 3.6	0.31
Women	57.0 (285)	58.5 (185)	54.4 (100)	0.36
Race				
White	86.4 (432)	88.6 (280)	82.6 (152)	0.06
Black	10.6 (53)	9.2 (29)	13.0 (24)	0.18
Asian, Native Hawaiian, other Pacific Islander	3.6 (18)	3.2 (10)	4.4 (8)	0.49
Other	2.8 (14)	2.9 (9)	2.7 (5)	0.93
Ethnicity, Latino	11.4 (57)	12.0 (38)	10.3 (19)	0.56
Education				0.03
Some high school or less	4.7 (21)	3.0 (8)	7.5 (13)	
High school graduate or GED	29.2 (130)	27.7 (75)	31.6 (55)	
Some college or 2-year degree	31.2 (139)	31.7 (86)	30.5 (53)	
4-year college graduate or more	34.8 (155)	37.6 (102)	30.5 (53)	
Geography				0.04
Large urban	15.1 (75)	18.5 (58)	9.3 (17)	
Mid-size urban	20.9 (104)	19.1 (60)	24 (44)	
Suburban	48.7 (242)	49 (154)	48.1 (88)	
Small town	7.7 (38)	7 (22)	8.7 (16)	
Rural	7.7 (38)	6.4 (20)	9.8 (18)	
Years since diagnosis				
Adults, >2 years ago	76.6 (340)	79.0 (214)	72.8 (126)	NA
Children, <2 years ago	77.8 (42)	79.5 (35)	70.0 (7)	NA
Disease activity in last month, mean ± SD	6.3 ± 2.7	6.4 ± 2.8	6.1 ± 2.5	
More active (rating 7–10)	52.6 (263)	53.8 (170)	50.5 (93)	0.48
Less active (rating 0–6)	47.4 (237)	46.2 (146)	49.5 (91)	
SF-12 PCS score, mean ± SD†	39.2 ± 9.5	39.4 ± 9.8	39.1 ± 9.1	0.76
SF-12 MCS score, mean ± SD†	46.4 ± 10.6	47.5 ± 10.4	44.7 ± 10.7	0.01
JAFAR-P score, mean ± SD‡	10.6 ± 9.8	9.8 ± 9.6	13.9 ± 10.4	0.24

* Values are the percent (number) unless indicated otherwise. APP = advanced practice providers; GED = general educational development; JAFAR-P = Juvenile Arthritis Functional Assessment Report for Parents; MCS = mental component summary score; NA = not available; PCS = physical component summary score; SF-12 = Short Form 12 health survey.
† Average = 50 ± 10.
‡ Total scores range from 0 to 46, with lower scores representing higher functioning.

primary rheumatology providers in the last 6 months. These experiences concern accessing care, provider communication and care coordination, provider education and promotion of healthy behaviors, support for managing effects of arthritis and treatment, and participatory decision-making. The response scale is "never," "sometimes," "usually," "always" or "yes, definitely," "yes, somewhat," and "no." Respondents provide an overall rating of their provider on a scale from 0 to 10, with a higher score indicating a better rating.

Statistical analysis. The survey firm weighted the data to ensure the demographic profile of the sample matched the profile of the target population as estimated from the 2018 National Center for Health Statistics' National Health Interview Survey (NHIS) data (30). The first stage of weighting was the application of a base weight to account for different selection probabilities and response rates across sample strata. In the second stage of weighting, the firm used raking to match sample demographic characteristics to population parameters. A survey was considered complete if the respondent reached the final demographic question. On average, the survey took 13 minutes to complete and, to ensure data quality, cases were deleted if the length of the interview was <4.3 minutes. In addition, the survey firm deleted cases if the respondent selected the same answer for every question in ≥2 series of questions that presented the same set of response options.

We used Stata, version 15.1 for all analyses and calculated descriptive statistics for overall responses and by provider type. Nominal and ordinal data were described as percentage (number) and continuous data as mean ± SD. To identify differences between respondents primarily seeing rheumatologists versus APPs, we conducted 2-sample Wilcoxon rank-sum tests for ordinal data, 2-sample *t*-tests for continuous data, and chi-square tests for nominal data. In addition, we conducted analysis of covariance (ANCOVA) tests to further investigate differences in CAHPS composite scores and attitudes toward APPs while controlling for age, education, and general health ratings. *P* values less than 0.05 were considered significant.

RESULTS

Of the 500 survey respondents, 36.8% (184) reported having APPs as primary rheumatology providers (Table 1). Of those, 94.6% (174) were adults who reported having been diagnosed with autoimmune conditions, and 5.4% (10) were parents who reported the same for their child. Of the participants who reported seeing rheumatologists, 85.8% (271) were adults and 14.2% (45) were parents of children with autoimmune conditions.

More than half of the respondents were female, and 86% were White; there was no significant difference in terms of sex or race between respondents who saw rheumatologists and those who saw APPs. Adult respondents with at least a 4-year college degree (34.8% of adults) were more likely to see rheumatologists than APPs ($P = 0.03$) compared to those respondents with less education. Relatively few respondents lived in small towns (7.7%) or rural areas (7.7%), with most respondents concentrated in suburban areas or mid-sized to large urban areas. Respondents in large urban areas tended to see rheumatologists rather than APPs ($P = 0.04$).

The most common diagnosis for adults was RA (83%), while juvenile idiopathic arthritis was most common for children (65.5%), as shown in Table 2. Among adults with RA, a greater proportion of patients saw rheumatologists than APPs (87.5% versus 77.6%; $P = 0.01$). A notable proportion of children were diagnosed with juvenile lupus (29.1%), fibromyalgia (27.3%), and juvenile scleroderma (25.5%). Most adults were diagnosed with autoimmune conditions >2 years before, while most children had been diagnosed <2 years before. Reported disease activity varied, with 52% of respondents reporting that their (or their child's) disease had been active in the last month (Table 1). In

Table 2. Condition types of patients*

	All respondents	Rheumatologist patients	APP patients	*P*
Adult conditions				
Rheumatoid arthritis	83.6 (372)	87.5 (237)	77.6 (135)	0.01
Psoriatic arthritis	13.9 (62)	12.9 (35)	15.5 (27)	0.44
Ankylosing spondylitis	11.5 (51)	10.3 (28)	13.2 (23)	0.35
Juvenile arthritis persistent to adulthood	5.2 (23)	3.3 (9)	8.1 (14)	0.03
Lupus	9.4 (42)	10.7 (29)	7.5 (13)	0.26
Childhood conditions				
Juvenile idiopathic arthritis	65.5 (36)	71.1 (32)	40.0 (4)	0.06
Juvenile myositis	16.4 (9)	15.6 (7)	20.0 (2)	0.73
Juvenile lupus	29.1 (16)	26.7 (12)	40.0 (4)	0.40
Juvenile scleroderma	25.5 (14)	28.9 (13)	10.0 (1)	0.22
Vasculitis	18.2 (10)	20.0 (9)	10.0 (1)	0.46
Fibromyalgia	27.3 (15)	20.0 (9)	60.0 (6)	0.01

* Values are the percent (number) unless indicated otherwise. APP = advanced practice providers.

Table 3. Patient access to care*

	All respondents	Rheumatologist patients	APP patients	*P*
Minutes traveling to provider's office				<0.01
≤15	31.3 (155)	26.0 (82)	40.6 (73)	
16–30	39.3 (195)	39.9 (126)	38.3 (69)	
31–45	17.7 (88)	18.7 (59)	16.1 (29)	
45 to 1 hour	6.7 (33)	8.5 (27)	3.3 (6)	
>1 hour	3.6 (18)	4.8 (15)	1.7 (3)	
>2 hours	1.4 (7)	2.2 (7)	0.0 (0)	
Routine appointment scheduling time				0.02
≤1 day	20.4 (89)	18.1 (49)	24.1 (40)	
2 days to ≤1 week	47.4 (207)	46.5 (126)	48.8 (81)	
>1 week to ≤1 month	23.3 (102)	23.6 (64)	22.9 (38)	
>1 month	8.9 (39)	11.8 (32)	4.2 (7)	
Urgent appointment scheduling time, days				0.03
Same day	19.3 (35)	15.3 (15)	24.1 (20)	
1	19.9 (36)	19.4 (19)	20.5 (17)	
2–3	37.6 (68)	35.7 (35)	39.8 (33)	
4–7	14.4 (26)	18.4 (18)	9.6 (8)	
>7	8.8 (16)	11.2 (11)	6 (5)	
Provider spent enough time with patient				0.21
Never	2.0 (10)	1.9 (6)	2.2 (4)	
Sometimes	9.7 (48)	8.6 (27)	11.7 (21)	
Usually	25.8 (127)	25.2 (79)	26.8 (48)	
Always	62.5 (308)	64.3 (202)	59.2 (106)	

* Values are the percent (number) unless indicated otherwise. APP = advanced practice providers.

reporting adult health through the SF-12 health survey, there were no significant differences between the physical health scores of adults seeing each type of provider, with the mean ± SD PCS score (39.2 ± 9.5) for all adults. Respondents whose primary providers were rheumatologists scored 2.8 points higher on mental health than those seeing APPs (mean ± SD MCS score 47.5 ± 10.4 versus 44.7 ± 10.7; $P = 0.01$). Parents of children reported high physical functioning, with a mean ± SD score on the JAFAR-P of 10.6 ± 9.8 and no significant differences between physical function by provider.

Access to care. Most respondents traveled ≤30 minutes to their providers' offices (70.6%), with 31.3% traveling ≤15 minutes (Table 3). Compared to urban residents, respondents living in any of the 4 less-populated geographic areas were 2–3 times more likely to see an APP than a rheumatologist (P <0.01–0.03) (Table 4). Respondents seeing APPs were significantly more likely than those seeing rheumatologists to arrive at their provider's office in ≤15 minutes ($P < 0.01$). Notably, 72.6% of adults, but only 53.7% of children, were able to travel to their providers within 30 minutes. When scheduling routine appointments, 67.8% of respondents reported being able to get an appointment within 1 week, with those seeing APPs significantly more likely to be able to schedule an appointment sooner than those seeing rheumatologists ($P = 0.02$). Similarly, the majority of respondents were able to schedule an urgent appointment within ≤3 days (76.8%), with those seeing APPs significantly more likely to get an appointment sooner than those seeing rheumatologists ($P = 0.05$). Most respondents reported that their providers "usually" or "always" spent enough time with them (88.3%), with no significant differences between provider types.

Patient experience. There were no significant differences between patients of rheumatologists and APPs in patient experience of care or the types of activities providers performed. When reporting their experience with their providers in the last 6 months (using CAHPS composite scoring), respondents rated them highly on measures of care coordination provider communication, and provider communication with the child (mean ± SD 80.1 ± 22.5, 84.4 ± 19, and 78.3 ± 22.2, respectively) (Table 5). Respondents rated their providers slightly lower on access to care (mean ± SD 71.4 ± 25.1).

Overall, most respondents rated their providers highly; however, those respondents who saw rheumatologists rated their

Table 4. Likelihood of patient seeing an APP versus a rheumatologist in various population densities*

	OR (95% CI)	*P*
Urban†	–	
Mid-size urban	2.5 (1.29–4.87)	0.01
Suburban	1.9 (1.07–3.56)	0.03
Small town	2.5 (1.07–5.75)	0.03
Rural	3.1 (1.33–7.08)	0.01

* 95% CI = 95% confidence interval; APP = advanced practice providers; OR = odds ratio.
† Urban population density was the comparison group.

Table 5. Patient experience*

	All respondents	Rheumatologist patients	APP patients	*P*
CAHPS composite scoring				
Care coordination	80.1 ± 22.5	81.5 ± 21.8	77.7 ± 23.6	0.15
Provider communication	84.4 ± 19.0	85.6 ± 18.2	82.3 ± 20.1	0.20
Provider communication with child	78.3 ± 22.2	79.6 ± 21.5	73.3 ± 25.1	0.86
Access to care	71.4 ± 25.1	71.8 ± 26.2	71.0 ± 23.6	0.79
Provider rating	85.4 ± 16.9	88.2 ± 14.8	80.5 ± 18.9	<0.01

* Values are the mean ± SD unless indicated otherwise. Data adjusted for age, education, and overall health status. APP = advanced practice providers; CAHPS = Consumer Assessment of Healthcare Providers and Systems.

providers significantly higher (mean ± SD 88.2 ± 14.8) than did respondents who saw APPs (mean ± SD 80.5 ± 18.9; $P < 0.01$). In addition, most respondents seeing rheumatologists (93%) and APPs (88.4%) rated their providers as "usually" or "always" showing respect for what they had to say, although the difference was statistically significant ($P = 0.05$). Similarly, most respondents seeing rheumatologists (87.3%) and those seeing APPs (82.3%) reported that their providers "usually" or "always" knew their medical history, but those seeing rheumatologists were more likely to report this than those seeing APPs ($P = 0.02$).

Respondent attitudes toward APPs. Half of respondents who primarily saw rheumatologists accepted earlier appointments with APPs in the past 6 months. For the 6 questions assessing perspectives on APPs from the NPSS, the mean ± SD for respondents seeing APPs (mean ± SD 71.8 ± 19.8) was higher than for those seeing rheumatologists (mean ± SD 65.4 ± 20.3), even when adjusted for age, education, and health status ($P < 0.01$) (see Table 6, and Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24618). This difference in scores was driven by the fact that there were significant differences by provider type for 2 questions, including preference to see an APP over a rheumatologist and likelihood of recommending an APP. More than 64% of respondents primarily seeing APPs and 26.1% of respondents primarily seeing rheumatologists agreed or strongly agreed that they preferred to see an APP over a rheumatologist ($P < 0.01$). More than 84% of respondents who saw APPs agreed or strongly agreed that they would recommend an APP, compared with 58.9% of respondents seeing rheumatologists ($P < 0.01$). Conversely, there were no significant differences by provider type as to whether the respondents thought that APPs were skilled health care providers, were knowledgeable about health problems, or knew when to refer patients to or consult with a physician. The Chronbach's alpha value for the respondent attitudes survey was 0.84A.

DISCUSSION

This study aimed to understand whether patients with autoimmune conditions would accept APPs for their rheumatology care by comparing the attitudes and experiences of people whose primary rheumatology providers were APPs with those who primarily saw rheumatologists. While some research exists on the degree to which patients will accept care by APPs, studies have focused mainly on primary care and emergency medicine, with little attention paid to preferences or satisfaction across provider type in specialty areas like rheumatology. We found that most people with autoimmune conditions were positive about their care experience regardless of whether they primarily saw APPs or rheumatologists, and that >40% of people, including people seeing rheumatologists, would prefer to see APPs. These

Table 6. Patient attitudes toward APPs*

	All respondents	Rheumatologist patients	APP patients	*P*
Likely to recommend an APP				<0.01
Strongly agree	23.3 (116)	22.0 (69)	25.5 (47)	
Agree	45.2 (225)	36.9 (116)	59.2 (109)	
Disagree	20.7 (103)	26.4 (83)	10.9 (20)	
Strongly disagree	10.8 (54)	14.7 (46)	4.4 (8)	
Would prefer to see an APP				<0.01
Strongly agree	13.8 (69)	12.4 (39)	16.3 (30)	
Agree	26.5 (132)	13.7 (43)	48.4 (89)	
Disagree	37.1 (185)	43.5 (137)	26.1 (48)	
Strongly disagree	22.7 (113)	30.5 (96)	9.2 (17)	

* Values are the percent (number) unless indicated otherwise. Data adjusted for age, education, and overall health status, and questions selected and adapted from Nurse Practitioner Satisfaction Survey (ref. 29). No significant difference by provider type for whether respondent thought advanced practice providers (APPs) were skilled health care providers, were knowledgeable about health problems, or knew when to refer patients to consult with a physician.

findings suggest that, if more APPs were brought into rheumatology practices to address the projected shortage of 5,000 rheumatology providers in the next decade (1), a substantial number of people would be receptive to APPs as their primary rheumatology providers.

We found several significant differences in respondent characteristics across primary provider types although these differences did not indicate racial or ethnic disparities. People seeing APPs were less likely to have a 4-year degree. They were also less likely to live in urban areas than people who saw rheumatologists, perhaps because practices in more densely populated areas had less need to supplement their rheumatology workforce with APPs or because APPs were more willing to live in small town and rural settings. Respondents had similar autoimmune disease activity and severity regardless of primary provider type. Overall, adults had lower mental health scores than the general population, although adults seeing APPs compared to those seeing rheumatologists scored lower on the mental health component of the SF-12. There is no obvious reason for this difference.

In regard to access to care, this research reinforces that, for some, travel to a rheumatology provider may be burdensome, particularly for parents of children with autoimmune conditions. Respondents whose primary providers were APPs were more likely to spend less time traveling to appointments, indicating that increasing the supply of APPs with pediatric rheumatology training could reduce the travel burden on parents and children with rheumatic conditions. With more local providers, patients could also benefit from access to rheumatology providers who are knowledgeable about local health care resources (e.g., pharmacies, availability of physical therapy). In addition, when scheduling both routine and urgent appointments, patients were able to see APPs more quickly than rheumatologists. Once at their appointments, regardless of provider type, respondents reported being satisfied that their providers spent enough time with them.

Moreover, we found that APPs and rheumatologists performed essentially the same function and services for their patients, including physical exams, prescribing medications, educating patients, and recommending treatments. And regardless of primary provider type, patients rated their overall experience of care similarly high. We found no differences by provider on care coordination, provider communication, and access-to-care composite scores. Finally, respondents' overall rating of their APPs or rheumatologists was high, but ratings by those seeing rheumatologists were significantly higher.

Our findings support previous research in primary care, obstetrics and gynecology, and orthopedics, which have generally found no difference in satisfaction with the care provided by NPs, PAs, or physicians (32,33), with 1 exception. One study found a significant difference between provider types, with patients of physicians showing higher satisfaction (compared with patients who saw nurse practitioners) with their providers' technical skill, personal manner, and time spent with them; however, the researchers suggested that this difference was small. Other satisfaction factors showed no difference between the 2 provider types (16).

Overall, we found that patient attitudes toward APPs were positive regardless of provider type, although APP patients held more positive overall attitudes toward APPs than did rheumatologist patients. In addition, patients of rheumatologists demonstrated a willingness to accept appointments with APPs in order to be seen earlier. We also found that patients of APPs were more likely to prefer seeing and recommending APPs to others than were patients of rheumatologists. Patients of both provider types had a high level of trust in APPs.

Our results confirm previous research that shows many patients will accept care provided by an APP over that of a physician, depending on the context. For example, Dill et al found that nearly half of patients surveyed would either select an APP over a physician or had no preference when asked to choose a new primary care provider (34). However, more than half of patients would accept earlier appointments with an APP over waiting for an appointment with a physician for acute care. In addition, patients who had previous exposure to APPs were more likely to choose an appointment with an APP over a physician (34). A survey of emergency department patients found that >50% of patients were willing to see APPs in the context of a minor injury or illness, but patient acceptance declined when the injury or illness was moderate or severe (35).

Although this study fills a gap in studying patient attitudes toward working with APPs, it does have some limitations. First, because of the low prevalence of the autoimmune conditions included in the survey and the relatively few APP providers in rheumatology care, respondents were recruited through an opt-in panel rather than from the broader population. Recruiting through an opt-in panel was determined to be the only feasible approach to targeting this population, and the data were weighted to a similar population (as identified through the NHIS). In addition, because we oversampled in more densely populated areas with APPs to ensure sufficient representation of their patients, there are fewer respondents representing the experience of patients in rural areas. People living in rural areas may find that getting autoimmune care from a primary care physician is more feasible than traveling long distances for specialty care.

Finally, the survey was fielded contemporaneously with community and business closures early in the COVID-19 pandemic, which could have impacted participant responses. Because major health impacts of the pandemic were in a few major cities at this time, namely the Northeast and Northwest, any overall health impacts would likely have been limited. In advance of fielding the survey, we reviewed all questions with a focus on whether the pandemic might affect participant responses, identifying only those questions asking about the respondent's mental or emotional health. We do not believe that the pandemic meaningfully affected the fundamental conclusions of the present study.

In conclusion, because a substantial portion of people with autoimmune conditions seem receptive to APPs, increasing the supply of APPs could improve access to care and help ameliorate the current and projected workforce shortages. This is particularly promising because previous research on APPs in rheumatology practice found that APPs functioned independently, with nearly two-thirds (61%) of patient visits covered by APPs alone (6). With patients receptive to APPs, and these providers able to shoulder a significant proportion of patient care responsibilities, providing APPs with specialty training and incorporating them into rheumatology practice could improve access to care for people with rheumatologic conditions, particularly those in areas of the country where rheumatology care is difficult to schedule or unavailable.

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. Frazier 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. Frazier, Paez, Creek, Vinci.

Acquisition of data. Frazier, Paez.

Analysis and interpretation of data. Frazier, Paez, Amolegbe, Hasanbasri.

REFERENCES

1. 2015 American College of Rheumatology Workforce Study Group. Workforce study of rheumatology specialists in the United States. 2016. URL: https://www.rheumatology.org/portals/0/files/ACR-Workforce-Study-2015.pdf.
2. Dinse GE, Parks CG, Weinberg CR, Co CA, Wilkerson J, Zeldin DC, et al. Increasing prevalence of antinuclear antibodies in the United States. Arthritis Rheumatol 2020;72:1026–35.
3. Cooper RA, Henderson T, Dietrich CL. Roles of nonphysician clinicians as autonomous providers of patient care. JAMA 1998;280:795–802.
4. Hooker RS. The extension of rheumatology services with physician assistants and nurse practitioners. Best Pract Res Clin Rheumatol 2008;22:523–33.
5. Solomon DH, Bitton A, Fraenkel L, Brown E, Tsao P, Katz JN. Roles of nurse practitioners and physician assistants in rheumatology practices in the US. Arthritis Care Res (Hoboken) 2014;66:1108–13.
6. Solomon DH, Fraenkel L, Lu B, Brown E, Tsao P, Losina E, et al. Comparison of care provided in practices with nurse practitioners and physician assistants versus subspecialist physicians only: a cohort study of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2015;67:1664–70.
7. Subramanian U, Kerr EA, Klamerus ML, Zikmund-Fisher BJ, Holleman RG, Hofer TP. Treatment decisions for complex patients: differences between primary care physicians and midlevel providers. Am J Manag Care 2009;15:373–80.
8. Tsai CL, Sullivan AF, Ginde AA, Camargo CA Jr. Quality of emergency care provided by physician assistants and nurse practitioners in acute asthma. Am J Emerg Med 2010;28:485–91.
9. Lichtenstein BJ, Reuben DB, Karlamangla AS, Han W, Roth CP, Wenger NS. Effect of physician delegation to other healthcare providers on the quality of care for geriatric conditions. J Am Geriatr Soc 2015;63:2164–70.
10. Muench U, Guo C, Thomas C, Perloff J. Medication adherence, costs, and ER visits of nurse practitioner and primary care physician patients: evidence from three cohorts of Medicare beneficiaries. Health Serv Res 2019;54:187–97.
11. Agarwal A, Zhang W, Kuo Y, Sharma G. Process and outcome measures among COPD patients with a hospitalization cared for by an advance practice provider or primary care physician. PloS One 2016;11:e0148522.
12. Kuo YF, Goodwin JS, Chen NW, Lwin KK, Baillargeon J, Raji MA. Diabetes mellitus care provided by nurse practitioners vs primary care physicians. J Am Geriatr Soc 2015;63:1980–8.
13. Kuo YF, Chen NW, Baillargeon J, Raji MA, Goodwin JS. Potentially preventable hospitalizations in Medicare patients with diabetes: a comparison of primary care provided by nurse practitioners versus physicians. Med Care 2015;53:776–83.
14. Liu CF, Hebert PL, Douglas JH, Neely EL, Sulc CA, Reddy A, et al. Outcomes of primary care delivery by nurse practitioners: utilization, cost, and quality of care. Health Serv Res 2020;55:178–89.
15. Yang Y, Long Q, Jackson SL, Rhee MK, Tomolo A, Olson D, et al. Nurse practitioners, physician assistants, and physicians are comparable in managing the first five years of diabetes. Am J Med 2018;131:276–83.e2.
16. Mundinger MO, Kane RL, Lenz ER, Totten AM, Tsai WY, Cleary PD, et al. Primary care outcomes in patients treated by nurse practitioners or physicians: a randomized trial. JAMA 2000;283:59–68.
17. Martin-Misener R, Harbman P, Donald F, Reid K, Kilpatrick K, Carter N, et al. Cost-effectiveness of nurse practitioners in primary and specialised ambulatory care: systematic review. BMJ Open 2015;5:e007167.
18. Ohman-Strickland PA, Orzano AJ, Hudson SV, Solberg LI, DiCiccio-Bloom B, O'Malley D, et al. Quality of diabetes care in family medicine practices: influence of nurse-practitioners and physician's assistants. Ann Fam Med 2008;6:14–22.
19. Tapper EB, Hao S, Lin M, Mafi JN, McCurdy H, Parikh ND, et al. The quality and outcomes of care provided to patients with cirrhosis by advanced practice providers. Hepatology 2020;71:225–34.
20. Rymer JA, Chen AY, Thomas L, Stafford J, Enriquez JR, Goyal A, et al. Advanced Practice Provider Versus Physician-Only Outpatient Follow-Up After Acute Myocardial Infarction. J Am Heart Assoc 2018;7:e008481.
21. Virani SS, Maddox TM, Chan PS, Tang F, Akeroyd JM, Risch SA, et al. Provider type and quality of outpatient cardiovascular disease care: insights from the NCDR PINNACLE registry. J Am Coll Cardiol 2015;66:1803–12.
22. Dynata. Panel Book. 2020. URL: https://www.dynata.com/panel-book-form/?cid=7D84CE0F-D852-EC11-8C62-000D3A9DE12E&cid=7D84CE0F-D852-EC11-8C62-000D3A9DE12E&utm_source=google&utm_medium=cpc&utm_campaign=panel_book&utm_content=ad_panel_book&gclid=CjwKCAjwtIaVBhBkEiwAsr7-czH0eHtL35csfLrMEV3ObCitG0BaYZBsL1Tj-J-vzW2u-MSYetm5JhoCcuMQAvD_BwE.
23. Nikiphorou E, Radner H, Chatzidionysiou K, Desthieux C, Zabalan C, van Eijk-Hustings Y, et al. Patient global assessment in measuring disease activity in rheumatoid arthritis: a review of the literature. Arthritis Res Ther 2016;18:251.
24. Tugwell P, Idzerda L, Wells GA. Generic quality-of-life assessment in rheumatoid arthritis. Am J Manag Care 2008;14:234.
25. Orbai AM, Ogdie A. Patient-reported outcomes in psoriatic arthritis. Rheum Dis Clin North Am 2016;42:265–83.
26. Van Tubergen A, Black PM, Coteur G. Are patient-reported outcome instruments for ankylosing spondylitis fit for purpose for the axial spondyloarthritis patient? A qualitative and psychometric analysis. Rheumatology (Oxford) 2015;54:1842–51.

27. Holloway L, Humphrey L, Heron L, Pilling C, Kitchen H, Højbjerre L, et al. Patient-reported outcome measures for systemic lupus erythematosus clinical trials: a review of content validity, face validity and psychometric performance. Health Qual Life Outcomes 2014;12:116.
28. Moorthy LN, Peterson MG, Harrison MJ, Onel KB, Lehman TJ. Physical function assessment tools in pediatric rheumatology. Pediatr Rheumatol Online 2008;6:9.
29. Agosta LJ. Psychometric evaluation of the Nurse Practitioner Satisfaction Survey (NPSS). J Nurs Meas 2009;17:114–33.
30. Centers for Disease Control and Prevention. National Center for Health Statisticks. 2018 National Health Interview Survey Questionnaire. URL: https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm.
31. Maruish ME, editor. User's manual for the SF-12v2 Health Survey. 3rd ed. Lincoln, RI: Quality Metric Incorporated; 2012.
32. Hooker RS, Potts R, Ray W. Patient satisfaction: comparing physician assistants, nurse practitioners, and physicians. Perm J 1997;1: 38–42.
33. Hooker RS, Cipher DJ, Sekscenski E. Patient satisfaction with physician assistant, nurse practitioner, and physician care: a national survey of medicare beneficiaries. J Clin Outcomes Manag 2005;12: 88–92.
34. Dill MJ, Pankow S, Erikson C, Shipman S. Survey shows consumers open to a greater role for physician assistants and nurse practitioners. Health Aff (Millwood) 2013;32:1135–42.
35. Larkin GL, Hooker RS. Patient willingness to be seen by physician assistants, nurse practitioners, and residents in the emergency department: does the presumption of assent have an empirical basis? Am J Bioeth 2010;10:1–10.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1602–1607
DOI 10.1002/acr.24645

A Critical Look at Race-Based Practices in Rheumatology Guidelines

Rose McKeon Olson and Candace H. Feldman

Objective. To assess how race has been incorporated into rheumatology practice guidelines, including how race is defined and used in diagnostic and treatment recommendations.

Methods. We searched race and ethnicity terms in all clinical practice guidelines from the American College of Rheumatology (ACR) and European Alliance of Associations for Rheumatology (EULAR) that were published between 2010 and 2020 and publicly available on professional society websites. Findings were summarized and assessed through standardized data abstraction forms. Key themes were identified through a thematic analysis approach.

Results. A total of 23 ACR clinical practice guidelines and 42 EULAR recommendations were reviewed. In total, 16 of 65 (25%) of the guidelines used race terms in their text. No guideline clearly defined race, and race was often conflated with ethnicity and/or genetic ancestry. Reported racial categories varied substantially by guideline and often used classifications that oversimplified and excluded non-White races. Research with insufficient racial diversity was used to make race-based recommendations for Black patients that may not be generalizable. Additionally, recommendations using research on predominantly White populations reinforced data of White populations as normative and perpetuated race-based stereotypes, especially for rare diseases. Structural causes of identified racial disparities were not discussed in clinical guidelines.

Conclusion. There is an urgent need for standardized race reporting in rheumatology. Recommendations are provided to enhance consistency and accuracy of race and ethnicity terms, mitigate conflation of race with ethnicity or genetic ancestry, encourage a critical reanalysis of race-based diagnostic tools and treatment options, and better address the structural causes of racial disparities.

INTRODUCTION

The field of medicine is starting to be held accountable for how its institutions, practice guidelines, and clinical-decision tools have perpetuated structural racism (1). This moment in the field has led to well-founded scrutiny of how health professional organizations define race and provide recommendations for race-based diagnoses and treatments.

The use of race in medicine is complex; race correlates with a number of dynamic social, cultural, and economic factors that can have a powerful impact on disease (2). Understanding racial differences in medical research has potential benefits, such as recognizing those at higher risk for certain diseases to aid in earlier detection and intervention, increasing awareness of racism in medicine, targeting efforts to alleviate racial health disparities, and identifying populations that merit reparations. However, when race is not well-defined, systematically collected, and reported, it can lead to race-based clinical practices and research that preserve and perpetuate structural racism. We aimed to explore these aspects further by assessing how race has been incorporated into rheumatology practice guidelines.

MATERIALS AND METHODS

We searched race and ethnicity terms in all clinical practice guidelines from the American College of Rheumatology (ACR) and European Alliance of Associations for Rheumatology (EULAR) that were published between 2010 and 2020 and publicly available on their professional society websites (see Supplementary Appendix, available on the *Arthritis Care &*

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

Rose McKeon Olson, MD, Candace H. Feldman, MD, ScD: Brigham and Women's Hospital, Boston, Massachusetts.

Dr. Feldman's work was supported by the NIH (grants K23-AR-071500 and 3K23-AR-071500-04S1 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases) and the Office of Minority Health (grants CPIMP171141-01-00 and CPIMP18116801-00).

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

Address correspondence to Rose McKeon Olson, MD, 45 Francis Street, Boston, MA 02115. Email: rolson@bwh.harvard.edu.

Submitted for publication January 7, 2021; accepted in revised form May 4, 2021.

SIGNIFICANCE & INNOVATIONS

- Race is commonly described in rheumatology clinical guidelines, yet race reporting is unstandardized and often conflates race with ethnicity and genetic ancestry.
- Review of race reporting in rheumatology guidelines revealed multiple inaccuracies, oversimplifications, and stereotypes that may perpetuate structural racism, bias in care, and disparities in outcomes.
- Findings suggest that there should be standardized criteria for race reporting in rheumatology guidelines, and only those studies that uphold the highest degree of scrutiny for research should be included in recommendations that support race-based diagnostic and treatment strategies.
- Researchers and guidelines authors should investigate the root causes of racial disparities when they are identified.

Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24645). The following search terms were used: race, racial, ethnicity, ethnic, African American, Black, Caucasian, White, Latino, Latina, Hispanic, Asian, Native, Indigenous, and Pacific Islander. Inclusion criteria were the following: 1) reported in ACR or EULAR guidelines, 2) available on the professional website, and 3) included predesignated race term. Full-text assessments of guidelines that met criteria were performed and data were extracted (RMO), and decisions were made about study inclusion and exclusion (RMO and CHF). Data were summarized into standardized data abstraction forms and synthesized with a thematic synthesis approach. A spreadsheet was created of all the data extracted from these studies, and thematic analysis methods were used to develop broad themes.

RESULTS

A total of 23 ACR clinical practice guidelines and 42 EULAR recommendations were reviewed (RMO and CHF). In total, 16 of 65 (25%) guidelines screened used race terms in their text, including 4 guidelines on systemic lupus erythematosus (3–6), 4 on glucocorticoid-induced osteoporosis (7–10), 3 on gout (11–13), 3 on rheumatoid arthritis (14–16), 1 on giant cell arteritis (GCA) (17), and 1 on idiopathic inflammatory myopathies (18) (Table 1). Major themes emerged in our analysis of the clinical guidelines and are described below.

Race-based issues identified in rheumatology guidelines. *Ambiguous use of race, ethnicity, and genetic ancestry.* Race was found to be heterogeneously reported in rheumatologic clinical guidelines, and distinctions between race, ethnicity, and genetic ancestry were infrequently made. Race and ethnicity terms were often used interchangeably in the reviewed guidelines (e.g., using the term "ethnicity/race"), and these categories were not defined in the text (3,4,8,11). Additionally, race and genetics were conflated in several guidelines. Several guidelines on the clinical management of gout state that allopurinol has different risks based upon the patient's race (11–13), because the HLA–B*5801 allele is associated with an increased risk for allopurinol-hypersensitivity syndrome (AHS) in certain Asian and African-American populations. However, these 2 concepts are distinct – it is the presence of the genetic variant that directly determines risk of AHS, not race. Race is a crude approximation of a person's genetic makeup and can lead to clinical mismanagement when used as the only indicator of the likelihood of genetic variance.

Race is the socially ascribed interpretation of how someone looks, and its definition changes by geography and time period (19). This social construct does not organize neatly into discrete genetic groups. Someone socially defined as Black in America may have their genetic ancestry trace for generations back to Ethiopia, Brazil, Australia, or the Dominican Republic. Genetic ancestry, on the other hand, is the genetic origin of one's population (20). Genomic studies have shown that genetic ancestry is a better predictor of genetic variants (alleles).

Race is unlikely to measure 1 discrete variable; rather, it correlates with an amalgam of genetic, ancestral, social, cultural, and economic factors. For racial minorities, race is the embodiment of lived experiences of structural racism (21). Using race, a complex social construct, as a proxy for genetic makeup can lead

Table 1. Race reporting in clinical guidelines by rheumatologic condition and purpose for inclusion

	Total guidelines reporting on race	Purpose: diagnostics	Purpose: treatment selection	Purpose: other
Systemic lupus erythematosus	4	0	3	1 (Adherence)
Glucocorticoid-induced osteoporosis	4	4	0	0
Gout	3	0	3	0
Rheumatoid arthritis	3	0	2	1 (Education)
Giant cell arteritis	1	1	0	0
Idiopathic inflammatory myopathies	1	1	0	0

to misclassification and error. Nevertheless, there are known genetic variants, such as the APOL1 gene, that appear at higher rates in certain races, indicating there are situations where self-reported race can provide insight into the likelihood of genetic variants (22). Given this nuance and complexity, there is significant opportunity to improve shared understanding of race and its appropriate use in rheumatology. Careful consideration of the interrelated biological and structural factors that influence race will be necessary to ensure appropriate use of race in research and clinical practice and to mitigate structural racism in medicine.

Oversimplified race and ethnicity categories. Guidelines often oversimplified race and ethnicity categories, in part due to the absence of clearly defined categories and small sample sizes in the referenced studies (7–9). For example, guidelines on glucocorticoid-induced osteoporosis cited the Fracture Risk Assessment Tool calculator, which uses just 4 racial categories to assign risk, including Black, White, Hispanic, and Asian. The expert panel of 1 of the glucocorticoid-induced guidelines further merged these categories into Black versus non-Black. The tables include 4 groups, including White women, White men, Black women, and Black men, "to provide clinicians with examples of typical patients to match their individual patient with the most closely fitting category" (8). Lack of racial representation in these categories was acknowledged and attributed to limited available data.

Such oversimplifications fail to describe true population diversity. When categories were combined, it was consistently non-White races who lost granularity. A racial group designated as "Asian" fails to capture the unique and varying genetic ancestry and environmental and structural exposures that a recent Hmong immigrant embodies compared to a third generation Indian-American. The oversimplified ways race is described in clinical guidelines, as a reflection of the way it is presented in primary research studies, obscures the nuances between racial experiences and can further entrench racial stereotyping in medicine.

Lack of racial representation in research. There was frequent underrepresentation of racial minorities in the scientific research used to create guideline recommendations, which can lead to several racial inequities (16,18). When minorities were the primary population studied, guidelines often hesitated to recommend the findings universally. For example, a guideline on lupus nephritis states, "The absence of robust evidence on calcineurin-inhibitors in non-Asian populations…has led the committee to adopt a more cautious attitude," and "data have to be corroborated with longer duration studies in multiethnic populations" (4). However, guidelines based upon data from majority White populations often did not state this qualification, and there appeared to be less hesitancy to recommend the therapy across racial categories (3,16–18). This pattern reinforces data from White populations as normative while data from minorities require further validation and should interpreted with caution.

Underrepresentation of racial minorities can also lead to inaccurate racial stereotypes that may delay or miss diagnoses when they occur among racial minorities. For decades, the only large population-based study of polymyalgia rheumatica and GCA in the US was from the predominantly White geographic region of Olmsted County, Minnesota (23). This study led to the common teaching that GCA is largely a disease affecting White individuals. However, more recent and racially representative studies have demonstrated similar prevalence of GCA between Black and White patients (24). Despite this, ethnicity was still included as an important clinical factor in diagnosis of GCA in the ACR rheumatology guideline (17). Racial stereotypes become especially engrained for rare diseases where clinicians and students rely more heavily on research and educational examples to guide clinical reasoning.

Issues with race-based recommendations. The review identified race-based clinical recommendations for Black patients that used limited data that may not be broadly generalizable to entire racial groups. In a study cited by the lupus nephritis guidelines, the efficacy of mycophenolate mofetil (MMF) was compared to intravenous cyclophosphamide (IVC) for induction treatment of lupus nephritis (25). The study concluded that MMF may be more efficacious in Black patients, despite sample sizes of n = 26 for MMF and n = 20 for IVC among Black patients. Additionally, only 50% of Black patients received IVC treatment for the study period compared to 72% of the entire IVC treatment population. Efficacy results for the Black participants alone were not statistically different, but when researchers combined data from Black patients with a racial group defined as "other" there was statistical difference, from which authors concluded that MMF may be more efficacious than IVC in Black patients. These data were used to support claims found in several rheumatology guidelines that there are, "possible ethnic/racial differences," suggesting that MMF may be more efficacious in African–Americans (3,6). As a result, this is broadly accepted in clinical practice and part of the teaching provided to rheumatology trainees.

There has been growing controversy over the use of race in clinical decision-making tools and treatment strategies due to evidence that it can result in racially biased care. For example, many institutions have abandoned the use of the estimated glomerular filtrate rate (eGFR) due to awareness that race is a social construct, and evidence that it negatively impacts access to care (such as renal transplants) for Black patients with renal disease (2,26). In response, the American Society for Nephrology with the National Kidney Foundation are developing new joint guidance on the use of race in diagnosis of kidney disease. Similarly, the American Academy of Cardiology has faced conflict over the use of race-based antihypertensive regimens, including dissent among panel experts, and are currently calling for additional research to clarify guidance (27). These examples indicate that if race-based diagnostic tools or treatment strategies are being considered for inclusion in clinical guidelines, there should be

rigorous scrutiny of the evidence to avoid additional racial biases in clinical care, which means ensuring that race is accurately defined and that racial data are high quality and consistently and equitably collected.

Lack of rigorous analysis of racial disparities. Once differences between racial groups were identified, guidelines largely stopped at reporting the racial difference without further analysis of what caused the disparities seen. Guidelines often placed commentary about racial differences in the discussion section where a variety of reasons for differences were postulated, such as yet-to-be-discovered biological differences or hypotheses about varied social exposures. This common academic practice attempts to absolve researchers, guideline authors, and clinicians alike of further investigation that may reveal the root cause of the inequity. Once the structural inequities are identified, such as lack of access to appropriate housing, transportation, health care, healthy food, and voting centers, these areas can be targeted to relieve racial health disparities and improve quality of care for patients of color (28).

DISCUSSION

Our review of the representation of race in rheumatologic guidelines has led us to offer a set of recommendations. These recommendations are geared toward both future guidelines and the primary research that informs these guidelines.

First, we recommend the use of accepted, accurate definitions of race, ethnicity, and genetic ancestry and stating their appropriate use and limitations. Accurate definitions of race, ethnicity, and genetic ancestry should be systematically employed in rheumatology guidelines, along with appropriate use of racial and ethnic terms, drawing upon expert guidance (19,20). If race is recommended to be used as a proxy for the likelihood of genetic differences (e.g., assumption of presence of HLA–B*5801 based on self-reported Asian race), assumptions and limitations of this approach should be clearly stated in guideline texts. Our recommended definitions are the following: 1) Race: socially defined category based on the interpretation of how someone looks that varies by geography and time period and influences socioeconomic positioning and access to societal opportunities and resources (19,20); 2) Ethnicity: shared culture of a population, rooted in common language, values, norms, and/or traditions (19,20); 3) Genetic ancestry: genetic origin of one's population that is a better predictor than race for the likelihood of genetic variants (20).

Secondly, nuanced, standardized categories of race and ethnicity and acknowledgment of their limitations in capturing true population diversity is recommended. Where possible, race and ethnicity desegregation should be utilized in order to collect and analyze nuanced differences between subgroups. For example, Asian may become Chinese, Indian, Korean, Lao, Hmong, or Vietnamese, and Middle Eastern and North African may be included as separate racial categories (19). We also recommend requiring that racial and ethnic determinations be self-identified, not determined through "physician eyeball." If self-identified races or ethnicities do not fall discretely into existing categories, including them as "other" along with the self-identified category (19) is recommended.

Third, we recommend adding the country of origin to collected demographic data. Geographic origin is an important designation distinct from race, genetic ancestry, and ethnicity (19). It provides important information on regionalization that can have several implications for data generalizability and can improve understanding of observed subgroup differences.

Our fourth recommendation is to increase racial diversity in research. Prior studies demonstrate that the racial and ethnic composition of clinical trial enrollees are not consistent with the demographic distribution of rheumatic conditions (29). In order for guidelines to make appropriate recommendations, the diversity of clinical trial enrollees needs to be improved. Guideline authors should advocate for researchers to adjust inclusion/exclusion criteria in trials to include racially and ethnically diverse, representative cohorts and identify mechanisms to increase recruitment and retention of underrepresented racial and ethnic groups. When data is unavailable or underrepresented from racial and ethnic groups, this limitation should be clearly stated.

Rigorous analysis of any race-based recommendation, including potential unintended consequences, is our fifth recommendation. Such an analysis should include the removal of current race-based recommendations based upon studies that do not provide sufficient evidence or, at minimum, clearly citing limitations and advocating for larger, more diverse studies. When race-based clinical recommendations are being considered, accurate racial definitions and the highest degree of scrutiny of methodology, data quality, and its conclusions should be ensured. The unintended consequences of race-based recommendations, such as differential access to care based on race, must be considered. It is important to recognize the responsibility of rheumatologists to actively participate in ongoing discussions and research about race-based recommendations that apply directly to patients with chronic rheumatic conditions (such as eGFR adjustments by race).

Our sixth recommendation is to collect and report data on exposure to structural racism and its mechanisms. Researchers should collect and report data on patient experiences with racism and discrimination as well as exposure to various structural inequities, such as racial residential segregation and access to stable housing, transportation, health care, healthy food, and voting centers. These data can help identify the underlying structural causes of observed racial disparities in more concrete, actionable ways. Where disparities exist but research on its causes has not been performed, guideline committees should advocate for such research and consider including these considerations in future recommendations.

In conclusion, the findings of our investigation suggest there is an urgent need for standardized race reporting in rheumatology scholarship. This would enhance consistency and accuracy in racial terms, mitigate conflation of race with genetics, and encourage a critical reanalysis of race-based diagnostic tools and treatment options in rheumatology. Approximately 75% of reviewed guidelines did not mention race, when likely racial differences and disparities exist. Rather than resorting to color-blindness, guidelines should be grounded in data from racially diverse research populations and recommendations that accurately define and describe racial differences. When differences among races are found, we urge deep analysis of the social, environmental, and structural root causes that contributed to the racial disparities. We call on physicians, researchers, and professional organizations alike to conduct and support future research on racism in rheumatology research and include these findings in future guidelines.

AUTHOR CONTRIBUTIONS

Drs. Olson and Feldman 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. Olson 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. Olson, Feldman.
Acquisition of data. Olson.
Analysis and interpretation of data. Olson, Feldman.

REFERENCES

1. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight — reconsidering the use of race correction in clinical algorithms. N Engl J Med 2020;383:874–82.
2. Powe NR. Black kidney function matters: use or misuse of race? JAMA 2020;324:737–8.
3. Fanouriakis A, Kostopoulou M, Cheema K, Anders H-J, Aringer M, Bajema I, et al. 2019 update of the joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of lupus nephritis. Ann Rheum Dis 2020;79: 713–23.
4. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis J, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Ann Rheum Dis 2019;78:736–45.
5. Andreoli L, Bertsias GK, Agmon-Levin N, Brown S, Cervera R, Costedoat-Chalumeau, et al. EULAR recommendations for women's health and the management of family planning, assisted reproduction, pregnancy and menopause in patients with systemic lupus erythematosus and/or antiphospholipid syndrome. Ann Rheum Dis 2017;76: 476–85.
6. Bertsias GK, Tektonidou M, Amoura Z, Aringer M, Bajema I, Berden JH, et al. Joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of adult and paediatric lupus nephritis. Ann Rheum Dis 2012;71: 1771–82.
7. Buckley L, Guyatt G, Fink HA, Cannon M, Grossman J, Hansen K, et al. 2017 American College of Rheumatology guideline for the prevention and treatment of glucocorticoid-induced osteoporosis. Arthritis Rheumatol 2017;69:1521–37.
8. Grossman JM, Gordon R, Ranganath VK, Deal C, Caplan L, Chen W, et al. American College of Rheumatology 2010 recommendations for the prevention and treatment of glucocorticoid-induced osteoporosis. Arthritis Care Res (Hoboken) 2010;62:1515–26.
9. Lems WF, Dreinhöfer KE, Bischoff-Ferrari H, Blauth M, Czerwinski E, da Silva J, et al. EULAR/EFORT recommendations for management of patients older than 50 years with a fragility fracture and prevention of subsequent fractures. Ann Rheum Dis 2017;76:802–10.
10. Duru N, van der Goes MC, Jacobs JW, Andrews T, Boers M, Buttgereit F, et al. EULAR evidence-based and consensus-based recommendations on the management of medium to high-dose glucocorticoid therapy in rheumatic diseases. Ann Rheum Dis 2013;72: 1905–13.
11. FitzGerald JD, Dalbeth N, Mikuls T, Brignardello-Petersen R, Guyatt G, Abeles AM, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Rheumatol 2020;72: 879–95.
12. Khanna D, FitzGerald JD, Khanna PP, Bae S, Singh MK, Neogi T, et al. 2012 American College of Rheumatology guidelines for management of gout. Part 1: Systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. Arthritis Care Res (Hoboken) 2012;64:1431–46.
13. 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.
14. Smolen JS, Landewé RB, Bijlsma JW, Burmester GR, Dougados M, Kerschbaumer A, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis 2020; 79:685–99.
15. Smolen JS, Landewé R, Bijlsma J, Burmester J, Chatzidionysiou K, Dougados M, Nam J, 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.
16. Zangi HA, Ndosi M, Adams J, Andersen L, Bode C, Boström C, et al. EULAR recommendations for patient education for people with inflammatory arthritis. Ann Rheum Dis 2015;74:954–62.
17. Ehlers L, Askling J, Bijlsma HW, Cid MC, Cutolo M, Dasgupta B, et al. 2018 EULAR recommendations for a core data set to support observational research and clinical care in giant cell arteritis. Ann Rheum Dis 2019;78:1160–6.
18. Lundberg IE, Tjärnlund A, Bottai M, Werth VP, Pilkington C, de Visser M, et al. 2017 European League Against Rheumatism/ American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Arthritis Rheumatol 2017;69:2271–82.
19. Flanagin A, Frey T, Christiansen SL, Bauchner H. The reporting of race and ethnicity in medical and science journals: comments invited. JAMA 2021;325:1049.
20. Borrell LN, Elhawary JR, Fuentes-Afflick E, Witonsky J, Bhakta N, Wu AH, et al. Race and genetic ancestry in medicine — a time for reckoning with racism. N Engl J Med 2021;384:474–80.
21. Krieger N. Embodiment: a conceptual glossary for epidemiology. J Epidemiol Community Health 2005;59:350–5.
22. Parsa A, Kao WH, Xie D, Astor BC, Li M, Hsu C-Y, et al. APOL1 risk variants, race, and progression of chronic kidney disease. N Engl J Med 2013;369:2183–96.
23. Doran MF, Crowson CS, O'Fallon WM, Hunder GG, Gabriel SE. Trends in the incidence of polymyalgia rheumatica over a 30 year period in Olmsted County, Minnesota, USA. J Rheumatol 2002;29: 1694–7.

24. Gruener AM, Poostchi A, Carey AR, et al. Association of giant cell arteritis with race. JAMA Ophthalmol 2019;137:1175–9.
25. Isenberg D, Appel GB, Contreras G, Dooley MA, Ginzler EM, Jayne D, et al. Influence of race/ethnicity on response to lupus nephritis treatment: the ALMS study. Rheumatol Oxf Engl 2010;49:128–40.
26. Ahmed S, Nutt CT, Eneanya ND, Reese PP, Sivashanker K, Morse M, et al. Examining the potential impact of race multiplier utilization in estimated glomerular filtration rate calculation on African-American care outcomes. J Gen Intern Med 2021;36:464–71.
27. Colvin CL, King JB, Oparil S, Wright Jr JT, Ogedegbe G, Mohanty A, et al. Association of race/ethnicity-specific changes in antihypertensive medication classes initiated among Medicare beneficiaries with the Eighth Joint National Committee Panel member report. JAMA Netw Open 2020;3:e2025127.
28. Groos M, Wallace M, Hardeman R, Theall K. Measuring inequity: a systematic review of methods used to quantify structural racism. J Health Disparities Res Pract 2018. URL: https://digitalscholarship.unlv.edu/jhdrp/vol11/iss2/13
29. Falasinnu T, Chaichian Y, Bass MB, Simard JF. The representation of gender and race/ethnic groups in randomized clinical trials of individuals with systemic lupus erythematosus. Curr Rheumatol Rep 2018; 20:20.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1608–1613
DOI 10.1002/acr.24623

Body Composition in Patients With Psoriatic Arthritis and Changes During Interleukin-12/Interleukin-23 Inhibition

Julien Paccou,[1] Wallis Bavière,[1] Elisabeth Sornay-Rendu,[2] Pawel Szulc,[2] Nassima Ramdane,[1] Bernard Cortet,[1] Roland Chapurlat,[2] and René-Marc Flipo[1]

Objective. Little is known about body composition in patients with psoriatic arthritis (PsA). Our objective was to compare body composition parameters in PsA patients and healthy controls and then investigate the effects of ustekinumab (UST) on body composition in patients with PsA.

Methods. At baseline, 30 PsA patients were compared cross-sectionally with 60 healthy controls without PsA, matched for age, sex, menopausal status, and body mass index (BMI). Thirty active PsA patients treated with UST were included in a 6-month open follow-up study. Body composition parameters were measured at baseline and 6 months of treatment.

Results. Body composition parameters were different in PsA patients compared to healthy controls; in PsA patients, total and appendicular lean mass were lower ($P = 0.013$ and $P = 0.010$, respectively), whereas total fat mass was higher ($P < 0.001$). In 30% of the PsA patients, skeletal muscle mass was below the cutoff for low muscle quantity (men 7.26 kg/m^2, women 5.5 kg/m^2), whereas no such change was observed in the control group. After 6 months of treatment with UST, there was no significant change in BMI in 18 of the PsA patients. Total lean mass decreased slightly ($P = 0.046$), whereas fat mass tended to increase, but not significantly. No significant changes in appendicular lean mass and skeletal muscle mass index were observed.

Conclusion. In this study, we found that PsA patients had higher fat mass and lower lean mass than healthy controls. At 6-months of treatment, total lean mass decreased slightly, whereas fat mass tended to increase, but not significantly.

INTRODUCTION

Inflammatory rheumatic diseases, and particularly rheumatoid arthritis (RA), are characterized by adverse changes in body composition and bone mineral density (BMD). Patients with RA are usually found to have lower lean mass and BMD and higher adiposity than controls (1,2). Many factors may influence body composition in patients with inflammatory rheumatic diseases. These include aging, nutrition, physical activity, disease activity, and disease-modifying antirheumatic drugs (DMARDs).

A better understanding of body composition and BMD changes in patients with inflammatory rheumatic diseases, including those undergoing treatment with DMARDs, is important because of the many potential implications in terms of outcome, such as sarcopenia, osteoporosis, and cardiometabolic risk (3). Where biologic DMARDs (bDMARDs) are concerned, most of the data at our disposal relate to the impact of anti–tumor necrosis factor (anti-TNF) drugs in patients with RA and spondyloarthritis. In these patients, anti-TNF drugs are associated with an increase in body weight, body fat (especially visceral adiposity), and BMD, but their effect on lean mass is more controversial (4). On the other hand, treatment with the interleukin-6 (IL-6) inhibitor, tocilizumab, is associated with an increase in body weight and lean mass, but with no change in fat mass in RA patients (5).

Psoriatic arthritis (PsA), beyond joint and skin involvement, is associated with many comorbid conditions (e.g., obesity, diabetes mellitus, and dyslipidemia), thus adding to the burden of disease (6). The literature provides little information about body composition in patients with PsA, and limited data are available regarding changes in body composition in patients treated with bDMARDs (7,8). Nevertheless, there have been numerous advances in the treatment of PsA in recent years, and

Supported by an unrestricted grant from Janssen France.
[1]Julien Paccou, MD, PhD, Wallis Bavière, MD, Nassima Ramdane, PhD, Bernard Cortet, MD, PhD, René-Marc Flipo, MD, PhD: University of Lille, CHU Lille, Lille, France; [2]Elisabeth Sornay-Rendu, MD, Pawel Szulc, PhD, Roland Chapurlat, MD, PhD: INSERM UMR 1033, Université de Lyon and Hôpital E. Herriot, Lyon, France.

No potential conflicts of interest relevant to this article were reported.
Address correspondence to Julien Paccou, MD, PhD, Service de rhumatologie, Hôpital Roger Salengro, Rue Emile Laine, CHRU, 59037 Lille cedex, France. Email: julien.paccou@chru-lille.fr.
Submitted for publication November 16, 2020; accepted in revised form April 13, 2021.

SIGNIFICANCE & INNOVATIONS

- Patients with psoriatic arthritis (PsA) have higher adiposity and lower lean mass than healthy controls.
- No substantial change in body composition in PsA patients undergoing treatment with ustekinumab was observed.

ustekinumab, a humanized IL-12/IL-23 inhibitor, is now commonly used as a biologic antipsoriatic drug (9). Therapeutic agents targeting the IL-23/IL-17 axis, such as ustekinumab, have systemic effects and may influence the bone-fat-muscle interactions, leading to body composition and BMD changes in patients with PsA (10). IL-17 can mediate pleiotropic effects throughout the body since IL-17 receptors are expressed on most cell types. IL-17 contributes to skeletal-muscle contractility defects and weakness. Moreover, as fat mass increases, Th17 cells accumulate in adipose tissue, and IL-17 is likely to have multiple downstream effects on bone cells (11).

For all of these reasons, we sought to compare, cross-sectionally, body composition in PsA patients and healthy controls. We hypothesized that lean mass would be lower, and adiposity higher, in PsA patients than in healthy controls. Then, in a prospective pilot study, we investigated potential changes in body composition, BMD, and bone turnover markers in PsA patients during IL-12/IL-23 inhibition.

MATERIALS AND METHODS

Study design. A total of 42 patients with PsA were deemed eligible for the study and approached. Of these, 12 (28.6%) declined participation. At baseline, all 30 PsA patients were compared cross-sectionally with 60 healthy controls without PsA and then started new treatment with ustekinumab in an open, prospective 6-month follow-up study (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24623).

The study protocol was approved by the local Institutional Review Board (2018-A01552-53), and the study procedures complied with the ethical standards of the relevant institutional and national Human Experimentation Ethics Committees (#CPP 18/055-2). All patients provided their written informed consent.

Study population. Inclusion criteria were the presence of the Classification of Psoriatic Arthritis criteria and an indication for ustekinumab and age ≥18 years. Exclusion criteria were current treatment with oral glucocorticoids >10 mg prednisone/day, patients on or considering a restrictive diet during the study period, patients undertaking or planning to undertake an intense exercise program, a history of treatment with bone active substances such as bisphosphonates, and a weight >160 kg. In patients who were receiving bDMARDs, a 5 half-life wash-out period was required between bDMARDs interruption and inclusion in the study.

Controls. For baseline references, data were obtained from 60 healthy controls without PsA. All control cohorts were from the Center for the Prevention of Bone Disease in Lyon, France: female controls were recruited from the MODAM and OFELY cohorts, and all male controls from the STRAMBO cohort. Our healthy controls were matched with PsA patients for age (±5 years), sex, body mass index (BMI, ±3 kg/m^2), and menopausal status.

Study protocol. Demographic and clinical characteristics were recorded. Patients' disease assessment was rated according to the 28-joint Disease Activity Score (DAS28) adjusted for C-reactive protein (CRP) levels, and the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) and Bath Ankylosing Spondylitis Functional Index (BASFI), depending on the predominant site of involvement. Current use of conventional synthetic DMARDs, nonsteroidal antiinflammatory drugs, and glucocorticoids was determined. Past use of bDMARDs was recorded. BMD in PsA patients was measured at the lumbar spine (L1–L4) and at the nondominant hip by dual-energy X-ray absorptiometry (DXA) scan (HOLOGIC Horizon W S/N 300869M).

For the assessment of body composition parameters, we used the Adult Official Positions of the International Society for Clinical Densitometry as updated in 2019. All PsA patients underwent total body DXA scanning. Fat, lean, and bone masses for the total body and per region (arms, legs, and trunk) were measured and analyzed using the manufacturer's validated software (version 13.6.0.5). Body fat percentage was calculated as the proportion of total fat mass to total mass. Appendicular lean mass (kg) was computed as the sum of the tissue compartment (lean) of both arms and legs. The skeletal muscle mass index was calculated as appendicular lean mass divided by height squared (kg/m^2), and the fat mass index as total fat mass divided by height squared (kg/m^2). Visceral adipose tissue (cm^2) was recorded. Regarding Baumgartner's criteria, a skeletal muscle mass index below the cutoff for low muscle quantity was men 7.26 kg/m^2 and women 5.5 kg/m^2.

Laboratory variables included fasting (at least 8 hours) blood samples, procollagen type 1 intact N-terminal polypeptide, and serum cross laps, measured by chemiluminescence assay using the IDS-iSYS Multi-Discipline Automated Analyzer (Immunodiagnostic Systems). Plasma concentrations of leptin were measured by radio immunologic assay.

Study size. Given the exploratory nature of the study, no formal sample size calculations were performed. Using a posteriori power calculations, we found that with 30 PsA cases and 60 healthy controls, using a 2-sided test with a significance level of 0.05 and 80% power, we could detect, for body fat percentage

(predefined as our primary outcome), an effect size (Cohen's *d*) >0.63, considered in the literature as a medium effect size. For the SD of body fat percentage found in PsA patients (8.6%), the effect size corresponds to a mean difference of 5.4%.

Statistical analysis. Categorical variables are expressed as numbers (percentage) and continuous variables as means ± SDs. The normality of model residuals was assessed graphically and using a Shapiro-Wilk test. Since a difference in BMI was observed between PsA patients and healthy controls (standardized difference >10%), despite matching for BMI, body-composition comparisons were further adjusted for BMI.

Body composition and BMD measurements were available for 30 PsA patients at baseline and 18 patients at 6 months of treatment. Thus, we studied body composition changes during treatment in those 18 patients by comparing baseline and 6-month values using a paired *t*-test or Wilcoxon's signed-rank test, depending on the normality of intrapatient differences.

To assess the possibility of a selection bias with regard to those PsA patients who were excluded from the analysis of body composition changes after 6 months of treatment, the baseline characteristics of included and nonincluded PsA patients were compared using a chi-square test or Fisher's exact test for the categorical variables, and a Student's test or Mann-Whitney U test for continuous variables. Correlations between the quantitative parameters were analyzed using Pearson's coefficient or Spearman's coefficient. Statistical testing was done at the 2-tailed alpha level of 0.05. Data were analyzed using the SAS software package, version 9.4.

RESULTS

Baseline characteristics of patients with PsA and matched controls. The baseline characteristics are shown in Table 1. In the PsA group, age and disease duration were mean ± SD 51.5 ± 11.3 years and 9.6 ± 8.4 years, respectively. Peripheral forms of PsA were found in 90% of the patients (n = 27) and axial forms alone in 3 patients. The DAS28-CRP score was mean ± SD 3.7 ± 1.5 in 27 patients, and the BASDAI score was mean ± SD 61.5 ± 12.2 in 15 patients with axial manifestations. Thirteen patients received a subcutaneously administered 45-mg dose of ustekinumab, while 17 patients received a 90-mg dose. Three patients in the PsA group were currently receiving glucocorticoids (prednisone) at a dose of <10 mg/day (mean ± SD 7.3 ± 2.5 mg/day). Nineteen PsA patients (63.3%) had previously received at least 1 anti-TNF.

Comparison of body composition in PsA patients and matched healthy controls. The baseline body composition parameters of PsA patients and matched controls are shown in Table 2. After adjustment for BMI, body composition parameters in PsA patients exhibited alterations compared to those in healthy controls; in PsA patients, total lean mass and appendicular lean mass were lower (mean ± SD 53.1 ± 13.1 kg versus 56.7 ± 11.9 kg [$P = 0.013$] and 21.6 ± 6.3 kg versus 23.4 ± 5.0 kg [$P = 0.010$], respectively), whereas total fat mass and body fat percentage were higher (mean ± SD 32.5 ± 10.8 kg versus 25.2 ± 8.9 kg [$P < 0.001$] and 36.7% ± 8.6% versus 30.1% ± 8.1% [$P < 0.001$], respectively). Among the PsA patients, 30% had a skeletal muscle mass index below the cutoff for low muscle quantity.

Table 1. Baseline characteristics of the 30 patients with psoriatic arthritis and the 60 healthy controls matched for age (±5 years), sex, body mass index (±3 kg/m^2), and menopausal status for women*

Characteristic	Psoriatic arthritis (n = 30)	Healthy controls (n = 60)
Age, years	51.5 ± 11.3	51.7 ± 11.3
Male, no. (%)	16 (53.3)	32 (53.3)
Body weight, kg	88.3 ± 18.8	81.8 ± 15.3
Body mass index, kg/m^2	29.9 ± 6.1	29.0 ± 5.0
<18.5, no. (%)	0 (<1)	1 (1.6)
18.5 to <25, no. (%)	6 (20)	12 (20)
25 to <30, no. (%)	12 (40)	25 (41.7)
≥30, no. (%)	12 (40)	22 (36.7)
Current smoking, no. (%)	5 (16.7)	7 (11.7)
Excessive alcohol consumption (>3 units/day), no. (%)	2 (6.7)	2 (3.3)
Comorbidities, no. (%)		
Hypertension	12 (40)	0 (<1)
Diabetes mellitus	6 (20)	0 (<1)
Dyslipidemia	6 (20)	0 (<1)
Depression	6 (20)	0 (<1)
COPD/asthma	3 (10)	0 (<1)
Cancer (any)	1 (3.3)	0 (<1)
Disease duration, years	9.6 ± 8.4	–
At least 1 previous biologic	19 ± 63.3	–
Swollen joints (0–28)†	1.9 ± 2.2	–
Tender joints (0–28)†	4.8 ± 5.8	–
CRP (mg/liter)	11.9 ± 13.7	–
DAS28-CRP†	3.7 ± 1.5	–
BASDAI‡	61.5 ± 12.2	–
BASFI‡	44.3 ± 21.3	–
Treatment, no. (%)		
Current synthetic DMARD	3 (10)	–
Current glucocorticoid (<10 mg prednisone/day)	3 (10)	–
Current NSAID	9 (30)	–

* Values are the mean ± SD unless indicated otherwise. BASDI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; COPD = chronic obstructive pulmonary disease; DAS28-CRP = 28-joint Disease Activity Score determined according to C-reactive protein level; DMARD = disease-modifying antirheumatic drug; NSAID = nonsteroidal antiinflammatory drug.
† N = 27 patients.
‡ N = 15 patients.

Body composition changes in PsA patients during treatment with ustekinumab. Twelve participants were excluded at follow-up. Reasons for exclusion were discontinuation of ustekinumab treatment before 6 months had elapsed (n = 6), glucocorticoid treatment >10 mg prednisone/day (n = 2), loss to

Table 2. Baseline body composition of psoriatic arthritis patients and healthy controls*

	Psoriatic arthritis (n = 30)	Controls (n = 60)	Absolute difference†	*P*†	BMI-adjusted *P*†
Fat mass parameters					
Total fat mass, kg	32.5 ± 10.8	25.2 ± 8.9	+7.3	<0.001	<0.001
Body fat percentage, %	36.7 ± 8.6	30.1 ± 8.1	+6.6	<0.001	<0.001
Fat mass index, kg/m^2	11.2 ± 4.1	9.1 ± 3.5	+2.1	<0.001	<0.001
Lean mass parameters					
Total lean mass, kg	53.1 ± 13.1	56.7 ± 11.9	−3.6	0.013	0.002
Appendicular lean mass, kg	21.6 ± 6.3	23.4 ± 5.0	−1.8	0.010	0.002
Skeletal muscle mass index, kg/m^2	7.2 ± 1.6	8.2 ± 1.3	−1.0	<0.001	<0.0001

* Values are the mean ± SD unless indicated otherwise. BMI = body mass index.
† Statistically significant.

follow-up (n = 1), period of inactivity due to joint replacement (n = 1), on a diet during the study period (n = 1), and body composition assessment not performed within time limits (n = 1, beyond 8 months of treatment with ustekinumab). A comparison between the 18 patients who had completed the study at 6 months and the 12 noncompleters revealed no significant differences in demographic, disease, and clinical characteristics.

Changes in body composition are shown in Table 3. After 6 months of treatment with ustekinumab, no significant changes in BMI were observed, although there was a small but nonsignificant difference in total fat mass and body fat percentage. A significant decrease in total lean mass was observed (mean ± SD 54.5 ± 12.5 kg versus 52.7 ± 13.0 kg [*P* = 0.046]), while appendicular lean mass and skeletal muscle mass index remained unchanged.

Significant correlations were found between changes in body composition parameters (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24623). A significant negative correlation was found between total fat mass and total lean mass (r = –0.719, *P* = 0.0008) but not between total fat mass and appendicular lean mass (r = –0.408, *P* = 0.093). A significant positive correlation was also found between change in total fat mass and visceral adipose tissue (r = 0.745, *P* = 0.0004).

In 16 patients, the mean DAS28-CRP remained stable over 6 months, ranging from mean ± SD 3.3 ± 1.4 to 3.2 ± 1.1, whereas in 9 patients with axial manifestations, BASDAI and BASFI scores decreased from mean ± SD 60.7 ± 11.3 to 40.8 ± 26.0 and from 44.8 ± 22.8 to 34.4 ± 36.0, respectively (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24623). Moreover, no significant correlations were found between changes in DAS28-CRP and changes in body composition parameters (total fat mass, appendicular lean mass, total lean mass, and visceral adipose tissue) (see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24623).

Changes in BMD, bone turnover markers, and leptin. No changes in bone turnover markers, leptin, and BMD were observed at 6 months (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24623). A slight decrease in femoral neck BMD was observed, but there was no significant bone loss (defined as a decrease in BMD >0.03 gm/cm^2).

Table 3. Body composition changes in 18 patients with active psoriatic arthritis treated with ustekinumab for 6 months*

	Baseline	6 months	Absolute difference	*P*
Body weight, kg	90.3 ± 16.7	90.5 ± 18.1	↓ 0.2	0.712
BMI, kg/m^2	30.4 ± 6.0	30.5 ± 6.4	↓ 0.1	0.781
Fat mass parameters				
Total fat mass, kg	33.1 ± 10.0	35.0 ± 12.3	↑ 1.9	0.054
Body fat percentage, %	36.6 ± 8.8	38.5 ± 10.4	↑ 1.9	0.067
Fat mass index, kg/m^2	11.7 ± 4.4	12.0 ± 4.8	↑ 0.3	0.061
Visceral adipose tissue, cm^2	170.2 ± 77.8	183.5 ± 89.1	↑ 13.3	0.116
Lean mass parameters				
Total lean mass, kg	54.5 ± 12.5	52.7 ± 13.0	↓ 1.8†	0.047†
Appendicular lean mass, kg	22.3 ± 5.8	21.6 ± 5.9	↓ 0.7	0.196
Skeletal muscle mass index, kg/m^2	7.4 ± 1.4	7.1 ± 1.3	↓ 0.3	0.173

* Values are the mean ± SD unless indicated otherwise. BMI = body mass index.
† Statistically significant.

DISCUSSION

In PsA patients, body composition parameters exhibited alterations compared to those of healthy controls; in PsA patients, total lean mass and appendicular lean mass were lower, whereas total fat mass was higher. Moreover, a third of the patients with PsA requiring ustekinumab had a low skeletal muscle mass index. This study is also the first to investigate changes in body composition in active PsA patients undergoing treatment with ustekinumab. At 6 months of treatment with ustekinumab, total lean mass decreased slightly, whereas total fat mass tended to increase, but not significantly.

The participants included in our study were patients with established PsA, mean disease duration was ~10 years, and only one-third were bDMARD-naive. At baseline, mean DAS28-CRP was quite low (mean swollen joints ~2) and remained stable over 6 months, without improvement with ustekinumab treatment. Treatment with ustekinumab had practically no effect on chronic systemic inflammation as assessed by CRP level measurement (from 10.3 to 10.8 mg/liter). This finding might be partly explained by the fact that most of the patients were overweight/obese, with BMIs that remained stable over the study period. Data suggest that the observed association between CRP level and BMI is probably driven by BMI, with CRP level being a marker of elevated adiposity.

RA is the most widely studied condition in connection with body composition. Book et al demonstrated that changes in body composition occur very early on in RA (1). Appendicular lean mass was found to be lower in RA patients, regardless of sex, and total fat mass and BMI were higher than expected in women with RA. Body composition continues to change over time in RA. Giles et al reported on a cohort of 189 RA patients with a mean disease duration of 9 years (62% female) (2). BMI and fat mass index were significantly higher in women with RA but not in men, compared to controls matched for age, sex, weight, and ethnicity. Total lean mass did not differ significantly between patients with RA and controls, whereas appendicular lean mass was lower in men with RA, but not in women.

To the best of our knowledge, this is one of the first studies to assess DXA-scan body composition in PsA patients compared to healthy controls. Using bioelectrical impedance analysis, Krajewska-Włodarczyk et al reported a significantly lower lean mass and higher fat mass in 51 female patients with PsA compared to 44 controls (8). In a cross-sectional study conducted by Pedreira et al, the authors compared the DXA-scan body composition measurements of 45 women with PsA (mean age 60.5 years), 52 women with psoriasis, and 98 healthy female controls (matched for age, BMI, and ethnicity) (7). Body fat percentage was found to be higher, and total lean mass tended to be lower, in the PsA group than in the healthy controls, which is in keeping with our findings. The researchers also found that 11% of the patients in the PsA group had a skeletal muscle mass index below the cutoff for low muscle quantity (7). Furthermore, Ferguson et al (12) found that individuals with PsA have an adverse magnetic resonance imaging body composition phenotype, with higher levels of visceral fat and lower thigh muscle volumes. On the other hand, Toussirot et al (13) found no difference between patients with PsA and their controls either for lean mass or fat mass using DXA-scan body composition.

Only a few studies have been published on the impact of bDMARDs on body composition in patients with inflammatory rheumatic diseases. In patients with RA, anti-TNFs may lead to an increase in fat mass and no change in lean mass (4). On the other hand, tocilizumab could have an effect on body composition by increasing lean mass with no change in fat mass (5,14).

Until now, no data have been published on body composition changes in patients with PsA treated with bDMARDs. In a study involving patients with psoriasis, Galluzzo et al evaluated changes in body weight, BMI, and body composition assessed by bioelectrical impedance analysis in 53 participants treated with ustekinumab (15). However, no significant changes in BMI, fat mass, and fat-free mass were observed at 12 months of treatment (15). Therefore, we are the first to report results on body composition changes in PsA patients undergoing treatment with bDMARDs. After 6 months of treatment, we found no significant change in BMI and total fat mass, but a slight decline in total lean mass was observed. These preliminary findings need to be confirmed in a larger population.

The mechanisms responsible for changes in body composition in inflammatory rheumatic diseases have yet to be elucidated, but could involve aging, disease activity, chronic systemic inflammation, physical inactivity, inadequate nutrition, and medications (e.g., glucocorticoids and DMARDs). In turn, altered body composition is likely to impact health and causes sarcopenia and cardiometabolic abnormalities. Increased fat mass, especially abdominal fat mass, thus increases the risk of developing type 2 diabetes mellitus and cardiovascular diseases.

Our study has several strengths. First, there were no missing data at the 6-month follow-up, since all of the patients included in our study were followed up in a single center with standardized procedures for all outcomes. Also, all DXA-scan body composition and BMD measurements were performed in the same laboratory by the same investigator. We acknowledge several limitations. First, our population of PsA patients was nonhomogeneous (men and pre- and post-menopausal women). Second, the number of patients is small and the duration of the study was short, which may have been insufficient to detect additional effects of ustekinumab on body composition, BMD, and bone remodeling markers, but despite this limitation, several significant differences were found. Third, we acknowledge that the absence of a control follow-up group precludes any definitive conclusions on body composition changes under ustekinumab. Fourth, we used BASDAI and DAS28-CRP criteria to measure the disease activity in patients with PsA. However, these tools are not used

specifically for PsA, and minimal disease activity and Disease Activity Index for Psoriatic Arthritis or Composite Psoriatic Disease Activity Index criteria should be used for further studies. Last, muscle function and performance were not assessed in our study.

In conclusion, our study has shown that patients with PsA have higher adiposity and lower lean mass than healthy controls. We observed no substantial change in body composition in PsA patients undergoing treatment with ustekinumab, but we did observe a slight decline in total lean mass.

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. Paccou 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. Paccou, Cortet, Flipo.

Acquisition of data. Paccou, Bavière.

Analysis and interpretation of data. Paccou, Sornay-Rendu, Szulc, Ramdane, Cortet, Chapurlat, Flipo.

REFERENCES

1. Book C, Karlsson MK, Akesson K, Jacobsson LT. Early rheumatoid arthritis and body composition. Rheumatology (Oxford) 2009;48: 1128–32.
2. Giles JT, Bartlett SJ, Andersen RE, Fontaine KR, Bathon JM. Association of body composition with disability in rheumatoid arthritis: impact of appendicular fat and lean tissue mass. Arthritis Rheum 2008;59:1407–15.
3. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, et al. Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on sarcopenia in older people. Age Ageing 2010;39:412–23.
4. Marouen S, Barnetche T, Combe B, Morel J, Daïen CI. TNF inhibitors increase fat mass in inflammatory rheumatic disease: a systematic review with meta-analysis. Clin Exp Rheumatol 2017;35:337–43.
5. Tournadre A, Pereira B, Dutheil F, Giraud C, Courteix D, Sapin V, et al. Changes in body composition and metabolic profile during interleukin 6 inhibition in rheumatoid arthritis. J Cachexia Sarcopenia Muscle 2017;8:639–46.
6. Bavière W, Deprez X, Houvenagel E, Philippe P, Deken V, Flipo RM, et al. Association between comorbidities and quality of life in psoriatic arthritis: results from a multicentric cross-sectional study (PSAQUAL study). J Rheumatol 2020;47:369–76.
7. Pedreira PG, Pinheiro MM, Szejnfeld VL. Bone mineral density and body composition in postmenopausal women with psoriasis and psoriatic arthritis. Arthritis Res Ther 2011;13:R16.
8. Krajewska-Włodarczyk M, Owczarczyk-Saczonek A, Placek W. Changes in body composition and bone mineral density in postmenopausal women with psoriatic arthritis. Reumatologia 2017; 55:215–21.
9. Paccou J, Wendling D. Current treatment of psoriatic arthritis: update based on a systematic literature review to establish French Society for Rheumatology (SFR) recommendations for managing spondyloarthritis. Joint Bone Spine 2015;82:80–5.
10. Lubberts E. The IL-23-IL-17 axis in inflammatory arthritis. Nat Rev Rheumatol 2015;11:415–29.
11. Beringer A, Miossec P. Systemic effects of IL-17 in inflammatory arthritis. Nat Rev Rheumatol 2019;15:491–501.
12. Ferguson LD, Linge J, Dahlqvist Leinhard O, Woodward R, Barrientos PH, Roditi, et al. Psoriatic arthritis is associated with adverse body composition predictive of greater coronary heart disease and type 2 diabetes propensity: a cross-sectional study. Rheumatology (Oxford) 2021;60:1858–62.
13. Toussirot E, Aubin F, Desmarets M, Wendling D, Augé B, Gillard J, et al. Visceral adiposity in patients with psoriatic arthritis and psoriasis alone and its relationship with metabolic and cardiovascular risk. Rheumatology (Oxford) 2021;60:2816–25.
14. Toussirot E, Marotte H, Mulleman D, Cormier G, Coury F, Gaudin P, et al. Increased high molecular weight adiponectin and lean mass during tocilizumab treatment in patients with rheumatoid arthritis: a 12-month multicentre study. Arthritis Res Ther 2020;22:224.
15. Galluzzo M, D'Adamio S, Pastorino R, Andreoli A, Servoli S, Bianchi L, et al. Effect of anti-IL-12/23 on body composition: results of bioelectrical impedance analysis in Caucasian psoriatic patients. Expert Opin Biol Ther 2018;18:229–35.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1614–1622
DOI 10.1002/acr.24616

Mortality in Ankylosing Spondylitis According to Treatment: A Nationwide Retrospective Cohort Study of 5,900 Patients From Israel

Niv Ben-Shabat,[1] Aviv Shabat,[2] Abdulla Watad,[3] Khalaf Kridin,[4] Nicola Luigi Bragazzi,[5] Dennis McGonagle,[6] Doron Comaneshter,[7] Arnon D. Cohen,[8] and Howard Amital[9]

Objective. In this large population-based study we aimed: 1) to assess mortality in patients with ankylosing spondylitis (AS) compared to the general population, considering demographics, comorbidities, and treatment, and 2) to assess factors associated with mortality within patients with AS.

Methods. This study was designed as a retrospective cohort study using the electronic database of the largest health maintenance organization in Israel. All patients with AS diagnosed between 2002 and 2018 were included. Controls were matched by age, sex, clinic, and enrollment time. Follow-up continued until death or the end of the study.

Results. The study comprised 5,930 AS patients and 29,018 matched controls who were followed up for a median period of 7.5 years. There were 667 deaths within the AS cohort and 2,919 deaths within controls; the mean age at death was 76.9 years and 77.1 years, respectively ($P = 0.74$). A total of 3,249 AS patients (54.8%) were treated only with nonsteroidal antiinflammatory drugs, 1,760 (29.7%) were treated with tumor necrosis factor inhibitors (TNFi), and 1,687 (28.4%) with disease-modifying antirheumatic drugs (DMARDs). Mortality rates were increased among AS patients compared to controls, with an age- and sex-adjusted hazard ratio (HR) of 1.19 (95% confidence interval [95% CI] 1.10–1.30). The association was significant for men (HR 1.15 [95% CI 1.04–1.27]) and women (HR 1.32 [95% CI 1.13–1.54]), and after adjusting for background comorbidities (HR 1.14 [95% CI 1.05–1.24]). AS patients treated with TNFi or with a combination of TNFi and DMARDs did not have significant difference in mortality rates compared to controls (HR 0.67 [95% CI 0.38–1.18] and HR 0.93 [95% CI 0.69–1.25], respectively). Age, male sex, mean C-reactive protein (CRP) levels and general comorbidities were predictors of mortality within the AS cohort.

Conclusion. AS patients had an increased mortality risk compared to the general population after adjusting for age, sex, and baseline comorbidities. AS patients treated with TNFi did not demonstrate excess mortality compared to matched controls. Within the AS cohort, age, male sex, background comorbidities, and higher CRP levels were identified as risk factors for mortality.

INTRODUCTION

Ankylosing spondylitis (AS) is a chronic inflammatory disease affecting the axial and peripheral joints, with extraarticular manifestations such as psoriasis, inflammatory bowel disease (IBD), uveitis, and cardiac conduction abnormalities (1). The estimated prevalence of AS in the general population ranges from 9 to 30 per 10,000 persons (2), which makes it a significant

[1]Niv Ben-Shabat, MPH: Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv-Yafo, and Sheba Medical Center, Ramat Gan, Israel; [2]Aviv Shabat, MD: Hadassah Faculty of Medicine, The Hebrew University, Jerusalem, Israel; [3]Abdulla Watad, MD: Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv-Yafo, and Sheba Medical Center, Ramat Gan, Israel, and University of Leeds and Chapel Allerton Hospital, Leeds, UK; [4]Khalaf Kridin, MD, PhD: University of Lübeck, Lübeck, Germany, and Azrieli Faculty of Medicine, Bar-Ilan University, Safed, Israel; [5]Nicola Luigi Bragazzi, MD, PhD: Postgraduate School of Public Health, University of Genoa, Genoa, Italy, and York University, Toronto, Ontario, Canada; [6]Dennis McGonagle, MD, PhD: University of Leeds and Chapel Allerton Hospital, Leeds, UK; [7]Doron Comaneshter, PhD: Clalit Health Services Tel Aviv, Tel-Aviv, Israel; [8]Arnon D. Cohen, MD, PhD: Clalit Health Services Tel Aviv, Tel-Aviv, and Ben Gurion University of the Negev, Beer Sheva, Israel; [9]Howard Amital, MD, MHA: Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv-Yafo, Sheba Medical Center, Ramat Gan, and Ben Gurion University of the Negev, Beer Sheva, Israel.

Drs. Ben-Shabat and Shabat contributed equally to this work.

Dr. Cohen has received consulting fees, speaking fees, and/or honoraria from AbbVie, Amgen, Boehringer Ingelheim, Dexcel Pharma, Janssen, Kamedis, Eli Lilly, Neopharm, Novartis, Perrigo, Pfizer, Rafa, Samsung Bioepis, Sanofi, Sirbal, and Taro (less than $10,000 each) and research grants from Janssen, Novartis, AbbVie, Janssen, and Sanofi. Dr. Amital has received consulting fees, speaking fees, and/or honoraria from AbbVie, Boehringer Ingelheim, Bayer, Janssen, Eli Lilly, Neopharm, Novartis, Perrigo, Pfizer, Roche, AstraZeneca, Psyche, Sanofi, Teva, and Taro (less than $10,000 each) and research grants from Janssen, AbbVie, Pfizer, and Gilead. No other disclosures relevant to this article were reported.

Address correspondence to Howard Amital, MD, MHA, Head, Department of Medicine B and The Zabludowicz Center for Autoimmune Diseases, Sheba Medical Center Tel-Hashomer, 52621, Israel. Email: howard.amital@sheba.health.gov.il.

Submitted for publication August 19, 2020; accepted in revised form April 6, 2021.

SIGNIFICANCE & INNOVATIONS

- When adjusting to baseline comorbidities, ankylosing spondylitis patients are at increased risk of mortality compared to the general population in Israel.
- In ankylosing spondylitis patients ever treated with tumor necrosis factor inhibitors, no excess mortality was observed compared to matched controls.
- Age, male sex, higher mean C-reactive protein levels, and general comorbidities were predictors of mortality within the ankylosing spondylitis cohort.

burden on the individual person as well as on the health care system (3). Commonly employed therapeutic agents are nonsteroidal antiinflammatory drugs (NSAIDs), disease-modifying antirheumatic drugs (DMARDs), such as methotrexate and sulfasalazine, and biologic therapy such as tumor necrosis factor inhibitors (TNFi) and secukinumab, the emergent interleukin-17 inhibitor (4).

The TNFi agents were introduced 2 decades ago and besides having a role in decreasing disease activity, they have also shown protective effects in atherosclerosis (5,6). Some evidence suggests long-term survival benefits from TNFi treatment in other chronic inflammatory rheumatic diseases, such as rheumatoid arthritis (RA) (7,8) and psoriatic arthritis (PsA) (9). In AS, 1 study has demonstrated survival benefits of TNFi treatment compared to the general population (10), yet this study was based on data from a randomized controlled trial (RCT) population and lacked a comparison cohort.

There are relatively few studies in the post–spinal radiotherapy era addressing all-cause mortality in AS, with all of them showing excess mortality compared to the general population (11–17). Cardiovascular disease was the leading cause of death, followed by malignancy and infectious diseases (12,14–17). When comparing to the general population, higher proportions of death due to infectious, respiratory, and renal diseases, and a lower proportion of death due to malignancy, were observed (12,14,17). Most of these studies had significant limitations, including small sample-size, cross-sectional design, self-reported cases, and lack of a comparison cohort. In addition, there has been no study to date that adjusted for significant comorbidities, which increase mortality on one hand and are associated with AS on the other hand (10). Therefore, this large population-based study aimed to investigate mortality patterns in AS patients compared to the general population, assessing the role of TNFi treatment and significant comorbidities for the first time.

MATERIALS AND METHODS

Data source. Data were obtained from the Clalit Healthcare Services (CHS) electronic database. CHS is the largest health maintenance organization in Israel and serves ~4.5 million insured members (>50% of Israel's population) from heterogeneous ethnic groups and has continuous input from pharmaceutical, medical, and administrative operating systems. The database is used for administrative and clinical management and is available for use in epidemiologic studies. Patients' data can be automatically extracted from the database using data-mining techniques. The CHS database was validated as having a 90–100% degree of accuracy (18) and has previously been used in many studies (19–23), including one conducted with data of AS patients (24).

Sample and design. This study was approved by the CHS Ethics Committee in Beer-Sheva, Israel (approval #0212-17-COM). No informed consent was needed, as this was an existing database. The study was designed as a retrospective cohort study. Using the CHS's computerized database, we extracted a cohort consisting of AS patients first diagnosed between January 1, 2002 to December 31, 2018 and compared them with age-, sex- and clinic-matched controls. For AS patients, follow-up began at the date of the first recorded AS diagnosis and for controls on the date of their matched patient. Follow-up continued until death or the end of the study on July 1, 2019.

Study variables. AS patients were defined as such if they had at least 1 documented diagnosis of AS in their medical records as an outpatient, either by a primary care physician or a specialist, or if they were diagnosed with AS in their hospital discharge papers. Patients age <18 years at the time of diagnosis were excluded. Controls were randomly assigned from the CHS database, with the exclusion of AS patients. Approximately 5 controls were matched by age, sex, primary-care clinic, and enrollment time for each AS patient. Data available from the CHS database included variables such as age, sex, socioeconomic status (SES), body-mass index (BMI), chronic diseases, laboratory test results, and death date. The SES was defined according to the poverty index of the member's residence area as defined in the 2008 National Census. Specifically, the poverty index was computed based on household income, education, marital conditions, and car ownership, among others.

We divided the population into 3 categories based on quartiles (low: 25th percentile; medium: 25th–75th percentile; and high: 75th–100th percentile) (23). The BMI was calculated using height and weight measurements from the year of enrollment to the study (if available). Diagnoses of comorbidities were also obtained from the CHS electronic database and were considered only if first diagnosed before the enrollment date. The comorbidity load was assessed using the Charlson Comorbidity Index, which was validated in predicting mortality in longitudinal studies (25). Patients were defined as being ever treated with TNFi if they had been prescribed and dispensed any of the following agents: infliximab, etanercept, adalimumab, golimumab, and certolizumab; and were defined as being ever treated with DMARDs if they ever had dispensed methotrexate or sulfasalazine. NSAIDs treatment was defined as any prescribed dispensing of any of the NSAID

agents (including COX-2 inhibitors). The mean level of serum C-reactive protein (CRP) was calculated from all laboratory tests results that were done in the hospital or in the community during the follow-up period.

Statistical analysis. Differences in baseline characteristics between different groups of independent variables were compared using *t*-tests or Mann-Whitney U tests for continuous variables, and Pearson's chi-square tests for categorical variables. Survival analysis was done using multivariate the Cox proportional hazards method. The outcome was all-cause mortality, and the independent variable was AS diagnosis. Two models with different sets of adjustments were used: adjusting for age and sex only and adjusting for the Charlson Comorbidity Index as well. The models were repeated separately according to treatment and sex. Survival curves were obtained using the Kaplan-Meier method with a post hoc log-rank comparison. Predictors for mortality among AS patients were obtained using a multivariate Cox proportional hazards method. Each model accounted for age and sex and was repeated separately for each sex. Factors associated with mortality in AS patients were evaluated using a univariate and a multivariate binary logistic regression analysis and reported as odds ratios (ORs) and 95% confidence intervals (95% CIs). The multivariate analysis included variables that were found significant in the univariate analysis. Statistical analysis was performed using the commercial software Statistical Package for the Social Sciences, version 23.0.

RESULTS

Cohort characteristics and AS features. The study comprised 5,930 AS patients and 29,018 matched controls who were followed up for a median period of 7.5 years (interquartile range 3.5–11.6 years). The mean ± SD age at enrollment was 49.8 ± 16.6 years, and the female proportion was 36.6% and was the same for AS patients and controls (not shown). During follow-up, 667 AS patients (11.2%) and 2,919 controls (10.1%) had died ($P < 0.01$). The mean ± SD age at death was 76.9 ± 12.0 years for AS and 77.1 ± 12.3 years for controls ($P = 0.74$). AS patients had higher baseline rates of connective tissue diseases (4.6% versus 0.6%; $P < 0.01$), diabetes mellitus (14.9% versus 12.4%; $P < 0.01$), ischemic heart disease (11.2% versus 9.2%; $P < 0.01$), peptic ulcer disease (8.0% versus 5.3%; $P < 0.01$), and solid malignancies (7.3% versus 5.5%; $P < 0.01$) than controls (Table 1). Among AS patients, 348 (6%) had concurrent IBD, and 299 (5.1%) had psoriasis. During follow-up, 707 AS patients (12.1%) were

Table 1. Baseline characteristics: a comparison between ankylosing spondylitis patients and controls*

Characteristic	AS patients (n = 5,930)	Controls (n = 29,018)	*P*
Age at enrollment, mean ± SD years	49.4 ± 17.2	49.3 ± 17.2	0.63
Female	2,156 (36.4)	10,576 (36.4)	0.90
Follow-up time, median (IQR) years	7.4 (3.5–11.5)	7.5 (3.5–11.7)	0.42
Death during follow-up	667 (11.2)	2,919 (10.1)	<0.01
Age at death, mean ± SD years	76.9 ± 12	77.1 ± 12.3	0.74
Body mass index, mean ± SD kg/m^2†	27.5 ± 5.9	28.5 ± 5.9	0.75
Arab ethnicity	999 (17.2)	4,914 (17.3)	0.74
Recruitment periods			
2002–2007	1,754 (30.1)	8,576 (30.2)	0.85
2008–2013	2,103 (36.1)	10,212 (36)	0.90
2014–2018	1,968 (33.8)	9,568 (33.7)	0.95
Socioeconomic status‡			0.16
Low	816 (14.7)	4,246 (15.6)	–
Medium	3,893 (70.1)	18,951 (69.7)	–
High	847 (15.2)	3,990 (14.7)	–
Baseline comorbidities			
Cerebrovascular disease	203 (3.5)	901 (3.2)	0.23
Chronic kidney disease	158 (2.7)	710 (2.5)	0.36
Chronic pulmonary disease	207 (3.6)	748 (2.6)	0.37
Cirrhosis	15 (0.3)	53 (0.2)	0.27
Congestive heart failure	100 (1.7)	478 (1.7)	0.87
Connective tissue disease	270 (4.6)	181 (0.6)	<0.01
Diabetes mellitus	870 (14.9)	3,514 (12.4)	<0.01
Ischemic heart disease	651 (11.2)	2,614 (9.2)	<0.01
Peptic ulcer disease	466 (8.0)	1,498 (5.3)	<0.01
Peripheral vascular disease	129 (2.2)	550 (1.9)	0.17
Solid malignancy	424 (7.3)	1,562 (5.5)	<0.01
Hematologic malignancy	61 (1.0)	239 (0.8)	0.13

* Values are the number (%) unless indicated otherwise. AS = ankylosing spondylitis; IQR = interquartile range.
† Available for 67.1% of data.
‡ Available for 93.7% of data.

Table 2. Clinical and treatment-related characteristics of ankylosing spondylitis patients*

AS-related manifestations	Value
Inflammatory bowel disease	348 (5.9)
Osteoporosis	707 (11.9)
Psoriasis	299 (5.1)
Joint replacement	292 (4.9)
Treatment	
NSAIDs	
Any	5,648 (95.2)
Alone	3,249 (54.8)
TNF inhibitors	
Any	1,760 (29.7)
Without DMARDs	737 (12.7)
With methotrexate alone†	253 (4.3)
With sulfasalazine alone†	359 (6.1)
With methotrexate and sulfasalazine	411 (6.9)
DMARDs	
Any	1,687 (28.4)
Without TNF inhibitors	664 (11.2)
Methotrexate alone†	153 (2.6)
Sulfasalazine alone†	372 (6.3)

* Values are the number (%). AS = ankylosing spondylitis; DMARDs = disease-modifying antirheumatic drugs; NSAIDs = nonsteroidal antiinflammatory drugs; TNF = tumor necrosis factor.
† With or without NSAIDs.

diagnosed with osteoporosis, and 292 (4.9%) had joint replacement surgery. Regarding treatment, 5,648 AS patients (95.2%) were ever treated with NSAIDs, and for 3,249 (54.8%), NSAIDs were the only therapeutic agent prescribed. A total of 1,687 AS patients (28.4%) were ever treated with DMARDs (methotrexate or sulfasalazine), 664 (11.2%) without ever being prescribed TNFi, and 1,023 (17.6%) who switched or were treated with concurrent TNFi. In all, 1,760 AS patients (29.7%) were ever treated with TNFi, 737 (12.7%) without ever being prescribed DMARDs (Table 2). Features of the AS cohort according to treatment are shown in Table 3. A significant difference in enrollment age was noticeable when AS patients being treated with TNFi were relatively younger (TNFi alone [41.4 ± 13.1 years], for TNFi and DMARDs [44.6 ± 13.4 years], DMARDs alone [50.4 ± 14.8 years], and NSAIDs alone [53.8 ± 17.4 years]). Accordingly, these patients had a relatively lower age at death (TNFi alone [66.2 ± 12.8 years], TNFi and DMARDs [68.1 ± 10.8 years], DMARDs alone [71.1 ± 12.3 years], and NSAIDs alone [78.9 ± 11.1 years]).

Mortality in AS compared to the general population. All-cause mortality was increased in AS patients compared to matched controls (hazard ratio [HR] 1.19 [95% CI 1.10–1.30],

Table 3. A comparison of baseline characteristics of ankylosing spondylitis patients according to treatment*

Characteristic	NSAIDs alone	TNF inhibitors alone	DMARDs alone	TNF inhibitors + DMARDs	*P*
Age at enrollment, mean ± SD years	53.8 ± 17.4	41.4 ± 13.1	50.4 ± 14.8	44.6 ± 13.4	<0.001
Age at death, mean ± SD years	78.9 ± 11.1	66.2 ± 12.8	71.1 ± 12.3	68.1 ± 10.8	<0.001
Follow-up time, median (IQR) years	7.0 (3–11)	5.3 (3–8)	9.1 (5.2–13)	9.6 (5.3–13.6)	<0.001
Duration of treatment, mean ± SD years	–	3 ± 3.2	5.7 ± 6.8	–	–
Body mass index, mean ± SD kg/m^2†	27.8 ± 5.9	26.5 ± 5.3	28.2 ± 5.7	27.2 ± 5.9	0.986
Serum CRP, mean ± SD mg/liter	1.12 ± 2.3	1.75 ± 2.1	1.41 ± 1.6	2.02 ± 2.4	<0.001
Women	1,139 (35.1)	253 (34.3)	274 (41.3)	416 (40.7)	<0.001
Arab ethnicity	525 (16.2)	134 (18.4)	129 (19.4)	200 (19.6)	0.092
Recruitment periods					
2002–2007	955 (29.4)	89 (12.1)	261 (39.3)	420 (41.1)	<0.001
2008–2013	1,185 (36.5)	271 (36.8)	258 (38.9)	350 (34.2)	0.036
2014–2018	1,109 (34.1)	377 (51.2)	145 (21.8)	253 (24.7)	<0.001
Socioeconomic status‡					0.019
Low	404 (13.3)	110 (16)	103 (16.7)	145 (14.9)	–
Intermediate	2,182 (71.8)	481 (69.8)	436 (70.6)	648 (61.8)	–
High	451 (14.9)	98 (14.2)	79 (12.8)	179 (18.4)	–
Baseline comorbidities					
Cerebrovascular disease	148 (4.6)	12 (1.6)	24 (3.6)	17 (1.7)	<0.001
Chronic kidney disease	119 (3.7)	9 (1.2)	18 (2.7)	8 (0.8)	<0.001
Chronic pulmonary disease	141 (4.3)	15 (2.0)	32 (4.8)	15 (1.5)	<0.001
Cirrhosis	9 (0.3)	1 (0.1)	1 (0.2)	4 (0.4)	0.515
Congestive heart failure	80 (2.5)	3 (0.4)	10 (1.5)	3 (0.3)	<0.001
Connective tissue disease	54 (1.7)	12 (1.6)	69 (10.4)	133 (13.0)	<0.001
Diabetes mellitus	602 (18.5)	50 (6.8)	103 (15.5)	105 (10.3)	<0.001
Ischemic heart disease	473 (14.6)	27 (3.7)	76 (11.4)	61 (6.0)	<0.001
Peptic ulcer disease	280 (8.6)	26 (3.5)	75 (11.3)	78 (7.6)	<0.001
Peripheral vascular disease	94 (2.9)	7 (0.9)	15 (2.3)	11 (1.1)	0.023
Solid malignancy	273 (8.4)	17 (2.3)	55 (8.3)	75 (7.3)	<0.001
Hematologic malignancy	40 (1.2)	3 (0.4)	5 (0.8)	13 (1.3)	0.241
Charlson Comorbidity Index, mean ± SD	1.5 ± 1.2	1.2 ± 0.8	1.7 ± 1.3	1.5 ± 1.0	<0.001

* Values are the number (%) unless indicated otherwise. CRP = C-reactive protein; DMARDs = disease-modifying antirheumatic drugs; IQR = interquartile range; NSAIDs = nonsteroidal antiinflammatory drugs; TNF = tumor necrosis factor.
† Available for 67.1% of data.
‡ Available for 93.7% of data.

$P < 0.001$), for both men (HR 1.15 [95% CI 1.04–1.27], $P < 0.01$) and women (HR 1.32 [95% CI 1.13–1.54], $P < 0.001$) separately (Table 4). This trend is depicted graphically in the Kaplan-Meier survival curves as well (Figure 1). The association remained significant when adjusting for baseline comorbidities as well (overall HR 1.15 [95% CI 1.04–1.27], $P < 0.01$; men HR 1.15 [95% CI 1.04–1.27], $P < 0.01$; women HR 1.15 [95% CI 1.04–1.27], $P < 0.01$). A similar trend was demonstrated when comparing AS patients treated with NSAIDs alone to matched controls (overall HR 1.16 [95% CI 1.05–1.28], $P < 0.01$; men HR 1.13 [95% CI 1.02–1.27], $P < 0.05$; women HR 1.22 [95% CI 1.02–1.46], $P < 0.05$) and when comparing AS patients treated with DMARDs alone (with or without concurrent or prior NSAIDs treatment but without ever receiving TNFi) to matched controls (overall HR 1.69 [95% CI 1.36–2.09], $P < 0.001$; men HR 1.54 [95% CI 1.18–2.01], $P < 0.01$; women HR 2.12 [95% CI 1.48–3.05], $P < 0.001$). These results did not change considerably when adjusting for baseline comorbidities as well. When looking at patients treated with TNFi alone (with or without concurrent or prior NSAIDs treatment but without ever receiving DMARDs), no excess mortality was observed when adjusting for age and sex (HR 0.67 [95% CI 0.38–1.18]) and when adjusting for comorbidities as well (HR 0.66 [95% CI 0.37–1.17]). Similar results were demonstrated in AS patients treated with TNFi and DMARDs (age- and sex-adjusted model HR 0.93 [95% CI 0.69–1.25]; age-, sex-, and comorbidities-adjusted model HR 0.85 [95% CI 0.63–1.15]).

Factors associated with mortality within AS cohort. Within the AS cohort, age (10-year increments OR 3.02 [95% CI 2.79–3.27], $P < 0.001$) and male sex (OR 1.56 [95% CI 1.27–1.90], $P < 0.001$) were significantly associated with mortality, while ethnicity (Arab versus Jewish OR 1.15 [95% CI 0.84–1.57], $P = 0.393$) and SES (low versus intermediate/high OR 1.36 [95% CI 0.99–1.85], $P = 0.058$) were not found to significantly influence mortality rates (Table 5). All general comorbidities according to the Charlson Comorbidity Index were found to be significantly associated with mortality, except cirrhosis (OR 1.75 [95% CI 0.44–6.90], $P = 0.425$) and connective-tissue diseases (OR 1.33 [95% CI 0.88–2.00], $P = 0.175$). Regarding cirrhosis, the low prevalence (0.3%) might explain the lack of statistical significance. The score of the Charlson Comorbidity Index was itself significantly associated with mortality as well, with an OR of 1.52 (95% CI 1.43–1.62; $P < 0.001$) for every 1-point increment. When addressing AS-related manifestations, IBD (OR 1.47 [95% CI 0.96–2.27], $P = 0.08$), psoriasis (OR 1.30 [95% CI 0.85–1.99], $P = 0.233$), osteoporosis (OR 1.24 [95% CI 0.99–1.55], $P = 0.058$), and a history of joint replacement surgery (OR 1.14 [95% CI 0.81–1.61], $P = 0.438$) were not found to be associated with mortality in AS, while higher levels of serum CRP, which is a marker for disease activity, were found to significantly increase mortality (OR 1.86 [95% CI 1.42–2.44], $P < 0.001$). The time from AS diagnosis to the administration of the first TNFi agent had also a minor, yet significant, association with mortality (OR 1.01 [95% CI 1.00–1.02], $P = 0.002$).

Table 4. Cox-regression hazards ratios comparing all-cause mortality in ankylosing spondylitis patients to controls, stratified according to treatment*

	Adjusted for age and sex		Adjusted for age, sex, and comorbidities†	
	HR (95% CI)	*P*	HR (95% CI)	*P*
All AS patients				
Overall	1.19 (1.10–1.30)	<0.001	1.14 (1.05–1.24)	<0.01
Men	1.15 (1.04–1.27)	<0.01	1.10 (1.00–1.22)	0.05
Women	1.32 (1.13–1.54)	<0.001	1.22 (1.05–1.43)	<0.01
Treated with NSAIDs alone				
Overall	1.16 (1.05–1.28)	<0.01	1.12 (1.01–1.23)	<0.05
Men	1.13 (1.02–1.27)	<0.05	1.10 (0.98–1.23)	NS
Women	1.22 (1.02–1.46)	<0.05	1.14 (0.96–1.37)	NS
Treated with TNF inhibitors alone				
Overall	0.67 (0.38–1.18)	NS	0.66 (0.37–1.17)	NS
Men	0.81 (0.46–1.43)	NS	0.80 (0.45–1.42)	NS
Women‡	–	–	–	–
Treated with DMARDs alone§				
Overall	1.69 (1.36–2.09)	<0.001	1.47 (1.19–1.83)	<0.001
Men	1.54 (1.18–2.01)	<0.01	1.32 (1.01–1.73)	<0.05
Women	2.12 (1.48–3.05)	<0.001	1.95 (1.36–2.81)	<0.001
Treated with TNF inhibitors + DMARDs				
Overall	0.93 (0.69–1.25)	NS	0.85 (0.63–1.15)	NS
Men	0.76 (0.52–1.13)	NS	0.70 (0.48–1.04)	NS
Women	1.39 (0.87–2.23)	NS	1.25 (0.78–2.00)	NS

* 95% CI = 95% confidence interval; AS = ankylosing spondylitis; DMARDs = disease-modifying antirheumatic drugs; HR = hazard ratio; NS = nonsignificant; NSAIDs = nonsteroidal antiinflammatory drugs; TNF = tumor necrosis factor.
† According to Charlson Comorbidity Index.
‡ No deaths in subgroup.
§ Without TNF inhibitors and with and without NSAIDs.

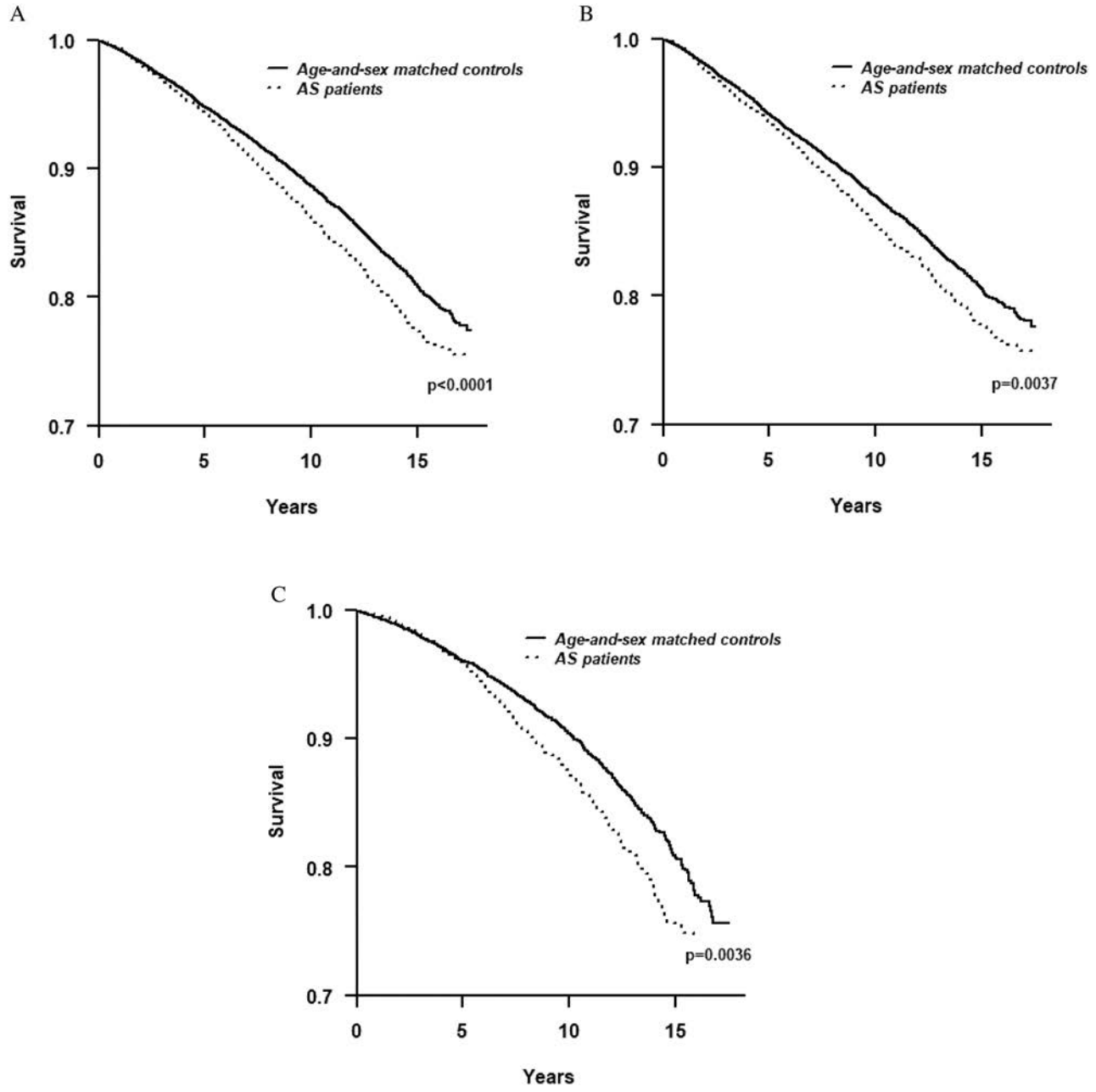

Figure 1. Kaplan-Meier survival curves comparing ankylosing spondylitis (AS) patients and age-and-sex matched controls: **A**, total, **B**, men, and **C**, women.

DISCUSSION

In this large population-based study, we found significantly increased all-cause mortality in AS patients compared with matched controls representing the general population in Israel. The association remained significant after adjusting for general comorbidities, which are associated with AS and with increased mortality. When examining AS patients ever treated with TNFi, alone or with combination with DMARDs, no excess mortality was observed. Among these patients, longer time for initiation of treatment was slightly yet significantly associated with increased mortality. Within the AS cohort, older age, male sex, general comorbidities, and increased CRP levels increased the risk for mortality, while Arab ethnicity, low-SES, and AS-related manifestations (IBD, psoriasis, history of joint replacement surgery, and osteoporosis) were not found significant.

Our results demonstrating increased mortality in AS patients were compatible with previous studies that showed mortality risks ranging between 1.3 to 1.8 (11–17). Only 2 of these studies were with relatively large sample sizes and used matched comparison cohorts (11,14). None of the prior studies adjusted for baseline

Table 5. Factors associated with mortality within ankylosing spondylitis cohort, logistic regression analysis*

	Rates in deceased AS patients (n = 667)	OR (95% CI)†	*P*
Demographics			
Age, mean ± SD 10-year increments	70.5 ± 12	3.02 (2.79–3.27)	<0.001
Male (vs. female)	468 (70.2)	1.56 (1.27–1.90)	<0.001
Ethnicity (Arab vs. Jewish)	60 (9.0)	1.15 (0.84–1.57)	0.393
Socioeconomic status (low vs. intermediate/high)	65 (9.7)	1.36 (0.99–1.85)	0.058
General comorbidities			
Cerebrovascular disease	78 (11.7)	1.71 (1.22–2.40)	0.002
Chronic kidney disease	78 (11.7)	2.45 (1.67–3.59)	<0.001
Chronic pulmonary disease	83 (12.4)	2.44 (1.74–3.42)	<0.001
Cirrhosis	4 (0.6)	1.75 (0.44–6.90)	0.425
Congestive heart failure	54 (8.1)	2.78 (1.75–4.41)	<0.001
Connective tissue disease	37 (5.5)	1.33 (0.88–2.00)	0.175
Diabetes mellitus	220 (33.0)	1.54 (1.25–1.89)	<0.001
Ischemic heart disease	238 (35.7)	1.72 (1.38–2.14)	<0.001
Peptic ulcer disease	121 (18.1)	1.38 (1.06–1.80)	0.015
Peripheral vascular disease	58 (8.7)	2.24 (1.49–3.35)	<0.001
Solid malignancy	162 (24.3)	4.34 (3.36–5.62)	<0.001
Hematologic malignancy	27 (4.0)	3.45 (1.87–6.38)	<0.001
Charlson Comorbidity Index (every 1-point increment)	–	1.52 (1.43–1.62)	<0.001
AS-related manifestations			
Inflammatory bowel disease	30 (4.5)	1.47 (0.96–2.27)	0.08
Psoriasis	31 (4.6)	1.30 (0.85–1.99)	0.233
Osteoporosis	181 (27.1)	1.24 (0.99–1.55)	0.058
Joint replacement	62 (9.3)	1.14 (0.81–1.61)	0.438
Serum CRP, mean ± SD 5-mg/liter increment	1.80 ± 2.4	1.86 (1.42–2.44)	<0.001
Time from enrollment to initiation of TNFi, median (IQR) 1-month increment	20.5 (6–32)	1.01 (1.00–1.02)	0.002

* Values are the number (%) unless indicated otherwise. 95% CI = 95% confidence interval; AS = ankylosing spondylitis; CRP = C-reactive protein; IQR = interquartile range; OR = odds ratio.
† Age-adjusted.

general comorbidities, which are found in higher rates in AS patients and are related to mortality, and therefore can act as confounders (26–28). In our study, we used the Charlson Comorbidity Index to quantify the comorbidity burden. This well-established score was validated as a prognostic factor in longitudinal studies (25).

Beneficial effects of TNFi on all-cause mortality were previously reported for other chronic inflammatory rheumatic diseases such as RA (7,8,10) and PsA (9). Regarding AS, many RCTs have demonstrated the efficacy of TNFi, in terms of disease activity (29), and its safety in terms of adverse events such as malignancies and infections (8). Yet regarding the effect on long-term survival, to our knowledge, only 1 study has previously been published (10). The study reported decreased all-cause mortality among AS patients treated with TNFi (standardized mortality ratio [SMR] 0.14 [95% CI 0–0.78]). However, the study referred only to adalimumab, had a small sample size with no deaths among female patients, and was based on a selected population from RCTs, which is not representative of a real-life population. Our study is not randomized, and each treatment group represents patients with different clinical features and cannot be directly compared. Therefore, we compared each group to matched controls from the general population and performed an adjustment for relevant factors. This cohort has the advantage of describing a real-world state yet is not indicative of efficacy or causality.

In our study, in accordance with previous reports, baseline comorbidity rates, such as chronic pulmonary disease, connective-tissue diseases, diabetes mellitus, ischemic heart disease, peptic ulcer disease, and malignancies, were higher in AS patients compared to controls (3,11,14). This observation might be explained by a shared pathogenic mechanism, reverse causality, or detection bias due to more frequent use of health care services by the AS cohort. Almost all general comorbidities, as expected, were found to be risk factors for mortality within the AS cohort. The rates of extraarticular manifestations (IBD and psoriasis) as well as disease-related conditions (osteoporosis and joint replacement surgery) were roughly the same as in previous studies and likewise did not significantly influence mortality (14,16). Another finding of our study was that higher levels of serum CRP, a marker for AS disease activity, were associated with increased mortality. This finding, which supports 2 previous reports (12,15), might indicate a positive association between the severity of AS and mortality. Notably, however, for the purpose of our study, we considered the mean level of all measurements done during the follow-up period, regardless of the context in which the test was taken. As the CRP level is influenced greatly by a variety of acute and chronic conditions, this measurement is not very specific.

The present study has several strengths. First is the large sample-size arising from a sampling frame including more than

half of Israel's population, which allows better precision of estimates for different subgroups and increases the external validity of the study. Moreover, the sampling of outpatients as well as of inpatients and the free medical insurance and services in Israel decrease the potential of referral bias. The use of an unexposed comparator from the general population to assess all-cause mortality is also an advantage, as it was shown to be more precise and less biased than SMRs (30). Finally, the use of CHS registries allowed us to include and adjust, accurately, for socioeconomic variables as well as other comorbidities and lifestyle habits, and enabled us direct access to all the subjects' pharmaceutical and laboratory data. Nonetheless, the use of a large, computerized database has several limitations. The diagnosis of AS was based on International Classification of Diseases codes, and we had no access to the criteria by which the diagnosis was made, and thus misclassification of AS is possible, although diagnosis of other spondyloarthropathies from the same database have been demonstrated previously to be highly valid (21,22).

Second, we had limited access to important clinical variables, such as joint involvement, the presence of anterior uveitis, disease activity measures (except for CRP level), and physical disability. Another limitation is that our analysis does not include a comparison between causes of deaths, as that information belongs to a governmental registry that we did not have access to. Finally, drugs such as NSAIDs, DMARDs, and TNFi are used for indications other than AS, and we could not account for the indication of treatment, and thus confounding by indication is possible. Furthermore, channeling bias due to the fact that subjects were not randomized according to treatment and a selection bias due to better medical care and follow-up for patients being treated with biologics may also influence the results. We performed an adjustment for the main variables we believed were significant in choosing a specific drug such as age, sex, and comorbidities. Yet regardless of the statistical methods used, in this study design, unmeasured confounders and channeling bias could not be overcome, and causality should not be addressed regarding treatment efficacy.

The observations made in this study suggest several directions for further investigations. One is the need to evaluate treatment strategies in chronic inflammatory conditions not only in terms of symptom and articular damage but in terms of survival benefits as well, given the understanding that inflammation is a systemic condition. For example, current guidelines recommend whether or not to start a treatment of TNFi based on the disease activity, CRP levels, and evidence of sacroiliitis (31,32). However, the benefits of TNFi treatment could be portrayed in manners other than articular damage through repair of microvascular damage and reduction of atherosclerosis (5,33). A similar association was only recently reported in a randomized controlled study exploring interleukin-1β targets (34). For this mechanism to be determined, further studies, preferably matched prospective or randomized controlled studies, are warranted.

In conclusion, in this study we found that AS patients had an increased mortality risk compared to the general population. This association remained significant after adjusting for age, sex, and baseline comorbidities. AS patients treated with TNFi did not demonstrate excess mortality compared to matched controls. Within the AS cohort, age, male sex, general comorbidities, and higher CRP levels were identified as risk factors for mortality. Further studies are warranted to explore the effects of disease activity, clinical manifestations, and treatment on AS mortality.

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. Amital 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. Ben-Shabat, Shabat, Watad, McGonagle, Amital.

Acquisition of data. Comaneshter, Cohen.

Analysis and interpretation of data. Ben-Shabat, Kridin, Luigi Bragazzi.

REFERENCES

1. Taurog JD, Chhabra A, Colbert RA. Ankylosing spondylitis and axial spondyloarthritis. N Engl J Med 2016;374:2563–74.
2. Wang R, Ward MM. Epidemiology of axial spondyloarthritis: an update. Curr Opin Rheumatol 2018;30:137–43.
3. Lee JS, Oh BL, Lee HY, Song YW, Lee EY. Comorbidity, disability, and healthcare expenditure of ankylosing spondylitis in Korea: a population-based study. PLoS One 2018;13:e0192524.
4. Sieper J, Poddubnyy D. Axial spondyloarthritis. Lancet 2017;390:73–84.
5. Van Eijk IC, Peters MJL, Serné EH, Van Der Horst-Bruinsma IE, Dijkmans BA, Smulders YM, et al. Microvascular function is impaired in ankylosing spondylitis and improves after tumour necrosis factor α blockade. Ann Rheum Dis 2009;68:362–6.
6. Liew JW, Ramiro S, Gensler LS. Cardiovascular morbidity and mortality in ankylosing spondylitis and psoriatic arthritis. Best Pract Res Clin Rheumatol 2018;32:369–89.
7. Listing J, Kekow J, Manger B, Burmester GR, Pattloch D, Zink A, et al. Mortality in rheumatoid arthritis: the impact of disease activity, treatment with glucocorticoids, TNFα inhibitors and rituximab. Ann Rheum Dis 2015;74:415–21.
8. Curtis JR, Mariette X, Gaujoux-Viala C, Blauvelt A, Kvien TK, Sandborn WJ, et al. Long-term safety of certolizumab pegol in rheumatoid arthritis, axial spondyloarthritis, psoriatic arthritis, psoriasis and Crohn's disease: a pooled analysis of 11 317 patients across clinical trials. RMD Open 2019;5:e000942.
9. Yang ZS, Lin NN, Li L, Li Y. The effect of TNF inhibitors on cardiovascular events in psoriasis and psoriatic arthritis: an updated meta-analysis. Clin Rev Allergy Immunol 2016;51:240–7.
10. Burmester GR, Panaccione R, Gordon KB, McIlraith MJ, Lacerda AP. Adalimumab: long-term safety in 23 458 patients from global clinical trials in rheumatoid arthritis, juvenile idiopathic arthritis, ankylosing spondylitis, psoriatic arthritis, psoriasis and Crohn's disease. Ann Rheum Dis 2013;72:517–24.
11. Dregan A, Chowienczyk P, Molokhia M. Cardiovascular and type 2 diabetes morbidity and all-cause mortality among diverse chronic inflammatory disorders. Heart 2017;103:1867–73.

12. Lehtinen K. Mortality and causes of death in 398 patients admitted to hospital with ankylosing spondylitis. Ann Rheum Dis 1993;52:174–6.
13. Radford EP, Doll R, Smith PG. Mortality among patients with ankylosing spondylitis not given X-ray therapy. N Engl J Med 1977;297:572–6.
14. Exarchou S, Lie E, Lindström U, Askling J, Forsblad-D'Elia H, Turesson C, et al. Mortality in ankylosing spondylitis: results from a nationwide population-based study. Ann Rheum Dis 2016;75: 1466–72.
15. Bakland G, Gran JT, Nossent JC. Increased mortality in ankylosing spondylitis is related to disease activity. Ann Rheum Dis 2011;70: 1921–5.
16. Buschiazzo EA, Schneeberger EE, Sommerfleck FA, Ledesma C, Citera G. Mortality in patients with ankylosing spondylitis in Argentina. Clin Rheumatol 2016;35:2229–33.
17. Prati C, Puyraveau M, Guillot X, Verhoeven F, Wendling D. Deaths associated with ankylosing spondylitis in France from 1969 to 2009. J Rheumatol 2017;44:594–8.
18. Rennert G, Peterburg Y. Prevalence of selected chronic diseases in Israel. Isr Med Assoc J 2001;3:404–8.
19. Watad A, McGonagle D, Bragazzi NL, Tiosano S, Comaneshter D, Shoenfeld Y, et al. Autoantibody status in systemic sclerosis patients defines both cancer risk and survival with ANA negativity in cases with concomitant cancer having a worse survival. Oncoimmunology 2019; 8:e1588084.
20. Yavne Y, Tiosano S, Watad A, Comaneshter D, Cohen AD, Amital H. Investigating the link between ischemic heart disease and Behcet's disease: a cross-sectional analysis. Int J Cardiol 2017;241:41–5.
21. Zisman D, Bitterman H, Shalom G, Feldhamer I, Comanesther D, Batat E, et al. Psoriatic arthritis treatment and the risk of herpes zoster. Ann Rheum Dis 2016;75:131–5.
22. Eder L, Cohen AD, Feldhamer I, Greenberg-Dotan S, Batat E, Zisman D. The epidemiology of psoriatic arthritis in Israel: a population-based study. Arthritis Res Ther 2018;20:3.
23. Ben-Shabat N, Tiosano S, Shovman O, Comaneshter D, Shoenfeld Y, Cohen AD, et al. Mortality among patients with giant-cell arteritis: a large-scale population-based cohort study. J Rheumatol 2020;47: 1385–91.
24. Ben-shabat N, Watad A, Shabat A, Bragazzi NL, Gan R. Low vitamin D levels predict mortality in ankylosing spondylitis patients: a nationwide population-based cohort study. Nutrients 2020;12:1400.
25. 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.
26. Szabo SM, Levy AR, Rao SR, Kirbach SE, Lacaille D, Cifaldi M, et al. Increased risk of cardiovascular and cerebrovascular diseases in individuals with ankylosing spondylitis: a population-based study. Arthritis Rheum 2011;63:3294–304.
27. Kang JH, Chen YH, Lin HC. Comorbidity profiles among patients with ankylosing spondylitis: a nationwide population-based study. Ann Rheum Dis 2010;69:1165–8.
28. Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. Arthritis Res Ther 2009;11:229.
29. Callhoff J, Sieper J, Weiß A, Zink A, Listing J. Efficacy of TNFα blockers in patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis: a meta-analysis. Ann Rheum Dis 2015;74: 1241–8.
30. Card TR, Solaymani-Dodaran M, Hubbard R, Logon RF, West J. Is an internal comparison better than using national data when estimating mortality in longitudinal studies? J Epidemiol Community Health 2006;60:819–21.
31. Ward MM, Deodhar A, Akl EA, Lui A, Ermann J, Gensler LS, et al. American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and Treatment Network 2015 recommendations for the treatment of ankylosing spondylitis and nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2016;68: 282–98.
32. Van Der Heijde D, Ramiro S, Landewé R, Baraliakos X, Van Den Bosch F, Sepriano A, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis 2017;76:978–91.
33. Zardi EM, Pipita ME, Giorgi C, Lichinchi D, Zardi DM, Afeltra A. Differences in carotid atherosclerosis between patients with ankylosing spondylitis treated with tumor necrosis factor-α antagonists and healthy matched controls. Medicine (Baltimore) 2018;97:e11250.
34. Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, et al. Antiinflammatory therapy with canakinumab for atherosclerotic disease. N Engl J Med 2017;377:1119–31.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1623–1630
DOI 10.1002/acr.24606

American College of Rheumatology
Empowering Rheumatology Professionals

Comparison of Responsiveness of British Isles Lupus Assessment Group 2004 Index, Systemic Lupus Erythematosus Disease Activity Index 2000, and British Isles Lupus Assessment Group 2004 Systems Tally

Chee-Seng Yee,[1] Caroline Gordon,[2] David A. Isenberg,[3] Bridget Griffiths,[4] Lee-Suan Teh,[5] Ian N. Bruce,[6] Yasmeen Ahmad,[7] Anisur Rahman,[3] Athiveeraramapandian Prabu,[2] Mohammed Akil,[8] Neil McHugh,[9] Christopher J. Edwards,[10] David D'Cruz,[11] Munther A. Khamashta,[12] and Vernon T. Farewell[13]

Objective. To compare the responsiveness of the British Isles Lupus Assessment Group 2004 index (BILAG-2004) and the Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) disease activity indices and to determine whether there was any added value in combining BILAG-2004, BILAG-2004 system tally (BST), or simplified BST (sBST) with SLEDAI-2K.

Methods. This was a multicenter longitudinal study of SLE patients. Data were collected on BILAG-2004, SLEDAI-2K, and therapy on consecutive assessments in routine practice. The external responsiveness of the indices was assessed by determining the relationship between change in disease activity and change in therapy between 2 consecutive visits. Comparison of indices and their derivatives was performed by assessing the main effects of the indices using logistic regression. Receiver operating characteristic curves analysis was used to describe the performance of these indices individually and in various combinations, and comparisons of area under the curve were performed.

Results. There were 1,414 observations from 347 patients. Both BILAG-2004 and SLEDAI-2K maintained an independent relationship with change in therapy when compared. There was some improvement in responsiveness when continuous SLEDAI-2K variables (change in score and score of previous visit) were combined with BILAG-2004 system scores. Dichotomization of BILAG-2004 or SLEDAI-2K resulted in poorer performance. BST and sBST had similar responsiveness as the combination of SLEDAI-2K variables and BILAG-2004 system scores. There was little benefit in combining SLEDAI-2K with BST or sBST.

Conclusion. The BILAG-2004 index had comparable responsiveness to SLEDAI-2K. There was some benefit in combining both indices. Dichotomization of BILAG-2004 and SLEDAI-2K leads to suboptimal performance. BST and sBST performed well on their own; sBST is recommended for its simplicity and clinical meaningfulness.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a complex multisystem disease, and assessment of this disease is challenging, given the multiple outcome domains to be considered. The 2 commonly used disease activity indices that allow the results from different cohorts of SLE patients to be compared in clinical trials or observational studies are the British Isles Lupus Assessment Group

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

Supported by Versus Arthritis (grant 16081), the Medical Research Council UK (grant U105261167), and Vifor Pharma/Aspreva Pharmaceuticals. The Birmingham SLE clinics were supported by Lupus UK. Some of the work for this study was carried out at the NIHR/Wellcome Trust Birmingham Clinical Research Facility. The work of Drs. Isenberg and Rahman was supported by the NIHR University College London Hospitals Biomedical Research Centre. Dr. Bruce's work was supported by Versus Arthritis, NIHR Manchester Biomedical Research Unit, and NIHR Manchester Wellcome Trust Clinical Research Facility.

[1]Chee-Seng Yee, PhD, FRCP(UK): Doncaster and Bassetlaw Teaching Hospitals NHS Foundation Trust, Doncaster, UK; [2]Caroline Gordon, MD, FRCP(UK), Athiveeraramapandian Prabu, MD, FRCP(UK): University of Birmingham, Birmingham, UK; [3]David A. Isenberg, MD, FRCP(UK), Anisur Rahman, PhD, FRCP(UK): University College London, London, UK; [4]Bridget Griffiths, MD, FRCP(UK): Newcastle-upon-Tyne Hospitals NHS Foundation Trust, Newcastle-upon-Tyne, UK; [5]Lee-Suan Teh, MD, FRCP(UK): Royal Blackburn Teaching Hospital, Blackburn, and University of Central Lancashire, Preston, UK; [6]Ian N. Bruce, FRCP(UK): University of Manchester, Manchester, UK; [7]Yasmeen Ahmad, PhD, MRCP(UK): Betsi Cadwaladr University Health Board, Bangor, Gwynedd, UK; [8]Mohammed Akil, FRCP(UK): Sheffield Teaching Hospitals NHS Trust, Sheffield, UK; [9]Neil McHugh, FRCP(UK): Royal National Hospital for Rheumatic Diseases NHS Trust, Bath, UK; [10]Christopher J. Edwards, MD, FRCP(UK): University Hospital Southampton NHS Foundation Trust, Southampton, UK; [11]David D'Cruz, MD, FRCP(UK): Guy's Hospital, London, UK; [12]Munther A. Khamashta, MD, FRCP(UK): King's College London, London, UK; [13]Vernon T. Farewell, PhD: MRC Biostatistics Unit, Cambridge, UK.

Dr. Yee has received consulting fees from Bristol Myers Squibb and Immupharma (less than $10,000 each). Dr. Gordon has received consulting

SIGNIFICANCE & INNOVATIONS

- Various ways of analyzing the British Isles Lupus Assessment Group 2004 index (BILAG-2004) and Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K), and their derivatives, have been employed in longitudinal studies of systemic lupus erythematosus (SLE), especially in clinical trials. However, a direct comparison of these 2 indices and their various combinations has not been made to determine the best way of using them without the addition of a physician's global assessment.
- The results of this analysis provide guidance on the use of these indices as disease activity outcome measures in longitudinal studies of SLE. The key findings from this analysis are: 1) both the BILAG-2004 index and the SLEDAI-2K have similar responsiveness, and there is some improvement when they are combined; 2) dichotomization of the BILAG-2004 index and the SLEDAI-2K may reduce performance as an outcome measure; and 3) the simplified BILAG system tally may have an advantage due to its simplicity and clinical meaningfulness.

2004 index (BILAG-2004) (1–5) and the Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) (6–8).

A strong correlation between the classic BILAG index and the original SLEDAI was demonstrated using patient vignettes, but there has been no direct comparison of the performance of BILAG-2004 and SLEDAI-2K using real-world clinical data (9,10). Various attempts have been made to combine SLEDAI (or its derivatives) with BILAG-2004 or classic BILAG indices in clinical trials, in the belief that a combination might be superior to either index on its own (11–14). However, few data are available to support this presumption, and concerns exist about the impact of variable recording of the physician's global assessment (PhGA) by different physicians (9) in composite responder indices such as the SLE Responder Index (SRI) and its derivatives (11,13–16) and the BILAG Composite Lupus Assessment (BICLA) (12,17,18). These composite clinical trial end points focus on changes, in particular on patients showing specific levels of improvement in 1 index at the final trial visit as compared to baseline visit and require no worsening in the alternative index and PhGA. Both SRI and BICLA are currently used as end points in clinical trials of SLE, but trial results have been inconsistent, including some with promising results in phase II studies but negative results in phase III or with disappointing results generally (12,15,17,19–21). One of the concerns with trials that failed was with the outcome measure used as the primary end point being not optimal (22). This study reports on the analysis comparing BILAG-2004 and SLEDAI-2K and tries to determine the best way of using these indices without PhGA in longitudinal studies.

We have previously demonstrated the external responsiveness of BILAG-2004 and SLEDAI-2K (4,23). Employing similar robust methodology (24), the analyses presented here examined whether the use of both indices improves the responsiveness of each alone using data from a large longitudinal study of SLE patients seen in routine practice. We also compared the performance of the BILAG-2004 systems tally (BST). BST is an alternative way of representing BILAG-2004 scores in a longitudinal assessment that combines the flexibility and simplification of overall numerical scoring of BILAG-2004 with the clinical intuitiveness of BILAG-2004 structure (25).

PATIENTS AND METHODS

Data from a multicenter prospective longitudinal study in the UK, which was primarily designed to validate BILAG-2004, were used in this analysis (4). This same data set was used to demonstrate the external responsiveness of SLEDAI-2K and to develop BST and simplified BST (sBST) (23,25). All patients satisfied the revised American College of Rheumatology criteria for classification of SLE (26,27). This study received multicenter research ethical approval and was carried out in accordance with the Helsinki Declaration. Written consent was obtained from all patients.

This study has been described in detail previously (4). In summary, patients were followed up prospectively in routine clinical practice and data (BILAG-2004 index, SLEDAI-2K score, and treatment) were collected for all consecutive visits and physician encounters. Previously we demonstrated, based on receiver operating characteristic (ROC) curve analyses, that BST, sBST, and BILAG-2004 global numerical variables (combination of change in BILAG-2004 global numerical score [5] and the score from the previous visit), were comparably related to change in therapy and provided better discrimination than a model including variables for changes in all 9 BILAG-2004 system scores (25). In the analyses presented here, we included disease activity as assessed by SLEDAI-2K, BILAG-2004 individual system scores, and global BILAG-2004 numerical score.

Changes in disease activity and treatment between 2 consecutive visits were analyzed. Each observation for the analysis was derived from 2 consecutive visits. A robust definition for change in therapy between consecutive visits was used as the external reference for change in disease activity as described previously (see Supplementary Appendix A, available on the *Arthritis*

fees from UCB and MGP (less than $10,000 each). Dr. Teh has received grants from Pfizer and AbbVie. Dr. Akil has received consulting or speaking fees from Action, AstraZeneca, and Gilead (less than $10,000 each). Dr. D'Cruz has received consulting or speaking fees from GlaxoSmithKline (less than $10,000). No other disclosures relevant to this article were reported.

Address correspondence to Chee-Seng Yee, PhD, FRCP(UK), Consultant Rheumatologist, Department of Rheumatology, Doncaster Royal Infirmary, Armthorpe Road, Doncaster, DN2 5LT UK. Email: csyee@ymail.com.

Submitted for publication July 20, 2020; accepted in revised form March 23, 2021.

Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606) (3–5,23,25). Three categories of changes in therapy were defined: no change, increase in therapy, and decrease in therapy. All statistical analyses were performed using Stata for Windows, version 8, and R (28). Robust variance estimation was used to allow for correlation between multiple assessments from the same patients (29).

External responsiveness was used to compare the performance of the indices in this longitudinal study (24). It assessed the extent to which changes in the index over time relate to corresponding changes in therapy between 2 consecutive visits. Therefore, clinically meaningful change was assessed. Change in therapy was chosen as the external reference, as there was no better objective alternative, and this criterion has been used in multiple validation studies on BILAG-2004, SLEDAI-2K, and BST (3–5,23,25). The pros and cons of using change in therapy as the external reference were discussed previously (3).

Maximum-likelihood multinomial and binary logistic regression were used to assess external responsiveness, with change in therapy as the outcome variable and changes in disease activity (as determined by the indices) as the explanatory variables. For comparison purposes, the main effects of the indices were assessed within a common regression model. The baseline comparator for change in disease activity used in the analysis was minimal or no change in activity, while the baseline comparator for change in treatment was no change in therapy. The results were reported as odds ratios (ORs) with 95% confidence intervals (95% CIs), and Wald tests were used for model comparison where needed.

In the multinomial regression analyses, the baseline category of no change in therapy was compared with both increase in therapy and decrease in therapy. There was no direct comparison between increase in therapy and decrease in therapy. An OR value of >1 for a 1-unit increase in the variable defined by the index of interest, within the comparison between increase in therapy and no change in therapy, indicated that the increase in the index score was associated with an increase in therapy. Conversely, an OR of <1 for the same comparison implied that the increase in the index score was associated with no change in therapy (and not with a decrease in therapy) or equivalently an inverse association with an increase in therapy. Similar interpretation was applicable to the reported OR for the comparison between decrease in therapy and no change in therapy.

Various combinations of SLEDAI-2K and BILAG-2004 (including BST and sBST) as dichotomized or regarded as continuous variables were examined and compared to determine whether there was added value in combining both of these indices. For some analyses, the BILAG-2004 global numerical score was calculated based on the system scores using a coding scheme of A = 12, B = 8, C = 1, D/E = 0 (5). ROC curves analysis was used to describe the performance of these indices and the various combinations (30). Logistic regression was used to estimate the sensitivity, specificity, positive predictive value, negative predictive value, and area under the curve (AUC). The analyses were performed from 2 perspectives: deterioration in scores as a predictor of increase in therapy and improvement in scores as a predictor of decrease in therapy. Calculation of an asymptotic confidence interval for AUC and comparison of AUCs were performed using a nonparametric approach (31). AUC, with a value from 0 to 1, quantified the performance of the index, with the value of 1 corresponding to the index providing perfect discrimination.

Deterioration in the BILAG-2004 score was defined to have occurred if there was worsening in the score to grades A or B in any of the systems. The deteriorations were classified (in order of ranking) as: 1) major deterioration: change from C/D/E to A or from D/E to B, 2) minor A deterioration: change from B to A, and 3) minor B deterioration: change from C to B. A change from D/E to C was considered minor and not clinically significant. Therefore, such a change was excluded from the definition of deteriorations.

Improvement in the BILAG-2004 score was deemed to have occurred if there was reduction in the score in any system in the absence of any deterioration in the other systems. The improvements were classified (in order of ranking) as 1) major improvement: change from A to C/D or B to D, 2) minor A improvement: change from A to B, and 3) minor B/C improvement: change from B to C or C to D. These classifications were used to define BILAG-2004–based explanatory variables in regression analyses. The definitions and gradations above were based on the principle of intent-to-treat that underlay BILAG-2004 scoring, whereby active disease requiring therapy was graded A or B depending on the item, while grade C usually required symptomatic therapy (1). We accepted that at the individual patient and organ level, there may be variation in the severity of the disease items and the need for change in therapy within each grade.

BST and its simplified version, sBST, were counts of systems with specified changes in scores between 2 assessments (25). BST comprised 6 components: 1) the number of systems with major deterioration (change of B/C/D/E to A, or D/E to B), 2) the number of systems with minor deterioration (change of C to B), 3) the number of systems with persistent significant activity (no change from A or B), 4) the number of systems with major improvement (change of A to C/D, or B to D), 5) the number of systems with minor improvement (change of A to B, or B to C), and 6) the number of systems with persistent minimal or no activity (change of C/D/E to C/D/E).

Simplified BST had 3 components: 1) the number of systems with active/worsening disease (systems with major deterioration, minor deterioration, and persistent significant activity); 2) the number of systems with improving disease (systems with major improvement and minor improvement); and 3) the number of systems with persistent minimal or no activity.

RESULTS

There were 347 SLE patients with 1,761 assessments that contributed 1,414 observations for this analysis. There was an increase in treatment in 22.7% of the observations, while 37.3% had therapy decreased, and in 40.0%, there was no change in treatment, as previously reported (4). The demographic characteristics and distribution of change in disease activity for each system are summarized in Supplementary Tables 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606.

Comparison of BILAG-2004 with SLEDAI-2K. To examine the combined performance of SLEDAI-2K and BILAG-2004, we undertook multinomial logistic regression analysis of change in therapy, using the changes in both BILAG-2004 and SLEDAI-2K with and without their respective values at the previous visit. We had demonstrated previously that although change in the SLEDAI-2K score was significantly associated with changes in treatment, the strongest relationship was observed in a model that included both the change in the SLEDAI-2K score and the score at the previous visit as continuous variables (hereby referred to as SLEDAI-2K variables) (23).

In the analysis of external responsiveness reported here, changes in the individual system scores of BILAG-2004 and SLEDAI-2K variables (as a continuous variable) were included as explanatory variables for the outcome variable of change in therapy. Table 1 shows that SLEDAI-2K variables and individual BILAG-2004 system scores retained independent relationships with change in therapy. Consistent with our earlier work (23), if only the change in SLEDAI-2K score was included (i.e., the SLEDAI-2K score of the previous visit was omitted), the change in SLEDAI-2K score was no longer significantly associated with change in therapy (increase or decrease), while changes in BILAG-2004 system scores maintained their significant association with change in therapy (data not shown).

When we undertook a multinomial logistic regression analysis of change in therapy using change in the numerical score of BILAG-2004 and in SLEDAI-2K, along with their respective values at the previous visit (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606), we observed the expected relationships between the changes in the numerical scores and changes in therapy. Both pairs of variables, the 2 based on BILAG-2004 and the 2 based on SLEDAI-2K, added predictive power for an increase in therapy (P = 0.02 for the addition of SLEDAI-2K variables to BILAG-2004 numerical score variables, and $P < 0.01$ for the addition of BILAG-2004 numerical score variables to SLEDAI-2K variables by Wald test). For decrease in therapy, the SLEDAI-2K variables did not provide additional predictive power (P = 0.50 by Wald test).

Table 1. External responsiveness of the combination of the BILAG-2004 and SLEDAI-2K indices with multinomial logistic regression (n = 1,414)*

Change in score	Increase in therapy†	Decrease in therapy†
BILAG-2004 index system score‡		
Constitutional		
Increase	∞	1.35 (0.87–2.08)
Decrease	0.86 (0.27–2.71)	2.26 (1.25–4.06)
Mucocutaneous		
Increase	7.52 (4.36–12.98)	0.63 (0.31–1.28)
Decrease	0.88 (0.56–1.38)	1.49 (1.09–2.05)
Neuropsychiatric		
Increase	∞	1.85 (0.49–7.02)
Decrease	0.98 (0.20–4.79)	1.96 (0.85–4.51)
Musculoskeletal		
Increase	11.93 (5.32–26.76)	1.10 (0.42–2.88)
Decrease	0.69 (0.44–1.08)	0.96 (0.69–1.33)
Cardiorespiratory		
Increase	2.88 (0.96–8.60)	0.71 (0.24–2.05)
Decrease	1.17 (0.50–2.75)	1.42 (0.82–2.47)
Gastrointestinal		
Increase	7.74 (0.67–89.43)	0
Decrease	0.64 (0.18–2.31)	1.14 (0.26–4.88)
Ophthalmic		
Increase	1.32 (0.01–270.27)	0
Decrease	4.25 (0.51–35.06)	1.47 (0.22–9.99)
Renal		
Increase	1.08 (0.32–3.72)	3.14 (0.95–10.40)
Decrease	0.63 (0.26–1.54)	1.76 (0.97–3.19)
Hematologic		
Increase	–§	–§
Decrease	0.95 (0.56–1.60)	0.90 (0.64–1.28)
Change in SLEDAI-2K score	1.17 (1.08–1.27)	0.93 (0.87–1.00)
Previous visit SLEDAI-2K score	1.18 (1.11–1.26)	0.96 (0.91–1.02)

* Values are the odds ratio (95% confidence interval). BILAG-2004 = British Isles Lupus Assessment Group 2004 index; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000; ∞ = infinity.
† Compared to no change in therapy.
‡ Compared to minimal change (including change of grade D/E to C) within each system.
§ No observation with increase in score.

As shown in Table 2, we observed that BST variables were related to changes in therapy in the expected manner, and that SLEDAI-2K variables provided additional predictive power for an increase in therapy (P = 0.007 based on the Wald test from separate logistic regression) but not for a decrease in therapy (P = 0.30 by Wald test). Similar results were obtained with sBST (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606).

Comparison of performance of combinations of BILAG-2004 and SLEDAI-2K. Table 3 summarizes the results of further analyses using various combinations of information from BILAG-2004 and SLEDAI-2K. The table shows the AUC measures based on ROC curves derived from binary regression analyses of both increase in therapy and decrease in therapy versus no change in therapy. For completeness, we performed similar

Table 2. External responsiveness of the combination of BILAG-2004 systems tally and SLEDAI-2K indices with multinomial logistic regression (n = 1,414)*

Change in score	Increase in therapy†	Decrease in therapy†
BILAG-2004 systems tally‡		
Major deterioration	14.35 (8.51–24.21)	0.85 (0.48–1.52)
Minor deterioration	5.72 (2.76–11.86)	1.04 (0.55–1.94)
Persistent significant activity	5.54 (3.26–9.43)	0.65 (0.41–1.05)
Major improvement	0.95 (0.58–1.56)	1.57 (1.11–2.23)
Minor improvement	1.47 (0.89–2.43)	1.27 (0.91–1.79)
Change in SLEDAI-2K score	1.08 (1.01–1.17)	0.96 (0.90–1.03)
Previous visit SLEDAI-2K score	0.99 (0.91–1.07)	1.00 (0.94–1.07)

* Values are the odds ratio (95% confidence interval). BILAG-2004 = British Isles Lupus Assessment Group 2004 index; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000.
† Compared to no change in therapy.
‡ Compared to persistent minimal or no activity (change of grade C/D/E to C/D/E).

analyses of increase in therapy versus no increase in therapy and decrease in therapy versus no decrease in therapy.

The comparison of AUC measures from this exploratory analysis for increase in treatment versus no increase in treatment are summarized in Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606, which provides the significance levels for the comparison of the various models. The *P* values should be regarded as illustrative, as no adjustment for multiplicity was performed. Similar results were obtained for analysis for increase in treatment versus no change in treatment (see Supplementary Table 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606). The analysis showed no evidence that either BILAG-2004 system scores or SLEDAI-2K variables were more predictive of changes in therapy individually than the other ($P = 0.89$ by Wald test). There was some improvement in the performance from the combination of both BILAG-2004 system scores and SLEDAI-2K variables ($P < 0.001$ for the addition of each to the other by Wald test). BST and sBST had comparable performance ($P = 0.107$ by Wald test) and were, respectively, similar to ($P = 0.128$ by Wald test) or slightly worse than ($P < 0.001$ by Wald test) BILAG-2004 numerical score variables (change in numerical score and previous visit numerical score). BST, sBST, and BILAG-2004 numerical score variables appeared to be more predictive of an increase in therapy compared to BILAG-2004 system scores ($P < 0.001$, $P = 0.002$, and $P < 0.001$, respectively, by Wald test) and SLEDAI-2K variables ($P < 0.001$, $P = 0.013$, and $P < 0$ .001, respectively, by Wald test). Furthermore, BST, sBST, and BILAG-2004 numerical score variables were comparable to or slightly better than the combination of BILAG-2004 system scores and SLEDAI-2K variables ($P = 0.63$, $P = 0.26$, and $P = 0.03$, respectively, by Wald test). Finally, the addition of SLEDAI-2K variables provided little improvement to the performance of BST, sBST, or BILAG-2004 numerical score variables ($P = 0.60$, $P = 0.16$, and $P = 0.22$, respectively, by Wald test).

Dichotomization of indices. Dichotomized versions of the BILAG-2004 and the SLEDAI-2K have been used for a variety of purposes. In the supplementary material using dichotomized variables to analyze deterioration of activity and improvement in

Table 3. Area under the curve values from receiver operating characteristics curves analysis of the BILAG-2004 index, SLEDAI-2K, and combination of the 2 indices*

	Increase in therapy		Decrease in therapy	
	Versus no change	Versus no increase	Versus no change	Versus no decrease
BILAG-2004 index system scores†	0.75 (0.70–0.79)	0.75 (0.71–0.78)	0.59 (0.56–0.62)	0.65 (0.62–0.67)
BST	0.82 (0.78–0.87)	0.83 (0.81–0.86)	0.57 (0.54–0.61)	0.66 (0.63–0.68)
Simplified BST	0.81 (0.78–0.84)	0.81 (0.78–0.84)	0.57 (0.54–0.60)	0.65 (0.63–0.68)
BILAG 2004 numerical score variables‡	0.84 (0.81–0.87)	0.85 (0.82–0.88)	0.58 (0.55–0.62)	0.67 (0.65–0.70)
SLEDAI-2K variables§	0.75 (0.71–0.78)	0.76 (0.73–0.79)	0.56 (0.53–0.60)	0.63 (0.60–0.66)
BILAG 2004 index system scores plus SLEDAI-2K variables	0.80 (0.77–0.83)	0.81 (0.78–0.84)	0.60 (0.57–0.64)	0.67 (0.64–0.70)
BST plus SLEDAI-2K variables	0.84 (0.81–0.86)	0.84 (0.81–0.87)	0.59 (0.55–0.62)	0.67 (0.64–0.70)
Simplified BST plus SLEDAI-2K variables	0.82 (0.79–0.85)	0.83 (0.80–0.86)	0.58 (0.55–0.62)	0.67 (0.64–0.69)
BILAG-2004 numerical score variables plus SLEDAI-2K variables	0.84 (0.82–0.87)	0.85 (0.83–0.88)	0.59 (0.56–0.62)	0.68 (0.65–0.71)

* Values are the area under the curve (95% confidence interval). BILAG-2004 = British Isles Lupus Assessment Group 2004 index; BST = BILAG-2004 systems tally; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000.
† 9 separate changes in system scores.
‡ Change in numerical score and previous visit numerical score.
§ Change in SLEDAI-2K score and previous visit SLEDAI-2K score.

activity (Supplementary Tables 7 and 8, respectively, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606), the results for clinically relevant dichotomizations were given for the 2 indices, separately and in combination. These were based on multinomial regressions with a single binary explanatory variable.

Two particular categorizations of changes in the combination of these measures that were of similar magnitude to those used in the definition of SRI (SLEDAI-2K score decrease of ≥4 and no BILAG-2004 deterioration) (11) and BICLA (all improvements in BILAG-2004 with no SLEDAI-2K score increase of ≥1) (12) were included in Supplementary Table 8, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606, examining improvement in disease activity, but without PhGA. The change was also between 2 consecutive visits (not between the start and end of study). The estimated sensitivities and specificities were 1.5% and 98.9%, respectively, for the SRI-like variable, and 48.2% and 70.0%, respectively, for the BICLA-like variable when used to predict a decrease in therapy (versus no decrease). The AUC values for these 2 variables were 0.50 and 0.59, respectively, compared with AUCs >0.65 for BST, sBST, and BILAG-2004 numerical variables (Table 3). Other dichotomized variables also did not perform as well as these numerical variables in relation to both decrease in therapy and increase in therapy.

DISCUSSION

This multicenter observational study directly compared the responsiveness of the BILAG-2004 index with the SLEDAI-2K in longitudinal fashion and assessed the potential value of combining the 2 indices using a comprehensive range of approaches. Our analyses showed that there was some nonoverlapping relationship with change in therapy when both BILAG-2004 and SLEDAI-2K were included in the model, confirming that both indices had similar responsiveness. Responsiveness was optimal if both the change in the SLEDAI-2K score and the SLEDAI-2K score of the previous visit were included in the model as continuous variables. The use of only change in the SLEDAI-2K score was associated with inferior performance.

Outcome in clinical trials is determined by 3 factors: efficacy of intervention, study design, and effectiveness of the outcome measure used. Our discussion is focused on properties of the outcome measure that would affect its ability to differentiate the efficacy of the different treatment arms. Discussing the other factors is beyond the scope of the present work.

Many clinical trials in SLE have reported their results using various combinations of the SLEDAI-2K or the Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA)–SLEDAI (and its variants) with the BILAG-2004 or the classic BILAG index as the primary end points (22). In the belimumab phase III trials, the SRI was used in which a response was defined as an improvement in the SELENA-SLEDAI score of at least 4 points with no new grade A and ≤1 new grade B classic BILAG system score, and no deterioration of PhGA (13,14). This combination was selected using the data set from the phase II trial of belimumab to derive the best separation in efficacy between belimumab and placebo, with the presumption that belimumab was effective (11). Using a similar combination of improvement in the SLEDAI-2K score of at least 4 points with no worsening of the BILAG-2004 system score to grade A or B, we found that this combination performed poorly when assessed using the reference of change in therapy. This finding was surprising as we would have expected these 2 indices to exert a greater role than PhGA, which is subject to variable reporting due to individual physicians scoring lupus manifestations differently from each other in the absence of a glossary, particularly in patients with >1 system involved (9). The indices used in this study were different from the original SRI (BILAG-2004 instead of classic BILAG index, SLEDAI-2K instead of SELENA-SLEDAI and no PhGA). The modified SRI used in the analysis was very similar to the indices used successfully in the phase 2 trial of ustekinumab (16), but which failed as the primary end point in phase 3 trials of anifrolumab (15) and Lupuzor (19). These trials used a modification of the SRI in which response was driven by a 4-point reduction in the SLEDAI-2K score with ≤1 new B grade in BILAG-2004 and ≤10% worsening of PhGA.

A different combination (BICLA) was used in other clinical trials, in which a response was defined as an improvement in the BILAG-2004 system score (in the absence of new grade A or B score) with no worsening of the SLEDAI-2K score (≥1) and no worsening of PhGA (12,15,17,18). The results of our study, shown in Supplementary Table 8, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24606, supported the use of this combination of BILAG-2004 and SLEDAI-2K indices, which although not successful in the epratuzumab phase 3 trial (17), was successful in phase 3 trials of anifrolumab as primary (TULIP-2 trial) and secondary end points (TULIP-1 trial) (15,18).

Currently, the combination of the 2 indices (BILAG-2004 with SLEDAI-2K) used in clinical trials involves dichotomizations of the outcome variables. Our data suggested that the benefit was minimal when combining these 2 indices in this specific way, and the value of PhGA was debatable (9). Dichotomization involves using a cutoff to determine whether a response is achieved (yes/no response). However, dichotomization of variables may result in loss of efficiency, as it does not allow for a graded response, and a partial response might be considered lack of response if the cutoff is not achieved (32). We demonstrated that dichotomization of both BILAG-2004 and SLEDAI-2K resulted in poorer responsiveness in our longitudinal study. With better efficiency and performance of the outcome measure used, fewer patients would be required in a study to demonstrate differences between groups, which then facilitates target recruitment and reduces the cost of running the study. In comparison to the use of a continuous outcome, the size of a trial may need to be increased by a factor of 30% if a binary outcome with a uniform distribution is used

with a median cutoff, with greater gains for a normal distribution (32). By using BST or sBST, which are based on counts of systems with specified transitions in BILAG-2004 scores, the problem of dichotomization could be avoided.

Although BILAG-2004 numerical score variables and the combination of the SLEDAI-2K variables with BST or sBST had a slightly better performance than BST or sBST alone, BST and sBST performed better than BILAG-2004 system scores and SLEDAI-2K. In addition, there was difficulty with interpretation of the clinical meaningfulness of BILAG-2004 numerical score variables and the combination of SLEDAI-2K variables with BST or sBST. Our analyses supported the use of BST or sBST alone and suggested minimal advantage of combining SLEDAI-2K with BST or sBST. Consequently, there could be simplification in study methodology by using only 1 disease activity index (BILAG-2004), which would avoid confusion and reduce errors due to differences in BILAG-2004 and SLEDAI-2K glossaries.

One limitation of this study that might affect the applicability of the results to clinical trials was the time reference used to define change in disease activity. This study looked at the changes between consecutive visits. In contrast, clinical trials generally compare the disease activity between the beginning and the end of the study (and not between consecutive visits), which might be 1 year apart. With a longer time interval, a larger effect is far more likely to occur. However, comparing the outcome measures at only 2 time points (the beginning and the end of study) ignores the level of disease activity between these 2 time points. The use of counts or a continuous variable over the study period (such as flare rate) could overcome this disadvantage. Another limitation was that BST and sBST were developed using the same data set, which might have provided an advantage. Validation of our result with an independent data set is needed.

In conclusion, both BILAG-2004 and SLEDAI-2K have similar responsiveness longitudinally. There is some benefit in combining the 2 indices, but dichotomization of the indices leads to suboptimal performance. BST and sBST performed well on their own and the addition of SLEDAI-2K variables only resulted in minimal improvement. There is no significant difference with the responsiveness of BST or sBST. Given that sBST has only 3 components, we would recommend the use of sBST in longitudinal analysis of disease activity for its simplicity and clinical meaningfulness.

ACKNOWLEDGMENTS

We would like to thank all the nurse specialists and doctors who contributed to this study at the participating centers.

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. Yee 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. Yee, Gordon, Isenberg, Griffiths, Teh, Bruce, Ahmad, Rahman, Prabu, Akil, McHugh, Edwards, D'Cruz, Khamashta, Farewell.

Acquisition of data. Yee, Gordon, Isenberg, Griffiths, Teh, Bruce, Ahmad, Rahman, Prabu, Akil, McHugh, Edwards, D'Cruz, Khamashta, Farewell.

Analysis and interpretation of data. Yee, Gordon, Isenberg, Farewell.

ROLE OF THE STUDY SPONSOR

Vifor Pharma/Aspreva Pharmaceuticals had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Vifor Pharma/Aspreva Pharmaceuticals.

REFERENCES

1. Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004: development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford) 2005;44:902–6.
2. Yee CS, Farewell V, Isenberg DA, Prabu A, Sokoll K, Teh LS, et al. Revised British Isles Lupus Assessment Group 2004 Index: a reliable tool for assessment of systemic lupus erythematosus activity. Arthritis Rheum 2006;54:3300–5.
3. Yee CS, Farewell V, Isenberg DA, Rahman A, Teh LS, Griffiths B, et al. British Isles Lupus Assessment Group 2004 Index is valid for assessment of disease activity in systemic lupus erythematosus. Arthritis Rheum 2007;56:4113–9.
4. Yee CS, Farewell V, Isenberg DA, Griffiths B, Teh LS, Bruce IN, et al. The BILAG-2004 index is sensitive to change for assessment of SLE disease activity. Rheumatology (Oxford) 2009;48:691–5.
5. Yee CS, Cresswell L, Farewell V, Rahman A, Teh LS, Griffiths B, et al. Numerical scoring for the BILAG-2004 index. Rheumatology (Oxford) 2010;49:1665–9.
6. Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol 2002;29:288–91.
7. Touma Z, Urowitz MB, Gladman DD. SLEDAI-2K for a 30-day window. Lupus 2010;19:49–51.
8. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH, and the Committee on Prognosis Studies in SLE. Derivation of the SLEDAI: a disease activity index for lupus patients. Arthritis Rheum 1992;35:630–40.
9. Wollaston SJ, Farewell VT, Isenberg DA, Gordon C, Merrill JT, Petri MA, et al. Defining response in systemic lupus erythematosus: a study by the Systemic Lupus International Collaborating Clinics group. J Rheumatol 2004;31:2390–4.
10. American College of Rheumatology Ad Hoc Committee on Systemic Lupus Erythematosus Response Criteria. The American College of Rheumatology response criteria for systemic lupus erythematosus clinical trials: measures of overall disease activity. Arthritis Rheum 2004;50:3418–26.
11. Furie RA, Petri MA, Wallace DJ, Ginzler EM, Merrill JT, Stohl W, et al. Novel evidence-based systemic lupus erythematosus responder index. Arthritis Care Res (Hoboken) 2009;61:1143–51.
12. Wallace DJ, Kalunian K, Petri MA, Strand V, Houssiau FA, Pike M, et al. Efficacy and safety of epratuzumab in patients with moderate/severe active systemic lupus erythematosus: results from

EMBLEM, a phase IIb, randomised, double-blind, placebo-controlled, multicentre study. Ann Rheum Dis 2014;73:183–90.

13. Furie R, Petri M, Zamani O, Cervera R, Wallace DJ, Tegzová D, et al. A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. Arthritis Rheum 2011;63:3918–30.
14. Navarra SV, Guzmán RM, Gallacher AE, Hall S, Levy RA, Jimenez RE, et al. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. Lancet 2011;377:721–31.
15. Furie RA, Morand EF, Bruce IN, Manzi S, Kalunian KC, Vital EM, et al. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. Lancet Rheumatol 2019;1:e208–19.
16. Van Vollenhoven RF, Hahn BH, Tsokos GC, Wagner CL, Lipsky P, Touma Z, et al. Efficacy and safety of ustekinumab, an IL-12 and IL-23 inhibitor, in patients with active systemic lupus erythematosus: results of a multicentre, double-blind, phase 2, randomised, controlled study. Lancet 2018;392:1330–9.
17. Clowse ME, Wallace DJ, Furie RA, Petri MA, Pike MC, Leszczyński P, et al. Efficacy and safety of epratuzumab in moderately to severely active systemic lupus erythematosus: results from two phase III randomized, double-blind, placebo-controlled trials. Arthritis Rheumatol 2017;69:362–75.
18. Morand EF, Furie R, Tanaka Y, Bruce IN, Askanase AD, Richez C, et al. Trial of anifrolumab in active systemic lupus erythematosus. N Engl J Med 2020;382:211–21.
19. National Institutes of Health. US National Library of Medicine: ClinicalTrials.gov. A 52-week, randomized, double-blind, parallel-group, placebo-controlled study to evaluate the efficacy and safety of a 200-mcg dose of IPP-201101 plus standard of care in patients with systemic lupus erythematosus: study results. 2019. URL: https://clinicaltrials.gov/ct2/show/results/NCT02504645?view=results.
20. Furie R, Khamashta M, Merrill JT, Werth VP, Kalunian K, Brohawn P, et al. Anifrolumab, an anti–interferon-α receptor monoclonal antibody, in moderate-to-severe systemic lupus erythematosus. Arthritis Rheumatol 2017;69:376–86.
21. Zimmer R, Scherbarth HR, Rillo OL, Gomez-Reino JJ, Muller S. Lupuzor/P140 peptide in patients with systemic lupus erythematosus: a randomised, double-blind, placebo-controlled phase IIb clinical trial. Ann Rheum Dis 2013;72:1830–5.
22. Merrill JT. For lupus trials, the answer might depend on the question. Lancet Rheumatology 2019;1:e196–7.
23. Yee CS, Farewell VT, Isenberg DA, Griffiths B, Teh LS, Bruce IN, et al. The use of Systemic Lupus Erythematosus Disease Activity Index-2K to define active disease and minimal clinically meaningful change based on data from a large cohort of systemic lupus erythematosus patients. Rheumatology (Oxford) 2011;50:982–8.
24. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459–68.
25. Yee CS, Gordon C, Isenberg DA, Griffiths B, Teh LS, Bruce IN, et al. The BILAG-2004 systems tally: a novel way of representing the BILAG-2004 index scores longitudinally. Rheumatology (Oxford) 2012;51:2099–105.
26. 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.
27. 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.
28. Hornik K. R FAQ; frequently asked questions on R. 2020. URL: https://cran.r-project.org/doc/FAQ/R-FAQ.html.
29. Williams RL. A note on robust variance estimation for cluster-correlated data. Biometrics 2000;56:645–6.
30. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry 1993;39:561–77.
31. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837.
32. Farewell VT, Tom BD, Royston P. The impact of dichotomization on the efficiency of testing for an interaction effect in exponential family models. J Am Stat Assoc 2004;99:822–31.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1631–1639
DOI 10.1002/acr.24644

American College of Rheumatology
Empowering Rheumatology Professionals

Pregnancy Outcomes in Undifferentiated Connective Tissue Disease Compared to Systemic Lupus Erythematosus: A Single Academic Center's Experience

Katherine P. Kaufman, Amanda M. Eudy, Nathaniel Harris, Laura Neil, and Megan E. B. Clowse

Objective. Systemic lupus erythematosus (SLE) patients have more pregnancy complications than healthy patients. Data regarding pregnancy outcomes in women with undifferentiated connective tissue disease (UCTD) are more limited, and existing studies are concentrated in Italy and predominantly in patients with a new diagnosis. Our objective was to compare pregnancy outcomes for UCTD and SLE patients with established disease.

Methods. Between 2008 and 2017, patients with UCTD and SLE at an academic medical center were recruited to a prospective pregnancy registry. UCTD was defined as a positive autoantibody plus connective tissue disease symptoms not meeting criteria for another rheumatic diagnosis. SLE was defined by American College of Rheumatology or Systemic Lupus International Collaborating Clinics classification criteria or by physician diagnosis. Data were collected throughout pregnancy and postpartum. Comparator groups included UCTD, low-activity SLE, and high-activity SLE.

Results. A total of 150 SLE and 51 UCTD pregnancies were analyzed. Disease activity was low in most patients, although more patients with SLE had severe activity during pregnancy (12% versus 2%; $P = 0.05$). The frequencies of prematurity and preeclampsia were significantly lower in UCTD than in high-activity SLE patients (preterm 17% versus 45% [$P = 0.004$] and preeclampsia 6% versus 34% [$P = 0.0008$]), although similar to low-activity SLE patients. More infants who were small for gestational age were born to SLE than UCTD patients (33% versus 7% [$P = 0.0005$]), regardless of disease activity level.

Conclusion. Pregnancies in women with UCTD managed by a rheumatologist have a high rate of pregnancy success and fewer risks than those in women with active SLE.

INTRODUCTION

When compared to healthy women, women with systemic lupus erythematosus (SLE) are known to have more complications in pregnancy, including higher rates of miscarriage, preterm delivery, preeclampsia, intrauterine growth restriction, and infants who are small for gestational age (SGA) (1–3). Women with SLE more often deliver by C-section or have labor induced (1,3,4). The risk of pregnancy complications tends to be higher in patients with active SLE than in patients with low disease activity (5,6).

Undifferentiated connective tissue disease (UCTD) is a term used to describe connective tissue diseases that do not meet criteria for another well-defined connective tissue disease (7). Definitions of UCTD vary but generally include the presence of symptoms of a connective tissue disease as well as a positive antinuclear antibody test (8–10). UCTD is typically a milder condition with a more benign course than SLE, so the outcomes studies of pregnancy in SLE may not be applicable to patients with UCTD.

Previous studies have evaluated pregnancy outcomes in patients with UCTD (9,11–14). Not all studies compared UCTD to other groups; however, the results of the comparative studies suggest that the risks of pregnancy complications in UCTD lie between those of healthy patients and those of patients with a

The Duke Autoimmunity in Pregnancy Registry was funded by Building Interdisciplinary Careers in Women's Health (grant K12-HD-043446), an Arthritis Foundation Arthritis Investigator Award, and the Duke Health Private Diagnostic Clinic ENABLE award. Dr. Eudy's work was supported by an NIH National Center for Advancing Translational Sciences award (grant 1KL2-TR-002554).

Katherine P. Kaufman, MD, Amanda M. Eudy, PhD, MD, Nathaniel Harris, MD, PhD, Laura Neil, BA, Megan E. B. Clowse, MD, MPH: Duke University, Durham, North Carolina.

Dr. Clowse has received consulting fees from AstraZeneca and UCB (less than $10,000 each) and has received medical education funding from GlaxoSmithKline. No other disclosures relevant to this article were reported.

Address correspondence to Megan E. B. Clowse, MD, Box 3535 Trent Drive, Durham, NC 27710. Email: megan.clowse@duke.edu.

Submitted for publication July 8, 2020; accepted in revised form May 4, 2021.

SIGNIFICANCE & INNOVATIONS

- This report is the first analysis of pregnancy outcomes of a cohort of women with undifferentiated connective tissue disease (UCTD) in the US.
- The findings here demonstrate that pregnancy outcomes of women with UCTD are similar to those of otherwise healthy women.
- The findings also demonstrate that women with well-controlled SLE at the start of pregnancy have good pregnancy outcomes.
- This study suggests that women with UCTD and women with well-controlled SLE can be counseled that they have a very high likelihood of having a healthy pregnancy and baby.

defined rheumatic condition (Table 1). All of these studies were done in Italy and thus represent a more racially homogeneous population, distinct from those seen in other parts of the world. Additionally, half of these studies evaluated outcomes in patients without a prior diagnosis of UCTD who, during pregnancy, screened positive for the diagnosis as part of the study (12–14). Whether these benign pregnancy outcomes would be found in women with symptoms significant enough to have sought rheumatologic care, to obtain this diagnosis, is not clear.

The objective of our study was to compare pregnancy outcomes of our cohort of patients with UCTD to pregnancy outcomes of patients with SLE. We hypothesized that the outcomes of UCTD patients would be more favorable than those of SLE patients, regardless of SLE activity in pregnancy. Our goal was to provide clinically useful information to women with UCTD who wish to pursue pregnancy.

PATIENTS AND METHODS

Patients and recruitment. Patients for this prospective cohort study were recruited into the Duke Autoimmunity in Pregnancy (DAP) registry (Duke University Hospital Institutional Review Board #Pro00000756) from the Duke rheumatology clinic by providing written informed consent at their first clinic visit in pregnancy. Patients were recruited from 2008 to 2017 and may be included in the analysis more than once for distinct pregnancies.

Patients were included if they had a singleton pregnancy, a diagnosis of either UCTD or SLE, and were enrolled into the DAP registry prior to 33 weeks. In our cohort, we diagnosed UCTD if patients had at least 1 positive autoantibody and symptoms suggestive of an autoimmune disease without meeting criteria for another rheumatic condition. SLE was defined according to American College of Rheumatology criteria, Systemic Lupus International Collaborating Clinics criteria, or physician diagnosis (15–17). Exclusion criteria included the patient having a primary diagnosis other than UCTD or SLE, including patients with a diagnosis of primary antiphospholipid syndrome (APS). Women with both SLE and APS were included in the study. Multiple gestation pregnancies (n = 9) are reported below but were excluded from statistical analysis of pregnancy and neonatal outcomes.

Data collection. Data collected at baseline included patient demographics and autoantibody profile. At each visit throughout pregnancy, data on medications, current lupus-related laboratory reports, and disease activity using the Systemic Lupus Erythematosus Pregnancy Disease Activity Index (18) and the physician global assessment (PhGA) were recorded at each visit. The maximum disease activity during pregnancy was used to classify patients into low, moderate, and severe activity groups. A PhGA of 0 to <1 was considered low activity, PhGA of 1 to <2 was moderate activity, and PhGA ≥2 was severe activity. High-activity SLE was defined as moderate or severe activity. Neonatal and maternal outcomes were collected via chart review and patient report soon after delivery. These outcomes included pregnancy loss for women enrolled prior to 20 weeks, preterm delivery (<37 weeks), early preterm birth (<30 weeks), weeks of gestational age at delivery, method of delivery, SGA (defined as birthweight <10th percentile for gestational age at birth, based on US population reference percentiles of birth weight, stratified by infant sex), very low birthweight (<1,500 gm), infant APGAR (appearance, pulse, grimace, activity, respiration) scores at 1 and 5 minutes, APGAR ≤6 at 5 minutes, neonatal intensive care unit (NICU) admissions, and complete heart block (19).

Statistical analysis. Outcomes were compared between pregnancies in women with UCTD versus SLE, as well as between pregnancies in which the mother had high-activity SLE (PhGA ≥1) during pregnancy, low-activity SLE during pregnancy, or UCTD. Due to significant differences in maternal race between women with SLE and UCTD and the known impact of non-Caucasian race on pregnancy outcomes, odds ratios (ORs) were adjusted by maternal race. Differences in proportions were estimated by Fisher's exact test. Differences in continuous variables were estimated by *t*-test or Wilcoxon's rank-sum test, depending on the distribution of the data. To account for the lack of independence in pregnancy outcomes for multiple pregnancies in the same woman, a sensitivity analysis was performed excluding all women with multiple pregnancies in the registry. Data were analyzed using SAS software, version 9.4.

RESULTS

Patients. Data were collected on a total of 201 pregnancies, including 150 pregnancies in 134 women with SLE and 51 pregnancies in 47 women with UCTD (Table 2). In patients with SLE, there were 8 twin pregnancies and 1 triplet pregnancy. There were significant differences in the racial composition of the

Table 1. Prior studies of pregnancies in women with undifferentiated connective tissue disease*

Author, year (ref.) [range], study design	Groups	Pregnancy loss	Mean GA at birth, weeks	Preterm	SGA or IUGR	Preeclampsia
Mosca, 2002 (11), cohort	UCTD (A), pregnant: n = 25 (20 patients)	12%	36.7	n = 3 (12%)	IUGR: n = 2 (twin gestation)	NR
Castellino, 2011 (9) [1/2003–12/2008], case–control†	UCTD (B), pregnant: n = 55 pregnancies (50 patients)	n = 3 (5.4%)	38.6	NR	IUGR: n = 2 (3.6%)	n = 2 (3.6%)
Spinillo, 2008 (12) [3/2004–6/2007], case–control	New diagnosis UCTD (C), pregnant: n = 41	NR	38.9‡	NR	SGA: 12/40 (30%)‡	n = 3 (7.5%)
	Healthy controls: n = 82	NR	39.6	NR	SGA: 11/80 (13.8%)	n = 1 (1.3%)
Spinillo, 2012 (13) [5/2005–4/2010], case–control	Defined CTD, pregnant: n = 24	n = 3 (12.5%)	37.8	n = 3 (12.5%)	IUGR: n = 3 (12.5%); SGA: n = 4 (16.67%)	n = 0/21
Spinillo, 2012 (13) [5/2005–4/2010], case–control	New diagnosis UCTD (D), pregnant: n = 62	n = 2 (3.23%)‡	38.7‡	Preterm, prior to 34 WGA; n = 2 (3.23%)‡	IUGR: n = 5 (8.06%)‡; SGA: n = 12 (19.35%)‡	n = 7/60 (11.3%)‡
Spinillo, 2012 (13) [5/2005–4/2010], case–control	Insufficient criteria for diagnosis, pregnant: n = 57	n = 2 (3.51%)	39.4	n = 0	IUGR: n = 0; SGA: n = 3 (5.26%)	n = 0/55
Spinillo, 2012 (13) [5/2005–4/2010], case–control	Asymptomatic controls, pregnant: n = 211	n = 2 (0.95%)	39.4	n = 1 (0.47%)	IUGR: n = 4 (1.90%); SGA: n = 19 (9.00%)	n = 3/209 (1.42%)
Spinillo, 2016 (14) [3/2009–6/2014], case–control	Defined CTD, pregnant: n = 68	n = 0	37.7	n = 5 (7.4%)	SGA: n = 18 (26.5%)	n = 20 (29.4%)
Spinillo, 2016 (14) [3/2009–6/2014], case–control	New diagnosis UCTD (E), pregnant: n = 131	n = 0	38.4‡	Preterm, prior to 34 WGA; n = 6 (4.6%)	SGA: 23 (17.5%)	n = 30 (22.9%)
Spinillo, 2016 (14) [3/2009–6/2014], case–control	Insufficient criteria for diagnosis, pregnant: n = 150	n = 1 (0.6%)	38.9	n = 22 (14.6%)	SGA: n = 17 (11.3%)	n = 9 (6%)
Spinillo, 2016 (14) [3/2009–6/2014], case–control	Asymptomatic controls: n = 597	n = 0	39	n = 17 (2.8%)	SGA: n = 46 (7.7%)	n = 19 (3.2%)
Zucchi, 2020 (25) [2000–2018], cohort	UCTD (A), pregnant: n = 100 (81 patients)	n = 11 (11%)	38.8	n = 9 (10%)	SGA: 10%	n = 1 (1.1%)
Radin, 2020 (24) [2010–2019], cohort	UCTD (F), pregnant: n = 224 (133 patients)	Sab: n = 45 (20.1%); SB: n = 2 (0.9%)	36.2	n = 30 (16.9%) late preterm: n = 22 (12.4%)	SGA: n = 21 (11.9%); IUGR: n = 6 (2.7%)	n = 5 (2.2%)

* (A) = signs, symptoms suggestive of connective tissue disease (CTD) lasting ≥1 year, not fulfilling criteria for a defined CTD plus positive antinuclear antibodies (ANAs), on 2 separate occasions; (B) = signs, symptoms suggestive of CTD lasting ≥1 year, not fulfilling criteria for a defined CTD plus positive ANAs, on 2 separate occasions >1:80; (C) = signs, symptoms suggestive of CTD lasting >6 months, not fulfilling criteria for defined CTD plus positive ANAs; (D) = signs, symptoms suggestive of CTD lasting ≥3 years, not fulfilling criteria for a defined CTD plus positive ANAs, on 2 separate occasions; (E) = signs, symptoms suggestive of CTD lasting ≥3 years, not fulfilling criteria for a defined CTD plus positive ANAs; (F) = undifferentiated connective tissue disease (UCTD) diagnosis plus positive ANAs. GA = gestational age; IUGR = intrauterine growth restriction; NR = not reported; ref. = reference; Sab = spontaneous abortion; SB = stillbirth (intrauterine death after 20 weeks gestation); SGA = small for gestational age; WGA = weeks gestational age.
† Control group = UCTD patients, not pregnant. Outcomes not reported here.
‡ Significant result.

diagnostic groups, with the majority of women with UCTD being Caucasian (67%) and half of women with SLE being Black (51%). The average maternal age at delivery and gestational age at study entry were not significantly different between women with UCTD and SLE.

SLE/UCTD features. The frequency of anti-Ro and anti-La antibodies was similar between women with UCTD and SLE, but anti-RNP and anti-Sm antibodies were both significantly more common in women with SLE. Antiphospholipid antibodies were not common within the cohort, and the frequency was not different for women with UCTD and SLE. Five pregnancies in 5 women with SLE met criteria for APS. In SLE patients, 28% had a history of lupus nephritis.

Disease activity. Most patients had low disease activity at the first clinic visit and throughout pregnancy. However, severe activity during pregnancy occurred in 2% of women with UCTD and

12% of women with SLE (P = 0.05). More patients with SLE than UCTD had abnormalities in complement and double-stranded DNA during pregnancy. Proteinuria prior to 20 weeks gestation was only found in patients with active SLE.

Medications. The large majority of women with UCTD and SLE took immunomodulatory medications during pregnancy, most commonly hydroxychloroquine (HCQ). Women with SLE were more likely to be prescribed azathioprine (none with UCTD compared to 20% with SLE; P = 0.0001) and glucocorticoids. Steroids of some form (oral, including dexamethasone, intramuscular, intraarticular, intravenous) were used by 47% of SLE patients and 18% of UCTD patients (P = 0.0002).

Table 2. Demographic and baseline characteristics for 150 pregnancies in 134 SLE patients and 51 pregnancies in 47 UCTD patients. Data include 141 singleton gestations, 8 twin gestations, and 1 triplet gestation in the SLE group. All pregnancies in the UCTD group were singleton*

Characteristic	UCTD	SLE	P†	High SLE	Low SLE
Total patients	47	134‡	–	58	81
Total pregnancies	51	150	–	61	89
Age, mean ± SD (range) years	31.4 ± 5.5 (17–42)	30.0 ± 5.5 (19–45)	0.1	28.9 ± 5.5 (20–42)	30.8 ± 5.4 (19–45)
Gestational age at entry, mean ± SD weeks	15.0 ± 7.7	12.6 ± 6.6	0.05	13.5 ± 6.2	12.0 ± 6.8
Race					
Black	10 (20)	76 (51)	0.0001	33 (54)	43 (48)
Caucasian	34 (67)	64 (43)	0.003	26 (43)	38 (43)
Asian	3 (6)	5 (3)	–	1 (2)	4 (4)
Other	3 (6)	5 (3)	–	1 (2)	4 (4)
Unknown	1(2)	0 (0)	–	0 (0)	0 (0)
Hispanic ethnicity (n = 196)	3 (6)	6 (4)	0.7	4 (7)	2 (2)
History of lupus nephritis	0 (0)	42 (28)	<0.0001	23 (38)	19 (21)
Autoantibodies (ever)					
Antinuclear antibody (n = 189)	44 (92)	139 (99)	0.04	54 (100)	85 (98)
Anti-Ro (n = 196)	18 (37)	75 (51)	0.1	25 (42)	50 (57)
Anti-La (n = 195)	9 (19)	23 (16)	0.7	12 (20)	11 (13)
Anti-RNP (n = 194)	5 (10)	64 (44)	<0.0001	29 (48)	35 (41)
Anti-Sm (n = 193)	0 (0)	47 (32)	<0.0001	25 (42)	22 (26)
aPL (ever) (n = 171)§	3 (8)	21 (16)	0.3	10 (18)	11 (14)
Antiphospholipid syndrome	0 (0)	5 (3)	0.3	3 (5)	2 (2)
Disease activity					
At first visit					
Low	45 (88)	110 (73)	0.1	21 (34)	89 (100)
Moderate	5 (10)	30 (20)	–	30 (49)	0 (0)
Severe	1 (2)	10 (7)	–	10 (16)	0 (0)
Highest during pregnancy					
Low	40 (78)	89 (59)	0.02	–	89 (100)
Moderate	10 (20)	43 (29)	–	43 (70)	–
Severe	1 (2)	18 (12)	–	18 (30)	–
Inflammatory markers during pregnancy					
Low C3 and/or C4 (n = 187)	2/42 (5)	50/145 (34)	<0.0001	26/59 (44)	24/86 (28)
dsDNA+ (n = 187)	0/42 (0)	35/145 (24)	<0.0001	14/59 (24)	21/86 (24)
Low platelets (n = 193)	2/45 (4)	6/148 (4)	1.0	5/60 (8)	1/88 (1)
Proteinuria (n = 192)	0/44 (0)	13/148 (9)	0.04	13/60 (22)	0/88 (0)
Proteinuria before 20 weeks (n = 155)	0/30 (0)	10/125 (8)	0.2	10/50 (20)	0/88 (0)
Medications					
Immunomodulatory					
Hydroxychloroquine	33 (65)	125 (83)	0.009	52 (85)	73 (82)
Azathioprine	0 (0)	30 (20)	0.0001	20 (33)	10 (11)
Steroids¶	9 (18)	70 (47)	0.0002	46 (75)	24 (27)
None	15 (29)	17 (11)	0.003	4 (7)	13 (15)
Antihypertensive	5 (10)	25 (17)	0.3	14 (23)	11 (12)
Aspirin	25 (49)	112 (75)	0.001	45 (74)	67 (75)
Anticoagulation	2 (4)	19 (13)	0.1	10 (16)	9 (10)

* Values are the number (%) unless indicated otherwise. aPL = antiphospholipid antibody; dsDNA = double-stranded DNA; SLE = systemic lupus erythematosus; UCTD = undifferentiated connective tissue disease.
† The P value compares UCTD and SLE pregnancies with statistical significance set at $P \leq 0.05$.
‡ 5 patients had both low and high SLE pregnancies.
§ Positive (≥40 IU) anticardiolipin IgG, IgM; positive (≥40 IU) anti-beta 2 glycoprotein IgG, IgM; positive lupus anticoagulant test.
¶ Oral (including dexamethasone), intramuscular, intravenous, or intraarticular steroids.

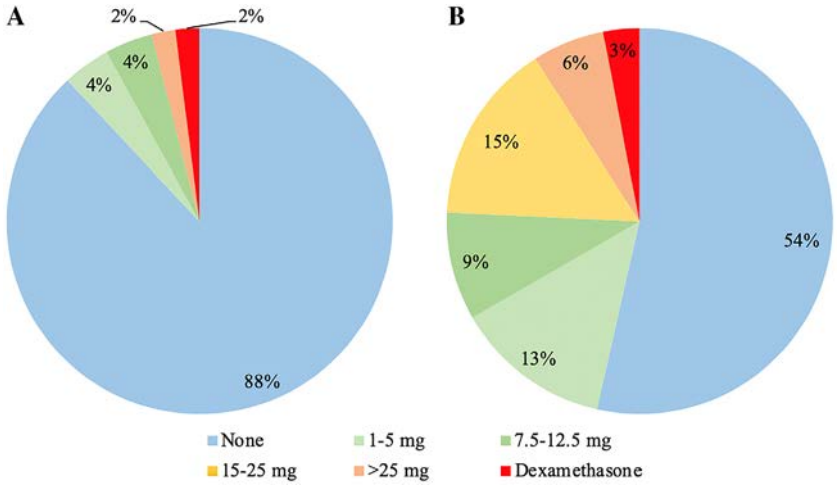

Figure 1. Maximum dose of oral steroids used in pregnancy among: **A**, undifferentiated connective tissue disease patients; and **B**, systemic lupus erythematosus patients. The majority of patients in both groups did not use oral steroids. A small percentage of patients in each group took dexamethasone as the only oral steroid for early neonatal heart block.

Table 3. Pregnancy and neonatal outcomes in singleton pregnancies*

Outcome	UCTD	All SLE	*P*†	High SLE	Low SLE
Total pregnancies, no.‡	51	141	–	59	82
Live birth	47 (92)	121 (86)	0.3	49 (83)	72 (88)
Losses§					
Miscarriage¶	2 (5)	12 (10)	0.5	2 (4)	10 (15)
Stillbirth#	1 (3)	3 (3)	1.0	3 (6)	0 (0)
Termination (medically indicated)	0 (0)	5 (4)	0.3	5 (10)	0 (0)
Termination (malformation)	1 (3)	0 (0)	0.2	0 (0)	0 (0)
Pregnancy outcomes**	n = 47	n = 121		n = 49	n = 72
Preterm (<37 weeks)	8 (17)	34 (28)	0.2	22 (45)	12 (17)
Early preterm (<30 weeks)	1 (2)	5 (4)	1.0	5 (10)	0 (0)
Weeks gestational age					
Mean ± SD	37.7 ± 2.5	36.8 ± 3.0	0.06	35.4 ± 3.9	37.8 ± 1.7
Range	27–40	24–40	–	24–40	32–40
Preeclampsia/eclampsia (n = 170)††	3 (6)	25 (20)	0.02	17 (34)	8 (11)
C-section (n = 167)	19 (40)	60 (50)	0.3	29 (60)	31 (43)
Induction (n = 167)	18 (38)	77 (64)	0.003	35 (73)	42 (58)
Medical induction‡‡	10 (56)	53 (69)	0.3	29 (83)	24 (57)
Induced for dates‡‡	8 (44)	24 (21)	0.3	6 (17)	18 (43)
Neonatal outcomes**					
SGA (n = 161)§§	3/45 (7)	38/116 (33)	0.0005	17/47 (36)	21/69 (30)
Very low birthweight (<1,500 gm)	1/45 (2)	8/116 (7)	0.4	6/47 (13)	2/69 (3)
APGAR 5 minutes, median (IQR) (n = 117)	9 (9–9)	9 (8.5–9)	1.0	9 (8–9)	9 (9–9)
APGAR 5 minutes ≤6	1/37 (3)	5/80 (6)	0.7	4/32 (13)	1/48 (2)
NICU admission (n = 131)	4/36 (11)	27/95 (28)	0.04	14 (37)	13 (23)
NICU length of stay, median (IQR) days	8 (2.5–29.5)	5 (2–12)	0.8	9 (4–16)	3 (2–5)

* Values are the number (%) unless indicated otherwise. APGAR = appearance, pulse, grimace, activity, respiration; IQR = interquartile range; NICU = neonatal intensive care unit; SGA = small for gestational age; SLE = systemic lupus erythematosus; UCTD = undifferentiated connective tissue disease.
† The *P* value compares UCTD and SLE pregnancies.
‡ Outcomes presented for singleton pregnancies only (n = 8 twin and n = 1 triplet pregnancy occurred in 9 women with SLE).
§ The rates of losses were calculated including only pregnancies enrolled prior to 20 weeks gestation. No losses occurred in pregnancies that enrolled after 20 weeks gestation.
¶ Miscarriage = fetal loss before 20 weeks gestation.
# Stillbirth = fetal loss at or after 20 weeks gestation.
** Pregnancy and neonatal outcomes are reported for live births.
†† Among pregnancies with a gestational age at outcome of at least 20 weeks.
‡‡ Percentage includes only those who were induced.
§§ Birthweight <0.1 percentile for gestational age.

Table 4. Pregnancy outcomes adjusted for maternal race*

Outcome	Unadjusted		Adjusted for race	
	High SLE vs. UCTD	Low SLE vs. UCTD	High SLE vs. UCTD	Low SLE vs. UCTD
Gestational age, β (95% CI)	−2.32 (−3.39, −1.25)	0.08 (−0.90, 1.06)	−2.30 (−3.39, −1.21)	0.09 (−0.91, 1.10)
Preterm birth	3.97 (1.54, 10.23)	0.98 (0.37, 2.60)	3.85 (1.41, 10.53)	0.92 (0.32, 2.64)
Preeclampsia	7.73 (2.09, 28.55)	1.88 (0.47, 7.46)	5.74 (1.50, 21.88)	1.33 (0.32, 5.53)
SGA	7.93 (2.13, 29.51)	6.13 (1.71, 22.00)	6.92 (1.82, 26.31)	5.36 (1.46, 19.69)
C-section	2.25 (0.99, 5.11)	1.11 (0.53, 2.35)	2.30 (0.98, 5.39)	1.14 (0.53, 2.49)

* Values are the odds ratio (95% confidence interval) unless indicated otherwise. SGA = small for gestational age; SLE = systemic lupus erythematosus; UCTD = undifferentiated connective tissue disease.

The frequency of glucocorticoid use was significantly higher in women with active SLE compared to those with low-activity SLE throughout pregnancy (high-activity SLE 75% versus low-activity SLE 27%; $P < 0.0001$). Chronic daily steroid use of ≤5 mg was uncommon and only prescribed in 4% of women with UCTD and 13% with SLE (Figure 1). Six patients, 5 with SLE (3%) and 1 with UCTD (2%), received dexamethasone to prevent the development of complete heart block in the fetus as their only steroid exposure. More women with UCTD than SLE were not prescribed immunomodulatory medications.

Low-dose aspirin, which is recommended to decrease the risk for preeclampsia among high-risk pregnancies, was more frequently prescribed to women with SLE (75%) than UCTD (49%; $P = 0.001$). Anticoagulation medication was prescribed to women with APS or prior thrombosis at similar rates for women with UCTD and SLE. The frequency of antihypertensive therapy was not different between patients with UCTD and SLE.

Outcomes. The majority of the 192 singleton pregnancies resulted in a live birth, with 86% of SLE and 92% of UCTD patients delivering a live infant ($P = 0.3$) (Table 3). Pregnancy termination for medical indications, most commonly rheumatic disease activity, was most common in women with high-activity SLE.

There was no statistically significant difference in the frequency of preterm delivery or early preterm delivery between SLE and UCTD patients. When adjusting for race and stratifying by the degree of SLE activity in pregnancy, however, we found that women with high-activity SLE had a higher frequency of preterm birth compared to women with UCTD (OR 3.8 [95% confidence interval (95% CI) 1.41, 10.53]) and delivered, on average, 2.3 weeks earlier (β −2.30 [95% CI −3.39, −1.21]) (Table 4).

Preeclampsia was more common in patients with SLE (20%) than in patients with UCTD (6%; $P = 0.02$). When adjusting for race, the odds of having preeclampsia were not significantly higher in patients with SLE compared to UCTD (OR 2.84 [95% CI 0.79, 10.26]), but they was significantly higher for SLE patients with high-activity lupus than for patients with UCTD (OR 5.74 [95% CI 1.50, 21.88]).

The difference in the C-section rate between SLE and UCTD patients was not significant, and this finding did not change when stratifying by disease activity level. While the frequency of induction was similar among women with UCTD and SLE overall, the rate in pregnancies in women with UCTD (40%) was lower than in the pregnancies in women with active SLE (73%; $P = 0.0009$).

Table 5. Outcomes of patients who were prescribed dexamethasone for early fetal cardiac changes*

Patient no.	Maternal diagnosis	Taking HCQ	Abnormal cardiac findings	Gestational weeks at abnormal finding	Outcome
1	UCTD	No	Borderline 1st degree heart block	23	PR interval normalized
2	SLE	Started at 16 WGA	Borderline 1st degree heart block	27	PR interval normalized
3	SLE	Yes	Mild MR, trivial TR	19	MR and TR resolved
4	SLE	Yes	MR, TR, ectopy	25	MR and TR improved to mild
5	SLE	Started at first consult (19 WGA)	Ectopy	19	Ectopy resolved (patient also stopped caffeine)
6	SLE	No	Complete heart block	19	Complete heart block

* HCQ = hydroxychloroquine; MR = mitral regurgitation; PR = EKG measurement; SLE = systemic lupus erythematosus; TR = tricuspid regurgitation; UCTD = undifferentiated connective tissue disease; WGA = weeks gestational age.

The reasons for induction were different, however, with 44% of pregnancies in women with UCTD induced routinely when they reached term, while 83% of pregnancies in women with high-activity lupus were medically induced, either for the health of the mother or the infant.

Significantly more infants born to mothers with SLE were SGA, but there was no difference in infants born weighing <1,500 gm. When adjusted for race, SLE continued to be associated with a higher risk of SGA infants, irrespective of disease activity, compared to UCTD (high SLE activity versus UCTD OR 6.92 [95% CI 1.82, 26.31] and low-activity SLE versus UCTD OR 5.36 [95% CI 1.46, 19.69]).

The frequency of NICU admission was significantly higher for pregnancies in women with SLE (28% versus 11%; $P = 0.04$). For both SLE and UCTD patients, the primary driver for admission to the NICU was preterm delivery, with at least half of babies delivered early spending some time in the NICU (UCTD: 50% of preterm versus 6% of term infants in the NICU [$P = 0.05$]; SLE 56% of preterm versus 18% of term infants in the NICU [$P = 0.0007$]).

There were 9 multiple gestation pregnancies, all in patients with SLE. Eight of the pregnancies were twin pregnancies, and 1 was a triplet pregnancy. Other than 1 twin stillbirth at 21 weeks, the remainder resulted in live births. The average gestational age for the twin pregnancies was 34 weeks, with 4 delivering at 36 weeks and 1 at 37 weeks. The triplet pregnancy delivered at 34 weeks. Six of the 8 live births (twin and triplet gestations) were delivered by C-section, 3 for maternal health, 2 for infant health (growth retardation in triplet pregnancy and nonreassuring fetal heart tones in a twin pregnancy), and 1 for maternal history of prior C-section. Of the 5 pregnancies for which preeclampsia outcome was known, 3 of the pregnancies were complicated by preeclampsia. Of the 5 pregnancies in women with SLE and APS, 2 resulted in terminations for medical reasons, 1 in a preterm delivery with preeclampsia at 35 weeks gestation, 1 with an SGA infant at term without preeclampsia, and 1 with a normal-sized infant at term without preeclampsia.

Only 1 infant born to a woman with anti-Ro antibodies developed complete heart block (1.1% of pregnancies in women with anti-Ro antibodies). The complete heart block was discovered at 19 weeks gestational age and was not reversed with dexamethasone; the mother was not taking HCQ. Five additional patients received dexamethasone for early neonatal cardiac complications related to the presence of an anti-Ro antibody (Table 5). Two patients with a fetal first-degree heart block and 2 patients with findings concerning for carditis had reversal of these abnormalities with dexamethasone. One fetus was noted to have ectopy, which resolved with dexamethasone administration and caffeine avoidance. Of the 5 patients with early fetal cardiac changes, 2 patients took HCQ throughout pregnancy, 2 patients started HCQ after entering the second trimester, and 1 patient did not take HCQ. Overall, 80% of women with anti-Ro antibody took HCQ during pregnancy (SLE 85%, UCTD 56%; $P = 0.009$). The frequency of cardiac changes suggestive of neonatal lupus was similar between pregnancies in women with SLE (6.6%) and UCTD (5.5%). The comparisons between UCTD and SLE, as well as high and low SLE, were not different when 9 pregnancies in 5 women with UCTD and 29 pregnancies in 14 women with SLE were excluded from the analysis.

DISCUSSION

Most pregnancies in women with UCTD resulted in a healthy baby and healthy mother, with similar outcomes to women with mild to inactive SLE but with fewer pregnancy and neonatal complications than pregnancies in women with active SLE. As has been shown previously, among pregnancies in women with SLE, most of the poor pregnancy and neonatal outcomes were related to SLE activity; given the low incidence of high-activity disease in women with UCTD, a similar association is not found in this cohort (1,5,20). In this cohort, collected prospectively in a university clinic in the southeastern part of the US, there were significantly more Black patients with SLE compared to patients with UCTD. Because extensive data have demonstrated more pregnancy complications in Black women than Caucasian women, pregnancy and neonatal outcomes were adjusted by maternal race, still revealing significant differences in rates of preterm birth, preeclampsia, and SGA between pregnancies with UCTD and high-activity SLE.

Pregnancy loss is difficult to assess in prospective pregnancy registries, as pregnancy losses that occur prior to a woman signing informed consent are not included. For this reason, the miscarriage rate in this report is likely an underestimate of the actual rate. Stillbirth, when assessed for pregnancies enrolled prior to 20 weeks gestation, is more reliable. The rate of stillbirth in this study is comparable to other SLE and UCTD studies, and among women with active SLE is higher than among the general population (21–24).

Our study is notably different from prior studies of pregnancies in women with UCTD by including only women with symptomatic UCTD that led to a rheumatologic diagnosis, by being located within the US, and by including a racially diverse population. Multiple prior studies have primarily included women who were diagnosed with UCTD during pregnancy through a screening survey administered in the obstetric clinic. As these women had not presented for rheumatologic evaluation, they probably had a lower level of rheumatic symptoms and, thus, less systemic inflammation and fewer pregnancy complications than our patients. We identified rates of preterm delivery, preeclampsia, and SGA that were all within the range of those reported in prior studies, although the rate of preterm birth in women with UCTD was 17% in this cohort, at the top of the range of prior UCTD pregnancy reports (range 3–17%) (13,14,24,25). The rate of preterm birth in women with SLE in this cohort is comparable to previously published reports as well (1,26,27).

The rate of SGA infants was significantly higher in patients with SLE compared to those with UCTD. Interestingly, this finding was irrespective of disease activity and persisted after adjusting for race. The frequency of SGA in this cohort of women with UCTD (7%) was within the range of prior reports (5–30%) and the 10% in the general population. Infants are most commonly born SGA due to placental insufficiency, which has been associated with SLE pregnancies (28). In a comparison of placental histology between women with SLE and healthy women, the placentas from women with SLE, with and without preeclampsia, had abnormalities resembling those found in preeclampsia without SLE, including maternal vasculitis, laminar decidual necrosis, maternal-fetal interface hemorrhage, and nonocclusive fetal thrombotic vasculopathy (29). The low frequency of SGA in the UCTD population in this study suggests that the placental pathology seen in women with SLE may not be common in women with UCTD.

The infant outcomes for women with UCTD in this cohort were very reassuring, with a low rate of NICU admission, normal APGAR scores, and no increase in SGA rate. The primary cause for NICU admission was preterm delivery, as is seen in the general population.

The rate of complete heart block in our cohort of pregnancies in women with anti-Ro antibodies was as expected, with only 1 infant (1.1%) developing this complication. The majority of patients in both groups were taking HCQ. Four retrospective cohort studies and 1 prospective trial show that HCQ may reduce the risk of developing complete heart block, so the high rate of HCQ use in our cohort may be responsible for keeping the number of cases with complete heart block low (30–34). The use of dexamethasone in the patients with early signs of heart block may have prevented those fetuses from having progression of heart block, though the literature on fluorinated steroids (dexamethasone, betamethasone) to prevent or reverse early heart block due to anti-Ro antibodies is conflicting without randomized studies available (35).

The primary limitation of our study is the small number of pregnancies in the UCTD cohort and the absence of a healthy control group for comparison. Additionally, these data are derived from a single center and managed by an expert in reproductive rheumatology, which limits generalizability.

The findings reported here allow us to counsel women with UCTD that pregnancy is frequently safe, with low rates of pregnancy and neonatal complications. More than half of the women with UCTD in this cohort were maintained on medications throughout pregnancy, including HCQ and aspirin, which may have decreased their risk for poor pregnancy outcomes. As this was not a randomized trial, the impact of the medications on pregnancy outcomes is unclear; this study does not suggest that women with UCTD without rheumatologic management will have similar outcomes. Women with SLE had higher rates of pregnancy complications than women with UCTD, with particularly high rates of preterm birth and preeclampsia in women with moderate to severely active SLE during pregnancy. An important finding is that our patients with low-activity SLE had similar pregnancy outcomes to those of UCTD patients, with the exception of an elevated rate of SGA. This study suggests that women with UCTD, as well as women with well-controlled SLE, can be reassured that they have a very high likelihood of having a healthy pregnancy and baby.

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. Clowse 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. Kaufman, Eudy, Clowse.
Acquisition of data. Kaufman, Eudy, Harris, Neil, Clowse.
Analysis and interpretation of data. Kaufman, Eudy, Clowse.

REFERENCES

1. Petri M. Hopkins Lupus Pregnancy Center: 1987 to 1996. Rheum Dis Clin North Am 1997;23:1–13.
2. Lateef A, Petri M. Systemic lupus erythematosus and pregnancy. Rheum Dis Clin North Am 2017;43:215–26.
3. Pastore D, Costa M, Parpinelli M, Surita F. A critical review on obstetric follow-up of women affected by systemic lupus erythematosus. Rev Bras Ginecol Obstet 2018;40:209–24.
4. Eudy AM, Jayasundara M, Haroun T, Neil L, James AH, Clowse ME. Reasons for cesarean and medically indicated deliveries in pregnancies in women with systemic lupus erythematosus. Lupus 2018;27:351–6.
5. Clowse ME, Magder L, Petri M. Cyclophosphamide for lupus during pregnancy. Lupus 2005;14:593–7.
6. Buyon JP, Kim MY, Guerra MM, Laskin CA, Petri M, Lockshin MD, et al. Predictors of pregnancy outcomes in patients with lupus: a cohort study. Ann Intern Med 2015;163:153–63.
7. Kallenberg CG. Overlapping syndromes, undifferentiated connective tissue disease, and other fibrosing conditions. Curr Opin Rheumatol 1992;4:837–42.
8. Mosca M, Neri R. Undifferentiated connective diseases (UCTD): a review of the literature and a proposed preliminary classification criteria. Clin Exp Rheumatol 1999;17:615–20.
9. Castellino G, Capucci R, Bernardi S, Padovan M, Giacuzzo S, Pivato E, et al. Pregnancy in patients with undifferentiated connective tissue disease: a prospective case-control study. Lupus 2011;20:1305–11.
10. Mosca M, Tani C, Vagnani S, Carli L, Bombardieri S. The diagnosis and classification of undifferentiated connective tissue diseases. J Autoimmun 2014;48-49:50–2.
11. Mosca M, Neri R, Strigini F, Carmignani A, Totti D, Tavoni A, et al. Pregnancy outcome in patients with undifferentiated connective tissue disease: a preliminary study on 25 pregnancies. Lupus 2002;11:304–7.
12. Spinillo A, Beneventi F, Epis OM, Montanari L, Mammoliti D, Ramoni V, et al. The effect of newly diagnosed undifferentiated connective tissue disease on pregnancy outcome. Am J Obstet Gynecol 2008;199:1–6.
13. Spinillo A, Beneventi F, Ramoni V, Caporali R, Locatelli E, Simonetta M, et al. Prevalence and significance of previously

undiagnosed rheumatic diseases in pregnancy. Ann Rheum Dis 2012; 71:918–23.

14. Spinillo A, Beneventi F, Locatelli E, Ramoni V, Caporali R, Alpini C, et al. The impact of unrecognized autoimmune rheumatic diseases on the incidence of preeclampsia and fetal growth restriction: a longitudinal cohort study. BMC Pregnancy Childbirth 2016;16:1–8.
15. 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.
16. 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.
17. Petri M, Orbai AM, Alarcón 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.
18. Buyon JP, Kalunian KC, Ramsey-Goldman R, Petri MA, Lockshin MD, Ruiz-Irastorza G, et al. Assessing disease activity in SLE patients during pregnancy. Lupus 1999;8:677–84.
19. Oken E, Kleinman KP, Rich-Edwards J, Gillman MW. A nearly continuous measure of birth weight for gestational age using a United States national reference. BMC Pediatr 2003;3:1–10.
20. Ko HS, Ahn HY, Jang DG, Choi SK, Park YG, Park IY, et al. Pregnancy outcomes and appropriate timing of pregnancy in 183 pregnancies in Korean patients with SLE. Int J Med Sci 2011;8:577–83.
21. MacDorman MF, Gregory EC. Fetal and perinatal mortality: United States, 2013. Natl Vital Stat Rep 2015;64:1–24.
22. Clowse ME, Magder LS, Witter F, Petri M. Early risk factors for pregnancy loss in lupus. Obstet Gynecol 2006;107:293–9.
23. Smyth A, Oliveira GH, Lahr BD, Bailey KR, Norby SM, Garovic VD. A systematic review and meta-analysis of pregnancy outcomes in patients with systemic lupus erythematosus and lupus nephritis. Clin J Am Soc Nephrol 2010;5:2060–8.
24. Radin M, Schreiber K, Cecchi I, Bortoluzzi A, Crisafulli F, de Freitas CM, et al. A multicentre study of 244 pregnancies in undifferentiated connective tissue disease: maternal/fetal outcomes and disease evolution. Rheumatology (Oxford) 2020;59:2412–8.
25. Zucchi D, Tani C, Monacci F, Elefante E, Carli L, Parma A, et al. Pregnancy and undifferentiated connective tissue disease: outcome and risk of flare in 100 pregnancies. Rheumatology (Oxford) 2020;59: 1335–9.
26. Chakravarty EF, Colón I, Langen ES, Nix DA, El-Sayed YY, Genovese MC, et al. Factors that predict prematurity and preeclampsia in pregnancies that are complicated by systemic lupus erythematosus. Am J Obstet Gynecol 2005;192:1897–904.
27. Kroese SJ, Abheiden CN, Blomjous BS, van Laar JM, Derksen RW, Bultink IE, et al. Maternal and perinatal outcome in women with systemic lupus erythematosus: a retrospective bicenter cohort study. J Immunol Res 2017;2017:1–9.
28. Hanly JG, Gladman DD, Rose TH, Laskin CA, Urowitz MB. Lupus pregnancy: a prospective study of placental changes. Arthritis Rheum 1988;31:358–66.
29. Marder W, Knight JS, Kaplan MJ, Somers EC, Zhang X, O'Dell AA, et al. Placental histology and neutrophil extracellular traps in lupus and pre-eclampsia pregnancies. Lupus Sci Med 2016;3:1–9.
30. Izmirly PM, Kim MY, Llanos C, Le PU, Guerra MM, Askanase AD, et al. Evaluation of the risk of anti-SSA/Ro-SSB/La antibody-associated cardiac manifestations of neonatal lupus in fetuses of mothers with systemic lupus erythematosus exposed to hydroxychloroquine. Ann Rheum Dis 2010;69:1827–30.
31. Izmirly PM, Costedoat-Chalumeau N, Pisoni CN, Khamashta MA, Kim MY, Saxena A, et al. Maternal use of hydroxychloroquine is associated with a reduced risk of recurrent anti-SSA/Ro-antibody-associated cardiac manifestations of neonatal lupus. Circulation 2012;126:76–82.
32. Tunks RD, Clowse ME, Miller SG, Brancazio LR, Barker PC. Maternal autoantibody levels in congenital heart block and potential prophylaxis with antiinflammatory agents. Am J Obstet Gynecol 2013;208:64. e1–7.
33. Martínez-Sánchez N, Pérez-Pinto S, Robles-Marhuenda Á, Arnalich-Fernández F, Cameán MM, Zalvide EH, et al. Obstetric and perinatal outcome in anti-Ro/SSA-positive pregnant women: a prospective cohort study. Immunol Res 2017;65:487–94.
34. Izmirly P, Kim M, Costedoat-Chalumeau N, Friedman D, Saxena A, Copel J, et al. The Prospective Open Label Preventive Approach to Congenital Heart Block with Hydroxychloroquine (PATCH) study demonstrates a reduction in the recurrence rate of advanced block [abstract]. Arthritis Rheumatol 2019;71 Suppl 10. URL: https://acrabstracts.org/abstract/the-prospective-open-label-preventive-approach-to-congenital-heart-block-with-hydroxychloroquine-patch-study-demonstrates-a-reduction-in-the-recurrence-rate-of-advanced-block/.
35. Saxena A, Izmirly PM, Mendez B, Buyon JP, Friedman DM. Prevention and treatment in utero of autoimmune associated congenital heart block. Cardiol Rev 2014;22:263–7.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1640–1648
DOI 10.1002/acr.24612

Prevalence and Associated Factors of Electrocardiogram Abnormalities in Patients With Systemic Lupus Erythematosus: A Machine Learning Study

Zhuoran Hu,[1] Lin Wu,[1] Zhiming Lin,[1] Xiuhua Liu,[2] Changlin Zhao,[1] and Zhen Wu[1]

Objective. Electrocardiogram (EKG) abnormalities are predictive of subsequent cardiovascular events. Cardiac involvement is common in systemic lupus erythematosus (SLE). We aimed to determine the prevalence of EKG abnormalities in SLE patients and to examine the factors associated with EKG abnormalities with machine learning approaches.

Methods. Consecutive SLE patients' records were retrieved from the database of the hospital for the cross-sectional study. Abnormal EKGs with clinical significance were grouped by the presence of tachyarrhythmias, atrioventricular block, nonspecific ST segment changes, T wave abnormalities, ventricular hypertrophy, axis deviation, bundle branch block, and QT interval prolongation. Associated factors of the most common EKG abnormalities were assessed by comparing logistic regression and 4 other machine learning approaches.

Results. In the present study, 299 patients were enrolled, with 128 showing clinically significant abnormalities on EKG. T wave changes (52.3%), nonspecific ST segment–T wave (ST-T) changes (26.6%), and prolonged QT interval (8.6%) were the most prevalent abnormalities among patients with abnormal findings on EKG. Random forests models had the best performance in the discovery of associated factors. Age, disease duration, antinuclear antibody titer, disease activity (as measured by the Systemic Lupus Erythematosus Disease Activity Index 2000) were associated with nonspecific ST-T changes, prolonged QT interval, and T wave changes. Hypertension, positivity for anti-SSA antibodies, and secondary Sjögren's syndrome were influential factors for nonspecific ST-T changes, prolonged QT interval, and T wave changes, specifically.

Conclusion. ST-T and T wave changes were the most common abnormalities seen on EKGs of SLE patients. Our finding suggests that age, longer disease duration, higher disease activity, hypertension, anti-SSA antibody positivity, and secondary Sjögren's syndrome are important and influential factors in these EKG abnormalities.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic autoimmune disorder that may involve multiple organs (1). Cardiac involvement, representing one of the most important clinical manifestations of SLE, contributes to an increased risk of hospitalization, morbidity, and mortality, particularly in patients with later onset of disease activity (2–4). Atherosclerosis can occur prematurely in patients with SLE and is independent of traditional risk factors for cardiovascular disease (5). Therefore, it is essential for health care workers to identify cardiac involvement in SLE early on its disease course.

Given that resting electrocardiogram (EKG) is an inexpensive and noninvasive test independent of the operator and the patient's condition, it is performed routinely to detect cardiac abnormalities with important prognostic implications. Chou et al found that resting EKG abnormalities, particularly ST segment and/or T wave abnormalities, left ventricular hypertrophy, left axis deviation, left bundle branch block (BBB), and right BBB, were associated with subsequent cardiovascular events (i.e., sudden

[1]Zhuoran Hu, MM, Lin Wu, MD, Zhiming Lin, MD, Changlin Zhao, MD, Zhen Wu, MD: Third Affiliated Hospital of Sun Yat-sen University, Guangzhou City, China; [2]Xiuhua Liu, MM: Liuzhou Worker's Hospital/Fourth Affiliated Hospital of Guangxi Medical University, Liuzhou City, China.

Ms. Hu and Dr. Wu contributed equally to this work.

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

Address correspondence to Zhen Wu, MD, Division of Cardiology, Third Affiliated Hospital of Sun Yat-sen University, No. 600, Tianhe Road, Tianhe District, Guangzhou City, 51000, China. Email: wuzh@mail.sysu.edu.cn; or to Changlin Zhao, MD, Division of Cardiology, Third Affiliated Hospital of Sun Yat-sen University, No. 600, Tianhe Road, Tianhe District, Guangzhou City, 51000, China. Email: zhaoclzssy@126.com.

Submitted for publication March 2, 2021; accepted in revised form March 25, 2021.

SIGNIFICANCE & INNOVATIONS

- Patients with systemic lupus erythematosus (SLE) are vulnerable to cardiovascular diseases. Electrocardiogram (EKG) is a feasible method to detect cardiac involvement. Previous studies have reported an association between EKG abnormalities and subsequent cardiovascular events.
- To our knowledge, the present study is the first to investigate the prevalence of EKG abnormalities in SLE patients and the associated factors with machine learning approaches.
- If further studies support the findings of the present study, it might be possible to identify those at higher risk of developing EKG abnormalities earlier and to take appropriate action to prevent it.

cardiovascular death, nonfatal myocardial infarction) (6). Rates and risk factors of these significant abnormalities have been studied in some White SLE cohorts (7–10). In a Chinese cohort (11), however, relative studies are rare and involve limited EKG abnormalities and insufficient potential associated factors (sex, age, disease duration, and autoantibodies).

Moreover, conventional logistic regression has been the main method adopted by most previous studies. Although logistic regression is a fair process, it should be used under certain assumptions, such as no multicollinearity among variables. Machine learning is a robust approach that worked without the above assumption in logistic regression. Therefore, the present single-center, interdisciplinary study was conducted to explore the prevalence and potential relative factors of resting EKG abnormalities among SLE patients by using both logistic regression and machine learning methods.

PATIENTS AND METHODS

Patient selection. A cross-sectional study was conducted. Consecutive SLE patients' records were retrieved from the database of the Division of Rheumatology of the Third Affiliated Hospital of Sun Yat-sen University from August 2018 to November 2019. Eligible patients were ages ≥18 years, had a diagnosis of SLE based on the 1997 update of the SLE revised criteria as defined by the American College of Rheumatology (ACR) (12), and had available EKG records at the time they were in-patient during the above-mentioned period. Only one record would be collected if patients were hospitalized repeatedly. Patients were excluded from study participation if they had experienced/were experiencing the following: 1) a cardiovascular disease (CVD)–related event (myocardial infarction, angina, congestive heart failure, angioplasty, and pacemaker placement) or severe valvular disease prior to EKG; 2) hepatic and/or renal failure, overlapping syndrome, and/or abnormal serum electrolytes; 3) pregnancy at the time of study enrollment; 4) malignant tumors at the time of study enrollment; and 5) thyroid disorders, such as Grave's disease.

Ethical approval and consent to participate. Ethical approval of the present study was obtained from the Ethics Committee of the Third Affiliated Hospital of Sun Yat-sen University. The study was conducted in accordance with the Declaration of Helsinki. All patients provided informed consent for the collection and use of clinical and laboratory data. The study registration number was [2018] 02-283-01.

Study setting and data collection. A standard protocol (variables are shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24612) containing demographic characteristics, lifestyle information, complications, assessment of disease activity (as measured by the SLE Disease Activity Index 2000 [SLEDAI-2K] [13]), medication history, and laboratory tests, along with standard digitally available recorded 12-lead resting EKG reports, was used to collect data. Computer-assisted reading of EKGs was rechecked by 2 experienced cardiologists.

Age was categorized by decades. Patients were grouped into 3 categories according to body mass index (BMI; kg/m^2) (14): normal, underweight, and overweight or obese. Patients were considered to have a smoking habit if they were documented as a current smoker at any point in the 3 months preceding the date of EKG testing.

Secondary Sjögren's syndrome and lupus nephritis were diagnosed according to the 2016 ACR/European Alliance of Associations for Rheumatology criteria classification criteria for primary Sjögren's syndrome (15) or biopsy of the patients' kidneys and labial glands, respectively. Urine protein:creatinine ratio was calculated as urine total protein (mg)/urine creatinine (gm). When urine protein:creatinine ratio was >100, it was defined as a urine protein:creatinine ratio elevation. Patients were considered positive for antiphospholipid antibodies (aPLs) if they had positive results for 1) total anticardiolipin antibodies (aCLs) and individual aCLs (aCL IgG/aCL IgM), 2) lupus anticoagulant (LAC), or 3) individual β_2-glycoprotein I (β_2-GPI IgG/β_2-GPI IgM) (16).

Data on type 2 diabetes mellitus (diagnosed according to 1999 World Health Organization criteria), dyslipidemia (defined according to 2016 Chinese guidelines for the management of dyslipidemia in adults proposed by the National Expert Committee of guideline revision [17]), hypertension (systolic blood pressure reading of ≥140 mm Hg and/or a diastolic blood pressure reading of ≥90 mm Hg, or current administration of an antihypertensive drug), and hyperuricemia (diagnosed as having a serum uric acid [UA] level of ≥416 mmoles/liter for men and postmenopausal women or serum UA level of ≥357 mmoles/liter for premenopausal women) were extracted. Other candidate variables

are also shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24612.

Study design. EKGs were categorized into the normal group and the abnormal group. Abnormal EKGs were further grouped into "without clinical significance" and "with clinical significance" categories, with the "without clinical significance" group comprising EKGs that showed sinus tachycardia (heart rate of ≥100 beats per minute on EKG), first degree atrioventricular (AV) block (PR interval of ≥200 milliseconds on EKG), and sinus bradycardia (heart rate of ≤60 beats per minute on EKG). The "with clinical significance" group comprising EKGs with any of the following abnormalities: 1) tachyarrhythmias, including atrial fibrillation, atrial flutter, atrial tachycardia, and premature ventricular extrasystole; 2) second-to-third-degree AV block; 3) nonspecific ST segment and/or T wave abnormalities; 4) left ventricular hypertrophy; 5) left axis deviation; 6) left BBB; and 7) right BBB; and 8) QT interval prolongation. QT interval prolongation was defined as a corrected QT interval of ≥450 milliseconds (18). When more than one abnormality was observed in the same patient, each one was recorded separately.

Sample size. A systematic sampling design was used to select the study participants. Sample sizes were estimated using PASS 15 software (https://www.ncss.com), with the statistical power (1-β) set at 0.90, type I error (α) set at 0.05, assuming that the prevalence of EKG abnormalities was 30% (19) among SLE patients. The software calculated that a total sample size of at least 283 would suffice. Finally, we recruited 299 patients for the present study.

Missing data. There were respectively 9.8% and 9.6% missing data in "urine protein:creatinine ratio" and "antiphospholipid antibodies." Proportions of missing values were less than 5% across all other variables. Multiple imputations were implemented using the multivariate imputation by chained equations algorithm in R package "mice" (20) to account for missing data to minimize bias and precision reduction.

Statistical analysis. Statistical data were presented as the mean ± SE for continuous variables with normal distribution, median (interquartile range [IQR]) for those without normal distribution, and percentage for categorical variables. A 2-tailed *P* value of <0.05 was considered to indicate statistical significance.

Machine learning was performed, and methods included random forests algorithms (21), support vector machine (SVM) (22), least absolute shrinkage and selection operator (23), and multivariate adaptive regression spline (24). To assess the performance of each modeling approach, we randomly split the sample of SLE participants into training set and validation set at a ratio of 7:3 c-normality and abnormalities by synthetic minority oversampling technique, in order to reduce the negative effect of the class imbalance in the constructed models (25). Four commonly used machine learning methods and conventional logistic regression with a forward stepwise selection of variables were used to train models for assessing the influential factors of the 3 most common clinically significant EKG abnormalities (nonspecific ST segment–T wave [ST-T] changes, prolonged QT interval, and T wave changes). R software (version 3.6.1; R Core Team) was used to conduct statistical analysis, with the RandomForest package (26) for random forests models, glmnet package (27) for least absolute shrinkage and selection operator models, e1071

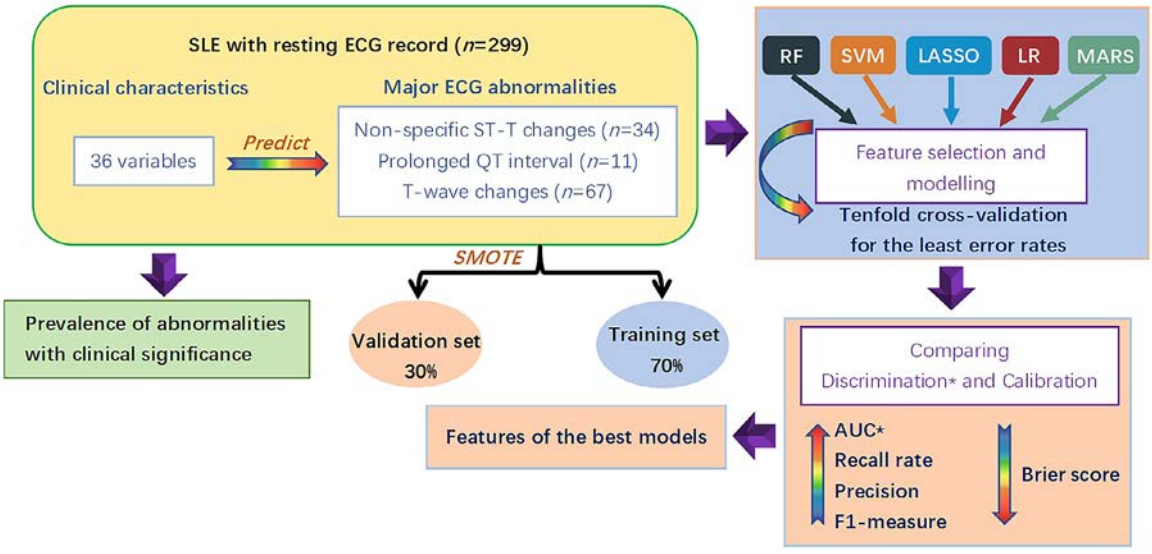

Figure 1. Flow diagram of the study. AUC = area under curve; ECG = electrocardiogram; LASSO = least absolute shrinkage and selection operator; LR = logistic regression; MARS = multivariate adaptive regression splines; RF = random forests; SLE = systemic lupus erythematosus; SMOTE = synthetic minority over-sampling technique; SVM = support vector machines. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24612/abstract.

Table 1. Demographic and clinical characteristics of the study patients*

Characteristics	Value
Female sex	269 (89.9)
Age, mean ± SD years	37.6 ± 15.3
Age groups, years	
≤30	107 (35.8)
31–40	69 (23.1)
41–50	54 (18.1)
>50	69 (23.1)
BMI, mean ± SD kg/m^2	21.6 ± 3.3
BMI groups	
Normal	201 (67.2)
Underweight	42 (14.0)
Overweight or obese	56 (18.7)
Disease duration, median (IQR)	2.0 (0.6, 5.0)
Disease duration	
<3 months	24 (8.0)
3–6 months	36 (12.0)
6 months to 1 year	75 (25.1)
1–3 years	65 (21.7)
>3 years	99 (33.1)
SLEDAI-2K, median (IQR)	8.0 (4.0, 14.0)
Disease severity	
No activity (0–4)	91 (30.4)
Mild activity (5–9)	69 (23.1)
Moderate activity (10–14)	70 (23.4)
Persistent activity and flare (≥15)	69 (23.1)
Lupus nephritis	32 (10.4)
Hypoproteinemia	76 (25.4)
Secondary Sjögren's syndrome	41 (13.7)
Photosensitization	16 (5.4)
Raynaud's phenomenon	28 (9.4)
Anemia, Coomb's test (negative)	113 (37.8)
CRP, median (IQR)	2.9 (0.9, 12.8)
ESR, median (IQR)	36.0 (16.0, 68.5)
Elevated markers of inflammation	
CRP	101 (33. 8)
ESR	191 (63.9)
ANA titer	
1:80	10 (3.3)
1:100	67 (22.4)
1:320	66 (22.1)
1:640	11 (3.7)
1:1,000	48 (16.1)
1:1,280	29 (9.7)
1:3,200	68 (22.7)
Urine protein:creatinine ratio elevation	80 (26.8)
Antibody positivity	
Anti-dsDNA	117 (39.1)
Anti-Sm	81 (27.1)
Anti-SSA/Ro	132 (44.1)
Anti-SSB/La	39 (13.0)
Anti-U1RNP	123 (41.1)
Anti-C1q	178 (59.5)
Anti-nucleosome	41 (13.7)
Anti-histone	90 (30.1)
ANCA	52 (17.4)
Antiphospholipid antibodies	17 (5.7)
Coomb's test	11 (3.7)
Long-term medications†	
Glucocorticoids	97 (32.4)
Hydroxychloroquine	112 (37.5)
Cyclophosphamide	16 (5.4)
Cyclosporine	10 (3.3)
Mycophenolate mofetil	38 (12.7)
Smoking, yes	17 (5.7)
Drinking, yes	10 (3.3)
Comorbidity	
Hypertension	42 (14.0)
Hyperuricemia	99 (33.1)
Dyslipidemia	109 (40.5)
Type 2 diabetes mellitus	12 (4.0)

* Unless indicated otherwise, values shown are the number (%) of study patients. ANA = antinuclear antibody; ANCA = antineutrophil cytoplasmic antibody; anti-dsDNA = anti–double-stranded DNA; BMI = body mass index; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; IQR = interquartile range; SLEDAI-2K = Systemic Lupus Erythematosus Disease Activity Index 2000.
† Long-term medication use was defined as a patient receiving medication for at least 3 months before undergoing an electrocardiogram.

package (28) for SVM models, and earth package (29) for multivariate adaptive regression spline; base generalized linear model function was used for logistic regression.

Evaluation and comparison of the prediction accuracy and receiver operating characteristic (ROC) area under the curve (AUC) of different models were performed. Precision, accuracy, recall rate, and F1 score were used to evaluate model performance. All candidate variables were input simultaneously into 4 machine learning approaches for the model. The flow diagrom of the study is shown in Figure 1, and Supplementary Methods are available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24612.

RESULTS

Patient characteristics. A total of 299 SLE patients participated in this study, and patient characteristics are shown in Table 1. Female patients dominated this study (89.9%), and 35.8% of participants were age ≤30 years (mean ± SD age 37.6 ± 15.3 years). More than half (53.6%) of patients had no or mild disease activity.

Prevalence of EKG abnormalities. Among all abnormalities observed on EKG, 128 (42.8%) of 299 SLE patients were found to have clinically significant abnormalities. Abnormalities without clinical significance, including sinus bradycardia (n = 10), sinus tachycardia (n = 22), and sinus irregularity (n = 9), were excluded from further analysis.

As shown in Figure 2, among all EKG abnormalities, abnormal ventricular repolarization, observed in 118 (39.5%) of 299 study participants, was the most common finding, followed by arrhythmia 15 (5%) and conduction abnormalities 8 (2.7%). In patients with abnormal EKG diagnosis (n = 128), T wave changes occurred in 67 (52.3%) of 128 patients, followed by nonspecific ST-T changes, which was found in 34 (26.6%) of 128 patients.

A prolonged QT interval was present in 11 (8.6%) of 128 patients. U wave elevation and pathologic Q wave were found in 4 (3.1%) and 2 (1.6%) of 128 patients, respectively. Frequent premature ventricular contractions (PVC) accounted for half of the patients with arrhythmia (7 of 15). Moreover, atrial fibrillation and frequent atrial premature beats were both found in 2 patients, and atrial flutter and occasional atrial premature beats only occurred in 1 patient. In patients with conduction abnormalities, shortened PR intervals, right BBB, second-degree AV block were found in 4 (3.1%), 3 (2.4%), and 1 (0.08%) of 128 patients, respectively.

The detection rate of axis deviation and rotation, including left axis deviation (n = 6), was the most detected among study patients, and was doubled for right axis deviation (n = 3) and tripled for both clockwise and counterclockwise rotation (n = 2). Only 1 patient had complications with preexcitation syndrome. Globally, 8 of all patients had more than 1 major clinically significant abnormality (i.e., prolonged QT intervals, nonspecific ST-T changes, and T wave changes).

Models' performance. Considering prevalence and clinical significance, we chose the 3 most common clinically significant EKG abnormalities, including nonspecific ST-T changes, prolonged QT intervals, and T wave changes, as outcomes to develop predictive models.

The models constructed by the 4 machine learning algorithms and logistic regression in the validating group were compared. Least absolute shrinkage and selection operator had the highest AUCs in the prediction of nonspecific ST-T changes (0.89) and prolonged QT intervals (0.97), respectively, and random forests modeling had the highest AUC in the prediction of T wave changes (0.93). ROC AUC analysis of the validating set and training set is shown in Supplementary Figures 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24612.

As shown in Supplementary Table 2 (http://onlinelibrary.wiley.com/doi/10.1002/acr.24612), the Brier scores of random forests models for 3 abnormalities were the lowest (0.14, 0.05, and 0.07, respectively). Moreover, recall rate, precision, accuracy, and F1-measure were also ranked the highest in random forests modeling of all abnormalities, except for the recall rate of nonspecific ST-T changes, which random forests modeling was close to the highest. Therefore, we chose random forests as the final predictive model.

Associated factors for EKG abnormalities. As shown in Figure 3, age, disease duration, disease severity, and ANA titer were consistently the 5 most influential factors affecting nonspecific ST-T changes, prolonged QT intervals, or T wave

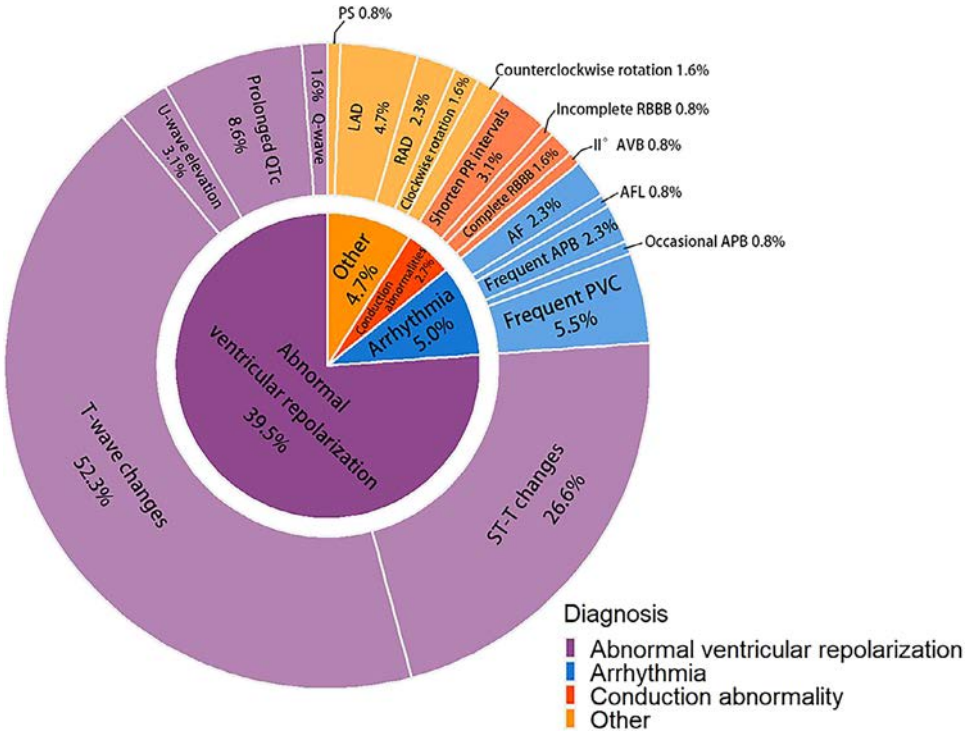

Figure 2. Resting electrocardiogram (EKG) abnormalities with clinical significance among inpatients with systemic lupus erythematosus (SLE). Inner ring is the proportion in each group of all in-patients with SLE (n = 299), and outer ring is the proportion in each subgroup of the abnormal findings on EKG (n = 128). AF = atrial fibrillation; AFL = atrial flutter; APB = atrial premature beats; AVB = atrioventricular block; LAD = left axis deviation; PS = preexcitation syndrome; PVC = premature ventricular contractions; QTc = QT interval; RAD = right axis deviation; RBBB = right bundle branch block; ST-T = nonspecific ST segment–T wave.

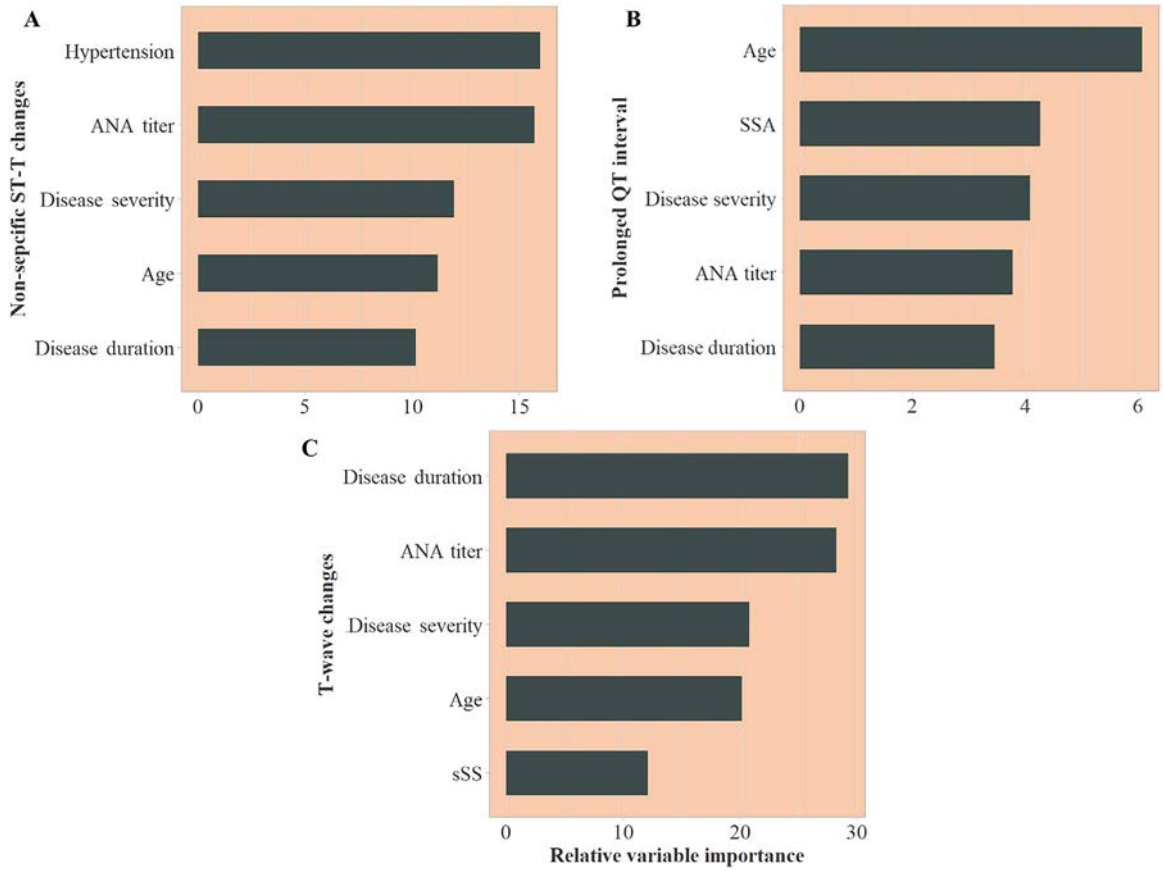

Figure 3. Top 5 important variables selected by random forests (RF) models. **A**, Nonspecific ST segment–T wave (ST-T) changes. **B**, Prolonged QT interval. **C**, T wave changes. ANA = antinuclear antibody; sSS: secondary Sjögren's syndrome. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24612/abstract.

changes—though their ranks differed among these abnormalities. Noticeably, hypertension was the most important associated factor for nonspecific ST-T changes; anti-SSA antibody positivity was also found to be related to the presence of a prolonged QT interval. Secondary Sjögren's syndrome was demonstrated to be relative to T wave changes. The 15 most important variables selected by random forests models were shown in Supplementary Figure 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24612.

DISCUSSION

Among the SLE cohort, the most common EKG abnormalities observed in SLE patients were abnormal ventricular repolarization (39.5%), which mostly consisted of T wave changes, nonspecific ST-T changes, and prolonged QT interval. No published study has revealed the prevalence of T wave changes individually. Only the prevalence of ST segment changes and/or T wave changes in all SLE patients have been reported previously (9), and this number was lower than that shown in the present study (15.0% versus 22.4%).

However, nonspecific ST-T changes in the present cohort (11.3%) were lower than those in other reported studies (30.9% to 44%) (10). Several factors can explain this difference. In the present study, the study population had a much shorter disease duration and higher disease activity levels (as measured by the SLEDAI-2K) than previous study cohorts, even without significant age differences between participants in the present study and those from previous study cohorts.

In the standard surface EKG, nonspecific ST-T changes are common findings even in patients without SLE. Previous studies have indicated that nonspecific ST-T abnormalities are significantly associated with cardiovascular morbidity and mortality (30–32). However, many studies did not separate T wave changes from nonspecific ST-T changes. In the present study, we found that T wave changes were more prevalent than nonspecific ST-T changes. Moreover, T wave abnormalities included amplitude and angle changes and were reported to serve as early markers of repolarization abnormalities in a hypertensive population (33). Although the clinical impact of nonspecific ST-T changes in patients who have SLE without CVD is unclear (10), it is enticing to speculate that these changes represent subclinical CVD.

QT interval prolongation is also an independent cardiovascular risk factor (34,35) and is mainly related to cardiac arrest. The proportion of SLE patients with complicated QT interval

prolongation in the present cohort was lower than that of the study carried out by Bourré-Tessier et al (3.7% versus 7.3%, respectively) (36). It should be noted that in the cohort studied by Bourré-Tessier et al, 38% of patients had positivity for anti-Ro/SSA antibodies. In the present study, anti-SSA antibody positivity was also found to be an influential factor for prolonged QT interval, similar to the survey performed by Bourré-Tessier and colleagues (36). Lazzerini et al (37) revealed an association between prolonged QT interval and anti-SSA/Ro antibodies. Their follow-up work showed that in patients with connective tissue diseases, QT interval prolongation is correlated with only 1 of the subtypes of anti-SSA/Ro antibodies (52-kD), which represent a clinically silent novel risk factor for torsades de pointes arrhythmia development via an autoimmune-mediated electrophysiologic interference with the human ERG potassium channel (38). Notably, characterized by anti-SSA/Ro and/or anti-SSB/La positivity, Sjögren's syndrome has also been found to be an important factor impacting T wave abnormalities. Thus, anti-SSA antibody positivity is assumed to impair cardiac ventricular repolarization.

The prevalence of other manifestations, such as atrial fibrillation, was lower in the present study than that of the study performed by Myung et al (2 of 299 versus 7 of 235, respectively) (7). This is probably due to the fact that the study population of the present work is younger than the cohort in Myung and colleagues' study (mean ± SD 38 ± 15 years versus 52 ± 15 years, respectively). Furthermore, the presence of hypertension and thyroid diseases, which are well-documented irritant causes of atrial fibrillation (39,40), was not assessed in the study by Myung et al. We excluded those with thyroid diseases in the present study population and included hypertension and diabetes mellitus as potential influential factors. Additionally, the results of the present work demonstrated that hypertension was strongly associated with nonspecific ST-T changes, which was consistent with a previous study (9).

Age, disease duration, ANA titer, and disease activity (SLEDAI-2K) were consistently important factors for 3 major EKG abnormalities in the study cohort. As disease duration is related to age, this factor has been reported to increase the risk of nonspecific ST-T changes by 5.8% in the general population, with an odds ratio [OR] of 1.058 (95% confidence interval [95% CI] 1.052–1.064) previously reported in a large-scale Chinese population study by Xiao et al (41). Another Chinese population–based study indicated that increasing age (after 35 years old) is a strong risk factor for QT interval prolongation (OR 1.228 [95% CI 1.168–1.290]) (42).

ANA titer was correlated with SLEDAI-2K score in the present cohort ($r_s = 0.204$, $P < 0.001$). Thus, ANA titer and level of disease activity could be important factors given the potential role of inflammatory activity in affecting cardiac rhythm and conduction. A systemic inflammatory response would probably evoke nonspecific ST-T changes seen, for instance, in pancreatitis (43). It has also been indicated that QT dispersion, a parameter of interlead variability in the QT interval, was significantly higher in SLE patients with high disease activity than in patients with mild-to-moderate disease activity (44).

Overall, machine learning analysis can be used when multicollinearity was found among variables. Therefore, random forests analysis could be used to explore and rank the relative importance of interactive variables with accuracy and efficiency. For instance, a consideration in future analyses could be the fact that disease duration may depend on the age of participating subjects. Another strength of the present study is the comparison between several machine learning approaches since each approach has its own unique strengths and weaknesses. As a more affordable cardiovascular examination, EKG may be recommended for all SLE patients in their first clinical visit as well as during follow-up visits because a multisystemic and chronic disease such as SLE needs interdisciplinary cooperation to assess the condition holistically and longitudinally. Assisted by the machine learning method, some objective information might be gained before inviting a cardiologist's consultations and further, expensive or intrusive examinations.

There were several limitations to the present study that can impact its generalizability to other populations as well as the interpretation of its clinical significance. First, the sample size was insufficient. Second, we only used resting 12-lead EKGs rather than 24-hour EKG monitoring (Holter), which can measure diurnal variations of EKG intervals, and this may have caused a higher "false-negative rate" when we diagnosed EKGs. Third, we cannot exclude the possibility of patient selection bias because only a single tertiary referral center participated in this study. Therefore, the prevalence of EKG abnormalities in the present study cannot represent the real rate of EKG abnormalities in China. Moreover, although EKG is an affordable, stable, and quick test that causes no harm to the recipient, the information offered by EKG testing is limited. Other cardiac imaging examinations such as echocardiogram, cardiac magnetic resonance imaging, or radionuclide perfusion could provide more details about cardiac lesions. Further longitudinal, prospective studies assessing the role of potential risk factors will help clarify the mechanism of EKG abnormalities among SLE patients.

The present data reveals that ST-T changes/T wave changes were the most common abnormalities seen in EKGs carried out in SLE patients. In addition, QT prolongation was also present. Our findings suggest that age, longer disease duration, higher disease activity, hypertension, anti-SSA antibody positivity, and secondary Sjögren's syndrome were important and influential factors for these EKG abnormalities. Efforts to monitor the EKGs of these key populations need to be instituted.

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. Zhen Wu had full access to all data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Lin, Zhao, Zhen Wu.
Acquisition of data. Hu, Lin Wu, Liu, Zhen Wu.
Analysis and interpretation of data. Hu, Lin.

REFERENCES

1. Tsokos GC. Systemic lupus erythematosus. N Engl J Med 2011;365: 2110–21.
2. Miner JJ, Kim AH. Cardiac manifestations of systemic lupus erythematosus. Rheum Dis Clin North Am 2014;40:51–60.
3. Bartels CM, Buhr KA, Goldberg JW, Bell CL, Visekruna M, Nekkanti S, et al. Mortality and cardiovascular burden of systemic lupus erythematosus in a US population-based cohort. J Rheumatol 2014;41:680–7.
4. Dhital R, Pandey RK, Poudel DR, Oladunjoye O, Paudel P, Karmacharya P. All-cause hospitalizations and mortality in systemic lupus erythematosus in the US: results from a national inpatient database. Rheumatol Int 2020;40:393–7.
5. Roman MJ, Shanker BA, Davis A, Lockshin MD, Sammaritano L, Simantov R, et al. Prevalence and correlates of accelerated atherosclerosis in systemic lupus erythematosus. N Engl J Med 2003;349: 2399–406.
6. Chou R, Arora B, Dana T, Fu R, Walker M, Humphrey L. Screening asymptomatic adults with resting or exercise electrocardiography: a review of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2011;155:375–85.
7. Myung G, Forbess LJ, Ishimori ML, Chugh S, Walalce D, Weisman MH. Prevalence of resting-ECG abnormalities in systemic lupus erythematosus: a single-center experience. Clin Rheumatol 2017;36: 1311–16.
8. Hosonuma M, Yajima N, Takahashi R, Yanai R, Matsuyama T, Toyosaki E, et al. Fragmented QRS complex in patients with systemic lupus erythematosus at the time of diagnosis and its relationship with disease activity. PloS One 2020;15:e0227022.
9. Al Rayes H, Harvey PJ, Gladman DD, Su J, Sabapathy A, Urowitz MB, et al. Prevalence and associated factors of resting electrocardiogram abnormalities among systemic lupus erythematosus patients without cardiovascular disease. Arthritis Res Ther 2017;19:31.
10. Geraldino-Pardilla L, Gartshteyn Y, Piña P, Cerrone M, Giles JT, Zartoshi A, et al. ECG non-specific ST-T and QTc abnormalities in patients with systemic lupus erythematosus compared with rheumatoid arthritis. Lupus Sci Med 2016;3:e000168.
11. Jia E, Geng H, Liu Q, Xiao Y, Zhang Y, Xie J, et al. Cardiac manifestations of Han Chinese patients with systemic lupus erythematosus: a retrospective study. Ir J Med Sci 2019;188:801–6.
12. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40:1725.
13. Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol 2002;29:288–91.
14. Wildman RP, Gu D, Reynolds K, Duan X, He J. Appropriate body mass index and waist circumference cutoffs for categorization of overweight and central adiposity among Chinese adults. Am J Clin Nutr 2004;80:1129–36.
15. 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.
16. Miyakis S, Lockshin MD, Atsumi T, Branch DW, Brey RL, Cervera R, et al. International consensus statement on an update of the classification criteria for definite antiphospholipid syndrome (APS). J Thromb Haemost 2006;4:295–306.
17. Joint Committee for Guideline Revision National Expert Committee on Cardiovascular Diseases, National Center for Cardiovascular Diseases, Chinese Society of Cardiology, Chinese Medical Association Chinese Diabetes Society, Chinese Medical Association, Chinese Society of Endocrinology, Chinese Medical Association, Chinese Society of laboratory Medicine, Chinese Medical Association. 2016 Chinese guidelines for the management of dyslipidemia in adults. J Geriatr Cardiol 2018;15:1–29.
18. Soliman EZ, Howard G, Cushman M, Kissela B, Kleindorfer D, Le A, et al. Prolongation of QTc and risk of stroke: The REGARDS (REasons for Geographic and Racial Differences in Stroke) study. J Am Coll Cardiol 2012;59:1460–7.
19. Bourré-Tessier J, Urowitz MB, Clarke AE, Bernatsky S, Krantz MJ, Huynh T, et al. Electrocardiographic findings in systemic lupus erythematosus: data from an international inception cohort. Arthritis Care Res (Hoboken) 2015;67:128–35.
20. Van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;2011;45:67.
21. Breiman L. Random forests, machine learning. J Clin Microbiol 2001; 2:199–228.
22. Noble WS. What is a support vector machine? Nat Biotechnol 2006; 24:1565–7.
23. Steyerberg EW, Eijkemans MJ, Harrell FE Jr, Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med 2000;19: 1059–79.
24. Friedman JH, Roosen CB. An introduction to multivariate adaptive regression splines. Stat Methods Med Res 1995;4:197–217.
25. Blagus R, Lusa L. Joint use of over- and under-sampling techniques and cross-validation for the development and assessment of prediction models. BMC Bioinformatics 2015;16:363.
26. Liaw A, Wiener M. Classification and regression by randomForest. Forest 2001;23.
27. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33: 1–22.
28. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) 2011;2:1–27.
29. Friedman JH. Multivariate adaptive regression splines. Ann Statist 1991;19:1–67.
30. Badheka AO, Rathod A, Marzouka GR, Patel N, Bokhari SS, Moscucci M, et al. Isolated non-specific ST-segment and T-wave abnormalities in a cross-sectional United States population and Mortality (from NHANES III). Am J Cardiol 2012;110:521–5.
31. Walsh JA III, Prineas R, Soliman EZ, Liu K, Ning H, Daviglus ML, et al. Association of isolated minor non-specific ST-segment and T-wave abnormalities with subclinical atherosclerosis in a middle-aged, biracial population: Coronary Artery Risk Development in Young Adults (CARDIA) study. Eur J Prev Cardiol 2013;20: 1035–41.
32. Kumar A, Lloyd-Jones DM. Clinical significance of minor non-specific ST-segment and T-wave abnormalities in asymptomatic subjects: a systematic review. Cardiol Rev 2007;15:133–42.
33. Dilaveris P, Gialafos E, Poloniecki J, et al. Changes of the T-wave amplitude and angle: an early marker of altered ventricular repolarization in hypertension. Clin Cardiol 2000;23:600–6.
34. Algra A, Tijssen JG, Roelandt JR, Pool J, Lubsen J. QTc prolongation measured by standard 12-lead electrocardiography is an independent risk factor for sudden death due to cardiac arrest. Circulation 1991;83:1888–94.

35. Robbins J, Nelson JC, Rautaharju PM, Gottdiener JS. The association between the length of the QT interval and mortality in the Cardiovascular Health Study. Am J Med 2003;115:689–94.
36. Bourré-Tessier J, Clarke AE, Huynh T, Bernatsky S, Joseph L, Belisle P, et al. Prolonged corrected QT interval in anti-Ro/SSA-positive adults with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2011;63:1031–7.
37. Lazzerini PE, Capecchi PL, Guideri F, Bellisai F, Selvi E, Acampa M, et al. Comparison of frequency of complex ventricular arrhythmias in patients with positive versus negative anti-Ro/SSA and connective tissue disease. Am J Cardiol 2007;100:1029–34.
38. Lazzerini PE, Yue Y, Srivastava U, Fabris F, Capecchi PL, Bertolozzi I, et al. Arrhythmogenicity of anti-Ro/SSA antibodies in patients with torsades de pointes. Circ Arrhythm Electrophysiol 2016;9:e003419.
39. Aronow WS, Fleg JL, Pepine CJ, Artinian NT, Bakris G, Brown AS, et al. ACCF/AHA 2011 expert consensus document on hypertension in the elderly: a report of the American College of Cardiology Foundation Task Force on Clinical Expert Consensus documents developed in collaboration with the American Academy of Neurology, American Geriatrics Society, American Society for Preventive Cardiology, American Society of Hypertension, American Society of Nephrology, Association of Black Cardiologists, and European Society of Hypertension. J Am Coll Cardiol 2011;57:2037–114.
40. Bruere H, Fauchier L, Bernard Brunet A, Pierre B, Simeon E, Babuty D, et al. History of thyroid disorders in relation to clinical outcomes in atrial fibrillation. Am J Med 2015;128:30–7.
41. Xiao L, Bai T, Zeng J, Yang R, Yang L. Nonalcoholic fatty liver disease, a potential risk factor of non-specific ST-T segment changes: data from a cross-sectional study. PeerJ 2020;8:e9090.
42. Ma Q, Li Z, Guo X, Guo L, Yu S, Yang H, et al. Prevalence and risk factors of prolonged corrected QT interval in general Chinese population. BMC Cardiovasc Disord 2019;19:276.
43. Redkar NN, Rawat KJ, Tiwari D, et al. Case series of pancreatitis with uncommon presentations. J Assoc Physicians India 2015;63:53–5.
44. Kojuri J, Nazarinia MA, Ghahartars M, Mahmoody Y, Rezaian GR, Liaghat L. QT dispersion in patients with systemic lupus erythematosus: the impact of disease activity. BMC Cardiovasc Disord 2012; 12:11.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1649–1658
DOI 10.1002/acr.24619

American College of Rheumatology
Empowering Rheumatology Professionals

Identifying Potential Classification Criteria for Calcium Pyrophosphate Deposition Disease: Item Generation and Item Reduction

Sara K. Tedeschi,[1] Tristan Pascart,[2] Augustin Latourte,[3] Cattleya Godsave,[4] Burak Kundakci,[4] Raymond P. Naden,[†] William J. Taylor,[5] Nicola Dalbeth,[6] Tuhina Neogi,[7] Fernando Perez-Ruiz,[8] Ann Rosenthal,[9] Fabio Becce,[10] Eliseo Pascual,[11] Mariano Andres,[11] Thomas Bardin,[3] Michael Doherty,[4] Hang-Korng Ea,[3] Georgios Filippou,[12] John FitzGerald,[13] Marwin Guitierrez,[14] Annamaria Iagnocco,[15] Tim L. Jansen,[16] Minna J. Kohler,[17] Frédéric Lioté,[3] Mark Matza,[17] Geraldine M. McCarthy,[18] Roberta Ramonda,[19] Anthony M. Reginato,[20] Pascal Richette,[3] Jasvinder A. Singh,[21] Francisca Sivera,[22] Alexander So,[10] Lisa K. Stamp,[23] Janeth Yinh,[17] Chio Yokose,[17] Robert Terkeltaub,[24] Hyon Choi,[17] and Abhishek Abhishek[4]

Objective. Classification criteria for calcium pyrophosphate deposition (CPPD) disease will facilitate clinical research on this common crystalline arthritis. Our objective was to report on the first 2 phases of a 4-phase process for developing CPPD classification criteria.

Methods. CPPD classification criteria development is overseen by a 12-member steering committee. Item generation (phase I) included a scoping literature review of 5 literature databases and contributions from a 35-member combined expert committee and 2 patient research partners. Item reduction and refinement (phase II) involved a combined expert committee meeting, discussions among clinical, imaging, and laboratory advisory groups, and an item-rating exercise to assess the influence of individual items toward classification. The steering committee reviewed the modal rating score for each item (range –3 [strongly pushes away from CPPD] to +3 [strongly pushes toward CPPD]) to determine items to retain for future phases of criteria development.

Results. Item generation yielded 420 items (312 from the literature, 108 from experts/patients). The advisory groups eliminated items that they agreed were unlikely to distinguish between CPPD and other forms of arthritis, yielding 127 items for the item-rating exercise. Fifty-six items, most of which had a modal rating of +/– 2 or 3, were retained for future phases. As numerous imaging items were rated +3, the steering committee recommended focusing on imaging of the knee and wrist and 1 additional affected joint for calcification suggestive of CPP crystal deposition.

Conclusion. A data- and expert-driven process is underway to develop CPPD classification criteria. Candidate items comprise clinical, imaging, and laboratory features.

INTRODUCTION

Calcium pyrophosphate deposition (CPPD) disease represents a symptomatic crystalline arthritis that affects an estimated 8–10 million adults in the US (1,2). Chondrocalcinosis, a radiographic finding that has been used to estimate CPPD prevalence, is present in approximately 10% of adults in Italy and the US age ≥65 years; CPPD accounts for a similar number of

Supported by the American College of Rheumatology and the European Alliance of Associations for Rheumatology. Dr. Tedeschi's work was supported by the NIH (grants K23-AR-075070 and L30-AR-070514). Dr. Neogi's work was supported by the NIH (grant K24-AR-070892). Dr. Perez-Ruiz's work was supported by the Cruces Rheumatology Association. Dr. Yokose's work was supported by the Rheumatology Research Foundation Scientist Development Award and the NIH (grant T32-AR-007258). Dr. Terkeltaub's work was supported by the Veterans Administration Research Service and the NIH (grants AR-060772 and AR-075990).

[1]Sara K. Tedeschi, MD, MPH: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [2]Tristan Pascart, MD, PhD: Lille Catholic University, Lille, France; [3]Augustin Latourte, MD, PhD, Thomas Bardin, MD, Hang-Korng Ea, MD, PhD, Frédéric Lioté, MD, PhD, Pascal Richette, MD, PhD: Hôpital Lariboisière, Université de Paris, Paris, France; [4]Cattleya Godsave, MD, Burak Kundakci, PT, MSc, Michael Doherty, MD, Abhishek Abhishek, MBBS, PhD: University of Nottingham, Nottingham, UK; [5]William J. Taylor, MBChB, PhD: University of Otago, Wellington, New Zealand; [6]Nicola Dalbeth, MD: University of Auckland, Auckland, New Zealand; [7]Tuhina Neogi, MD, PhD: Boston University School of Medicine, Boston, Massachusetts; [8]Fernando Perez-Ruiz, MD, PhD: Osakidetza, Ezkerraldea-Enkarterri-Cruces, Cruces University Hospital, Biocruces-Bizkaia Health Research Institute, and University of the Basque Country, Spain; [9]Ann Rosenthal, MD: Medical College of Wisconsin, Milwaukee; [10]Fabio Becce, MD, Alexander So, MD, PhD: Lausanne University Hospital, Lausanne, Switzerland; [11]Eliseo Pascual, MD, PhD, Mariano Andres, MD, PhD: Hospital General Universitario de Alicante and Alicante Institute of Sanitary and Biomedical Research, Alicante, Spain; [12]Georgios Filippou, MD, PhD: Luigi Sacco University Hospital, Milan, Italy; [13]John FitzGerald, MD, PhD: David Geffen School of Medicine, University of California, Los Angeles; [14]Marwin Guitierrez, MD: Instituto Nacional de Rehabilitación, Mexico City, Mexico; [15]Annamaria Iagnocco, MD: Università

SIGNIFICANCE & INNOVATIONS

- The calcium pyrophosphate deposition disease (CPPD) Classification Criteria Working Group is using established methodology that includes data-driven and expert-driven methods to develop validated CPPD classification criteria.
- This report describes the findings of phases I and II, item generation and item reduction, for the CPPD classification criteria project.
- The 56 candidate items retained in phase II are broadly categorized as clinical, imaging, and laboratory features of CPPD.

hospitalizations in France as does gout (3–5). CPPD was first recognized in the 1960s, yet clinical research on this disease has lagged far behind other common arthritides. This crystalline arthritis presents with a host of manifestations, including acute CPP crystal inflammatory arthritis, chronic CPP crystal inflammatory arthritis, and osteoarthritis with CPPD, and a patient may have >1 manifestation over time or simultaneously (6). Targeted treatments for CPP crystal deposition do not currently exist, and many patients with CPPD experience inadequately treated joint pain, swelling, and stiffness (7). Advances in CPP-related arthritis, including treatment trials, have been hampered by lack of validated classification criteria, a framework that has facilitated clinical research and trials in other rheumatic diseases (8–11).

The case definition of CPPD in research studies has varied, which poses a major setback to advancing knowledge about CPPD and to developing therapeutics. Diagnostic criteria for CPPD proposed in the 1970s require evidence of chondrocalcinosis on plain radiographs and CPP crystals on synovial fluid crystal analysis but have not been validated for use in clinical research and have several shortcomings (12,13). Plain radiography and synovial fluid crystal analysis each have limitations in their ability to detect CPP crystals, including low-to-moderate sensitivity of plain radiography, underutilization of arthrocentesis, interobserver differences in identifying CPP crystals via polarized light microscopy, and certain qualities of CPP crystals that pose challenges to identifying them via compensated polarized light microscopy (14–16). Sensitivity of synovial fluid CPP crystal identification has been variable in the literature, raising concerns that relying on this test to define CPPD for research could exclude a large percentage of subjects (17–19).

Classification criteria for a disease allow investigators to identify relatively homogenous populations with that disease for inclusion in clinical research (20,21). The European Alliance of Associations for Rheumatology (EULAR) generated working terminology and diagnostic recommendations in 2011, but formal classification criteria for CPPD do not exist (2). The American College of Rheumatology (ACR) and EULAR are jointly sponsoring multiphase development of CPPD classification criteria. The criteria system will aim to achieve high sensitivity while maximizing specificity for CPPD, ensuring that future study populations achieve a degree of homogeneity and represent a consensus definition of CPPD for research, reflecting the most common manifestations of CPPD. Classification criteria for CPPD will need to encompass the spectrum of symptomatic manifestations of this arthritis, though future investigators may choose to limit their study population to patients with particular types of CPPD manifestations.

The CPPD Classification Criteria Working Group is using established methodology that includes data-driven and expert-driven methods previously employed for developing classification criteria for gout, rheumatoid arthritis, systemic lupus erythematosus, systemic sclerosis, and other rheumatic diseases (8–11,22,23). This report describes the findings of phases I and II, item generation and item reduction, for the CPPD classification criteria project.

MATERIALS AND METHODS

CPPD classification criteria are being developed through a 4-phase process. The process began by generating a

degli Studi di Torino, Turin, Italy; [16]Tim L. Jansen, MD: VieCuri Medical Center, Venlo, Noord-Limburg, and University of Twente, Enschede, The Netherlands; [17]Minna J. Kohler, MD, Mark Matza, MD, Janeth Yinh, MD, Chio Yokose, MD, Hyon Choi, MD, PhD: Massachusetts General Hospital and Harvard Medical School, Boston; [18]Geraldine M. McCarthy, MD: Mater Misericordiae University Hospital, Dublin, Ireland; [19]Roberta Ramonda, MD, PhD: University of Padova, Padova, Italy; [20]Anthony M. Reginato, MD, PhD: Brown University, Providence, Rhode Island; [21]Jasvinder A. Singh, MD, MPH: University of Alabama at Birmingham and Birmingham Veterans Affairs Medical Center; [22]Francisca Sivera, MD, PhD: Hospital General Universitario Elda, Elda, and Universidad Miguel Hernandez, Elche, Spain; [23]Lisa K. Stamp, MD: University of Otago, Christchurch, New Zealand; [24]Robert Terkeltaub, MD: San Diego Veterans Administration Healthcare Service and University of California-San Diego.

†Dr. Naden is deceased.

Dr. Tedeschi has received consulting fees from NGM Biopharmaceuticals (less than $10,000). Dr. Latourt has received consulting fees from Novartis (less than $10,000). Dr. Dalbeth has received consulting and/or speaking fees from AstraZeneca, Dyve, Selecta, Horizon, Arthrosi, Cello Health, AbbVie, and Janssen (less than $10,000 each). Dr. Becce has received consulting fees from Horizon Therapeutics (less than $10,000) and has a research agreement with Siemens Healthineers. Dr. Jansen has received consulting and/or speaking fees from AbbVie, Amgen, Bristol Myers Squibb, Grunenthal, Olatec, and Sanofi Genzyme (less than $10,000 each). Dr. Kohler has received consulting and/or speaking fees from Novartis and Eli Lilly (less than $10,000 each). Dr. McCarthy has received consulting fees from PK Med (less than $10,000). Dr. Singh has received consulting fees from Crealta/Horizon, Medisys, Fidia, UBM LLC, Trio Health, Medscape, Adept Field Solutions, Clinical Care Options, Clearview Healthcare Partners, Putnam Associates, Focus Forward, Navigant Consulting, Spherix, and Practice Point Communications (less than $10,000 each), owns stock options in TPT Global Tech, Vaxart Pharmaceuticals, and Charlotte's Web Holdings, Inc., and is on the speaker's bureau of Simply Speaking. Dr. Terkeltaub has received consulting fees from SOBI, Horizon Therapeutics, and AstraZeneca (less than $10,000 each) and Selecta (more than $10,000) and has received research support from AstraZeneca. Dr. Abhishek has received consulting fees from NGM Biopharmaceuticals (less than $10,000). No other disclosures relevant to this article were reported.

Address correspondence to Sara K. Tedeschi, MD, MPH, Brigham and Women's Hospital Division of Rheumatology, Inflammation and Immunity, 60 Fenwood Road, Suite 6016, Boston, MA 02115. Email: stedeschi1@bwh.harvard.edu.

Submitted for publication January 12, 2021; accepted in revised form April 6, 2021.

comprehensive list of all possible items to be considered for classification of the disease (phase I: item generation), followed by reducing that list to a parsimonious number providing a comprehensive yet manageable set for collecting data from de-identified patient cases that form a derivation set (phase II: item reduction), reported herein. In phase III, the items will be structured into an evaluable framework, and the relative importance of items will be compared and item weights assigned through a multicriterion decision analysis exercise. A threshold score will then be determined for classifying a patient as having CPPD. The classification criteria will be validated in phase IV using an independent validation set of de-identified patient cases, and test performance characteristics will be determined (Figure 1). Study oversight and management in phases I and II were overseen by 4 co-PIs (WJT, RT, HC, and AA) and 8 other members of the steering committee with expertise in CPPD research and/or classification criteria methodology (SKT, TP, AL, RPN, ND, TN, FPR, and AR).

Item generation. The purpose of item generation was to supply a comprehensive list of items to be considered as classification criteria. This list should represent typical manifestations of CPPD and include items across a variety of categories, including imaging, clinical, and laboratory features.

Scoping literature review to begin item generation. A scoping literature review was conducted within Web of Science, Embase, MEDLINE, CINAHL, and AMED databases through May 31, 2019 (for the protocol, see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24619). The search identified randomized controlled trials, cohort studies, cross-sectional studies, and case–control studies of adults with CPPD without language restriction. Studies with <20 participants were excluded to increase generalizability of features extracted from the literature. Terms included in the search included CPPD, CPDD, chondrocalcinosis, pseudogout, pyrophosphate, and calcium pyrophosphate. Literature search and data management were performed by an experienced systematic reviewer (BK). Four investigators divided the identified studies, screened titles and abstracts for eligibility, and performed full-text review on eligible studies (SKT, TP, AL, and CG). A second investigator from among those 4 independently examined 10% of titles, abstracts, and eligible studies to assess for agreement with the primary reviewer. Data extracted from full-text review included all features that were positively or negatively associated with CPPD. The list of identified features included demographic characteristics; signs and symptoms; imaging findings on plain radiograph, ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI); and laboratory findings in synovial fluid and peripheral blood. Risk factors for CPPD were also included in the list.

Expert and patient input for item generation. The steering committee invited 23 additional CPPD experts to form a 35-member combined expert committee comprising 32 rheumatologists, 2 classification criteria methodologists, and 1 musculoskeletal radiologist with expertise in CPPD (see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24619). The combined expert committee members were asked to review the list of items generated from the literature and suggest additional items, including those that may not have been reported in the literature but that influence clinical judgment. Two patient research partners also reviewed the list of items from the literature and proposed additional items. The literature review list, expert suggestions, and patient suggestions were combined into 1 list with duplicates removed, yielding a list of unique potential items to consider in developing the classification criteria.

Item reduction. The item-reduction phase was guided by questions of specificity and generalizability. During each of 3 steps (Figure 1), experts were asked to consider the discriminating ability of items, i.e., whether they help distinguish CPPD from non-CPPD, a key feature of classification criteria. Experts

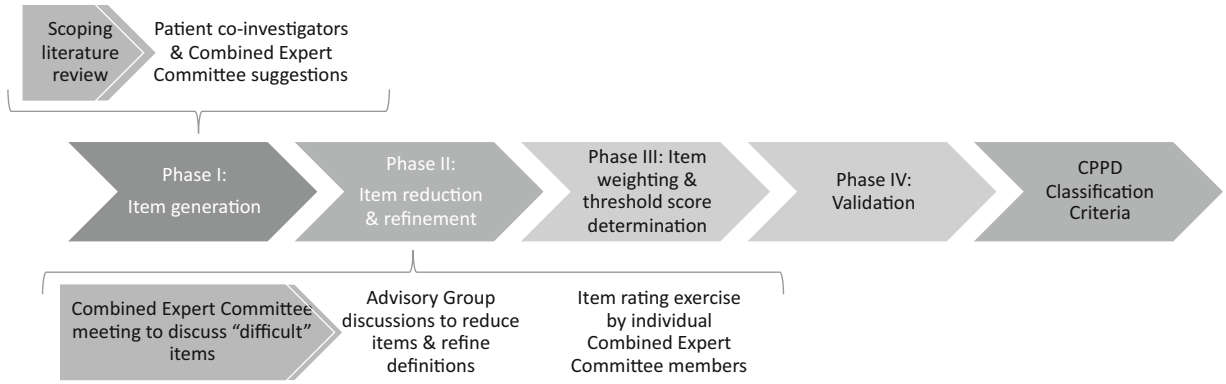

Figure 1. Overview of classification criteria development process, with a focus on subprocesses in phases I and II. CPPD = calcium pyrophosphate deposition disease.

were also asked to consider whether items might be highly specific for CPPD, but too rare to be implemented across a wide range of medical centers internationally.

The first step of this phase involved a 4-hour in-person meeting of the combined expert committee immediately prior to the American College of Rheumatology meeting in Atlanta in November 2019. Experts were presented with the list of items categorized according to clinical, imaging, and laboratory features and risk factors. The group held an open discussion as to whether each item could plausibly discriminate between CPPD and non-CPPD based on expert opinion. Consensus to eliminate an item was considered achieved if there were no objections to the proposed elimination.

The second step was led by 3 advisory groups working remotely from December 2019 to January 2020. Clinical, laboratory, and imaging advisory groups each included 6–10 members from the combined expert committee and were co-led by an expert in the topic and a junior faculty facilitator (see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24619). Each advisory group held 1–2 teleconferences to further reduce items following the same consensus-based discussion process that was used during the prior combined expert committee meeting. Advisory groups also refined item definitions to increase their precision. The goal was to consolidate the number of candidate items to approximately 150 to make the subsequent item-rating exercise more feasible. Each group was encouraged to retain all items that, based on expert opinion, might distinguish between CPPD and non-CPPD.

The third step involved an item-rating exercise in January 2020. Combined expert committee members were invited to use an online web-survey platform (www.surveymonkey.com) to rate each potential classification criterion on a 7-point scale, indicating how much the presence of that item would push them toward or against classifying a patient as having CPPD. Each item was rated from –3 to +3, with –3 signifying that it pushed the expert strongly away from classifying the patient as having CPPD, 0 indicating that did not influence them in classifying a patient as having CPPD, and +3 meaning it pushed them strongly toward classifying the patient as having CPPD. Experts were asked to select "not applicable" if they were unsure how to respond and were able to add comments as free text. The steering committee reviewed the distribution of ratings for each item and retained items with modal ratings of –3, –2, +2, or +3, as these scores indicated a general consensus that the item is an important discriminator for or against CPPD. Items with a modal rating of –1, 0, or +1 were discussed in a steering committee teleconference before removal from the list of candidate items; those items that were considered to be potentially important discriminators were retained. Items retained at the end of phase II (item reduction) will be included in subsequent phases of the classification criteria development.

RESULTS

Results of item generation. The scoping literature review yielded 67 manuscripts eligible for full-text review (Figure 2), from which 312 unique items were identified after removing duplicates and merging very similar items together. Twenty-five experts and 2 patients suggested 108 additional items, yielding a total of 420 candidate items. These 420 items were grouped into 3 broad categories: clinical (n = 264, 62.9%), imaging (n = 90, 21.4%), and laboratory (n = 66, 15.7%) The full list of candidate items is shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24619.

Item reduction at in-person meeting. Twenty-six combined expert committee members attended the in-person meeting, during which 132 items were removed by consensus decision (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24619). The group first discussed whether to exclude patients with other types of arthritis such as seropositive rheumatoid arthritis or sacroiliitis from being classified as having CPPD. The group decided that because CPPD can coexist with these conditions, other forms of arthritis should not be exclusion criteria for classification. Combinations of osteoarthritis in 1 joint plus chondrocalcinosis at another joint were removed from further consideration, as the group felt that complexity of these permutations would be impractical to operationalize as classification criteria. Other osteoarthritis items were retained for future discussion and refinement.

Age as a possible criterion was discussed at length. Some experts noted that age is to CPPD what serum urate is to gout, meaning that it is critical for disease pathogenesis. The group did not reach consensus on this issue and decided to retain age at this step and reconsider it later. A number of features of acute CPP crystal arthritis were eliminated since they were not felt to distinguish between CPPD and other forms of inflammatory arthritis, e.g., fever, response to colchicine, and knee pain. Other clinical items were eliminated as they were complex constructs of items that had been included individually (e.g., pseudo-OA with synovitis) or were characteristic of many types of arthritis (e.g., disability with ambulation). Comorbidities not thought to be pathogenically related to CPPD were eliminated, such as diabetes mellitus, end-stage renal disease, and hypertension. The group discussed whether certain medications might be pathogenic factors, and these were retained, including bisphosphonates and diuretics.

The macroscopic appearance of synovial fluid and white blood cell count were considered nonspecific for CPPD and all such features were eliminated. The absence of CPP crystals was discussed in depth. In the ACR/EULAR 2015 gout classification criteria, the absence of monosodium urate crystals from an aspirated joint carried negative weight (8). The group felt that the

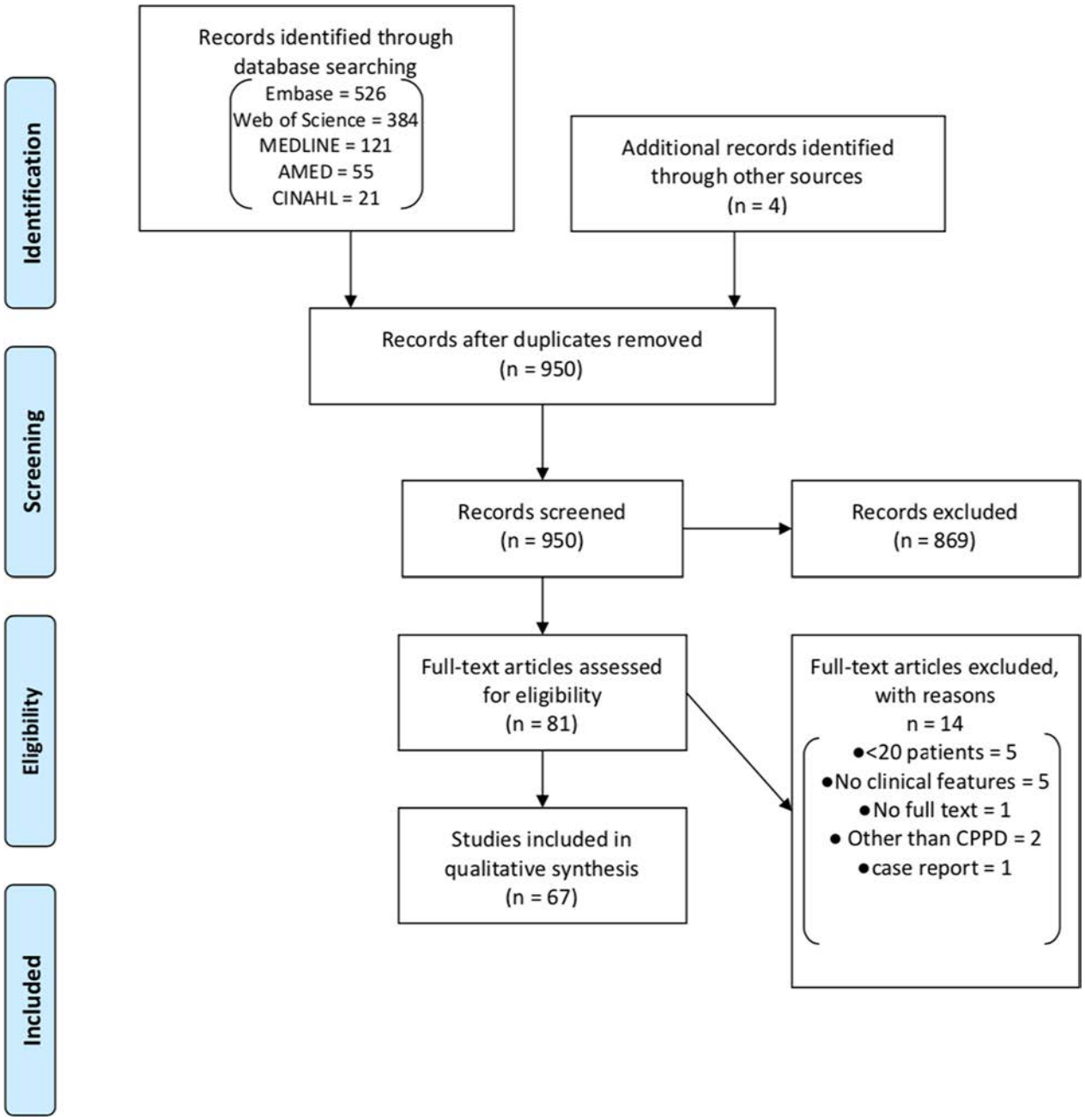

Figure 2. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (2009) for scoping literature review for calcium pyrophosphate deposition disease (CPPD).

absence of CPP crystals in synovial fluid may have a different impact on CPPD classification, given that CPP crystals can be difficult to identify via compensated polarized light microscopy, and there may be a high false negative rate. Experts favored creating a new item capturing repeatedly negative synovial fluid crystal analysis, distinct from an item characterized by a single negative synovial fluid crystal analysis. Monosodium urate crystals were retained as a potential criterion for 2 reasons: gout and CPPD can co-occur, CPP crystals can be found in gouty tophi, and there is a possibility that monosodium urate crystals could carry negative weight toward classifying a patient as having CPPD (24,25).

Many of the peripheral blood tests identified in the literature search were recognized to be nonspecific for CPPD (e.g., high white blood cell count, high erythrocyte sedimentation rate), and these were eliminated. Autoimmune serologic results were eliminated with the exception of rheumatoid factor and anti–citrullinated protein antibody (ACPA), as these may warrant consideration in future phases, for example, in potentially being assigned a negative weight toward CPPD classification. Genetic

Table 1. CPPD combined expert committee rating of influence on the likelihood of CPPD (–3 to +3) for candidate clinical and laboratory items*

Candidate criteria	Modal rating score ± SD
Age at symptom onset, years	
≥91	3 ± 0.81
81–90	3 ± 0.84
71–80	2 ± 0.82
61–70	1 ± 0.82
51–60	0 ± 1.03
<50	–2 ± 1.11
Acute inflammatory arthritis	
Crowned dens syndrome	3 ± 0.84
Knee	2 ± 0.90
Wrist	2 ± 0.77
Diffuse inflammatory hand swelling involving periarticular structures	1 ± 1.24
Ankle	0 ± 1.21
Spine	0 ± 1.50
Mid-foot	–1 ± 1.06
1st MTP joint	–2 ± 1.21
Recurrence and pattern of joint involvement	
One self-limiting episode of acute inflammatory arthritis (resolved without treatment)	2 ± 0.83
Recurrent attacks of acute inflammatory arthritis in a characteristic joint (e.g., wrist, knee)	2 ± 0.77
Additive attacks of inflammatory polyarthritis (i.e., the episode starts in 1 joint, then another joint becomes involved, then another, etc.)	1 ± 1.32
Asymmetric attacks of inflammatory polyarthritis	1 ± 1.12
Inflammatory oligoarthritis with asymmetric joint involvement	1 ± 0.98
Persistent inflammatory polyarthritis with intermittent flares	1 ± 1.12
Recurrent attacks of acute inflammatory arthritis in joints other than those that are typically involved by CPPD	0 ± 1.21
Persistent inflammatory polyarthritis without flares	–2 ± 1.32
Onset and treatment response	
Rapid onset acute inflammatory arthritis (time to maximal pain <24 hours)†	1 ± 1.01
Acute inflammatory arthritis that developed in the context of acute illness, surgery, or joint trauma	1 ± 0.97
Inflammatory monoarthritis with good response to colchicine	1 ± 1.23
Inflammatory oligoarthritis with good response to colchicine	1 ± 1.14
Inflammatory polyarthritis with good response to colchicine	1 ± 1.15
Acute inflammatory arthritis that developed in the context of recent bisphosphonate use	1 ± 1.06
Physical findings	
Erythema of the involved joints during acute inflammatory arthritis	1 ± 0.70
Hemarthrosis without other explanation (e.g., without joint injury, anticoagulation, or pigmented villonodular synovitis)	1 ± 1.23
Palpable subcutaneous tophus	–3 ± 1.00
Psoriasis	–2 ± 1.16
Comorbidities and family history	
History of familial CPPD	3 ± 0.81
Gitelman disease	
Hereditary hemochromatosis	3 ± 0.87
Wilson disease	2 ± 1.24
Ochronosis	2 ± 1.33
History of meniscectomy or arthroscopy	0 ± 1.26
Acromegaly	0 ± 1.10
Short bowel disease	0 ± 1.76
Osteoarthritis location and features	
Wrist	2 ± 0.90
Second and third MCP joints	2 ± 1.08
Second or third MCP joints‡	1 ± 0.84
Scaphotrapezium joint†	1 ± 0.69
Scaphotrapezium joint without OA of the trapeziometacarpal joint	1 ± 0.82
Scapholunate advanced collapse wrist†	1 ± 0.83
Elbow	1 ± 1.00
Patellofemoral joint without OA of the tibiofemoral joint	1 ± 1.09
Ankle, without history of trauma	1 ± 1.01
>1 of the following: ankle, wrist, shoulder, elbow	1 ± 0.97
OA at 1 or more joints with signs of joint inflammation on examination	1 ± 0.90
Patellofemoral joint with wrap-around patella	0 ± 1.14

(Continued)

Table 1. (*Cont'd*)

Candidate criteria	Modal rating score ± SD
Lateral tibiofemoral joint	0 ± 0.78
Patellofemoral and tibiofemoral joints	0 ± 0.89
Shoulder	0 ± 1.20
Scaphotrapezium joint without OA of the 1st MCP joint	0 ± 0.95
Scaphotrapezium and trapeziometacarpal joints	0 ± 1.00
Trapeziometacarpal joint	0 ± 0.61
Synovial fluid findings	
At least 1 joint aspirate demonstrating CPP crystals	3 ± 0.26
CPP crystals absent 1 occasion†	−1 ± 0.63
CPP crystals absent on 2 or more occasions	−2 ± 0.85
Monosodium urate crystals present, no CPPD crystals, and no microorganisms	−3 ± 0.83
Microorganisms on culture, no monosodium urate crystals, no CPPD crystals	−3 ± 0.99
Laboratory findings	
Hypomagnesemia	3 ± 0.83
Hyperparathyroidism	3 ± 0.78
Hypercalcemia	2 ± 0.92
Low-titer positive RF (1–3 times upper limit of normal)	0 ± 0.68
High-titer positive RF (>3 times upper limit of normal)	−2 ± 0.96
Low-titer positive ACPA (1–3 times upper limit of normal)†	0 ± 0.98
High-titer positive ACPA (>3 times upper limit of normal)	−3 ± 0.97

* ACPA = anti–citrullinated protein antibody; CPP = calcium pyrophosphate; CPPD = calcium pyrophosphate deposition disease; MCP = metacarpophalangeal; MTP = metatarsophalangeal; OA = osteoarthritis; RF = rheumatoid factor.
† Retained despite mode rating score −1, 0, or +1 per steering committee discussion.
‡ OA of second MCP joint and OA of third MCP joint were split into 2 discrete items. Retained despite mode rating score −1, 0, or +1 per steering committee discussion.

polymorphisms and mutations were considered highly specific, but too uncommon to be relevant for a classification criteria system and were eliminated. Metabolic abnormalities that are risk factors for CPPD were retained for future steps, such as high serum ferritin, hypomagnesemia, and hypophosphatasia. Laboratory assays that have been reported in research literature but are not routinely available for clinical use (e.g., high serum MMP3 enzyme and tissue-type plasminogen activator levels), were considered impractical for classification criteria and were omitted.

Candidate items in the imaging category were nearly uniformly retained during this discussion. The group felt that decisions about including imaging evidence of CPP crystal deposition at specific joints or combinations of joints were best left to future steps.

Advisory group meetings to further reduce and refine items. Each of the 3 advisory groups reached consensus on eliminating items that were not considered important discriminators between CPPD and non-CPPD. Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24619, presents the items that were retained and the updated definitions resulting from this step. Periarticular soft tissue swelling, asymmetric joint involvement, and hip pain are among the clinical items that were removed. The presence of CPP crystals specifically within white blood cells was removed, as laboratory experts reached consensus that extracellular and intracellular CPP crystals should carry equal importance in CPPD classification. Ultrasound findings characteristic of gout, such as the double-contour sign and "snowstorm" appearance, were eliminated as potential items weighing against CPPD classification. The advisory group discussions yielded 61 clinical items, 57 imaging items, and 9 laboratory items (127 items total) for further consideration.

Item-rating exercise. Twenty-nine combined expert committee members (82.9%) completed the item-rating exercise, which included 127 items. In total, 83 items (65.4%) had a modal rating of +3, +2, −2, or −3 and were retained for future phases based on this rating alone (Tables 1 and 2). Nearly all of the imaging items, synovial fluid CPP crystals, age >80 years at symptom onset, crowned dens syndrome, a history of familial CPPD, Gitelman syndrome, hemochromatosis, hyperparathyroidism, and hypomagnesemia were most strongly and consistently rated as influencing consideration of classifying as CPPD, with a modal rating of +3 and SD <1. Four items were rated as strongly indicating that the case was unlikely to be CPPD: palpable subcutaneous tophi, monosodium urate crystals (and no CPP crystals) in synovial fluid, positive synovial fluid culture, and high-titer ACPA (>3 times the upper limit of normal).

Seven respondents entered free-text comments indicating that the experience of the person performing polarized light microscopy is quite important, particularly when CPP crystals are not observed in synovial fluid. Thirteen respondents selected "not applicable" for questions related to MRI findings; aside from MRI, "not applicable" was rarely chosen.

Table 2. CPPD combined expert committee rating of influence on the likelihood of CPPD (–3 to +3) for candidate imaging items*

	Affected joint	Knee†	Wrist†	Second and/or third MCP†	Hip†	Symphysis pubis†	C1/C2†	Shoulder†
Plain radiograph: calcification in regions of fibro- or hyaline cartilage	3	3	3	3	3	3	3	3
Plain radiograph: calcification of the synovial membrane/capsule/tendon	1‡	1‡	2	2	–	–	3	–
Conventional CT: calcification in regions of fibro- or hyaline cartilage	3	3	3	3	3	3	3	3
Conventional CT: calcification of the synovial membrane/capsule/tendon	2	2	2	2	2	–	3	–
Ultrasound: CPP crystal deposition in fibro- or hyaline cartilage	3	3	3	3	–	–	–	3
Ultrasound: CPP crystal deposition in synovial membrane/capsule/tendon	2	2	2	3	–	–	–	–
Cartilage calcium aggregates in the affected joint on ultrashort echo time or in 3D dual-echo steady-state gradient-echo MRI sequences	2	1	2	3	3	3	2	2
Dual-energy CT: CPP crystal deposition in fibro- or hyaline cartilage	3	3	3	3	3	–	3	3
Dual-energy CT: CPP crystal deposition in synovial membrane/capsule/tendon	3	3	3	3	3	–	3	–

* The mode rating score is presented for each imaging modality and location. Imaging data from 1 affected joint, the knee (regardless of symptoms), and the wrist (regardless of symptoms) were ultimately recommended as candidate criteria by the steering committee. CPP = calcium pyrophosphate; CPPD = calcium pyrophosphate deposition disease; CT = computed tomography; MCP = metacarpophalangeal; MRI = magnetic resonance imaging.
† Regardless of symptoms at this joint.
‡ Plain radiograph of the affected joint and plain radiograph of the knee with calcification of the synovial membrane/capsule/tendon were retained per steering committee discussion.

The steering committee discussed 28 items (22.0%) with a modal rating of –1, 0, or +1. Eight of these were retained at this stage to facilitate data collection to evaluate the potential relevance of these items in a data-driven approach (Tables 1 and 2). The modal score was 0 for a history of meniscectomy or arthroscopy, and several experts commented that meniscectomy is prone to favor local disease but not systemic CPPD.

Because the number of candidate imaging items remained large and included numerous joints at this stage, the steering committee proposed a parsimonious approach for the imaging items: given that CPPD most commonly affects the knee and wrist, they agreed that imaging data regarding calcification suggestive of CPP crystal deposition should be routinely assessed for these joints. Including imaging data from 1 additional affected joint was felt to provide a reasonable balance between informativeness for potential classification as CPPD and workload for future steps of data collection. Imaging data from any additional joints were deemed to likely provide only minimal additional value for discriminating CPPD from non-CPPD.

Following the item-rating exercise, the steering committee and laboratory advisory group added an additional item, histopathologic evidence of CPP crystals, as they recognized that surgical specimens from knee replacement or spinal surgery may reveal these crystals in patients without arthrocentesis or synovial fluid crystal analysis. Therefore, 56 items were retained for future phases of classification criteria development.

DISCUSSION

We present results from the initial phases of CPPD classification criteria development, which will culminate in the first validated system to identify patients with CPPD for clinical research studies. A scoping review identified >300 unique clinical, laboratory, and imaging items from publications with a variety of study designs and a range of definitions of CPPD. Input from CPPD experts, some of whom have >5 decades of experience treating this disease, and CPPD patients was critical for identifying >100 additional items that are observed in clinical practice but may not have been published in the relatively scant literature on CPPD.

The item-reduction process achieved its goal of honing a manageable number of items for future data collection and testing. Expert opinion was key for the rating exercise. Candidate items for which the mode was +3 or –3 had an SD of <1, indicating consensus among individual raters and providing face validity for these items. For a handful of items with a modal rating of –1, 0 or +1, the steering committee took an inclusive approach to retain these items now and analyze them in a future data-driven step.

While the item-generation phase had produced a large number of clinical features, the item-reduction phase revealed that many of these occur in a large number of conditions, and as a result, do not discriminate between CPPD and non-CPPD, a key requirement for an item to be considered for inclusion in

classification criteria. Periarticular soft tissue swelling and ultrasound double-contour sign are among the items that were removed during item reduction (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24619). Factors that are pathogenic and central to the construct of disease, such as older age, hyperparathyroidism, and osteoarthritis of the knee, wrist, and second or third metacarpophalangeal joints were retained.

Notably, 40 of 49 items that were rated as strongly influencing experts toward classifying a patient as having CPPD (rating +3) were imaging findings of the affected joint or other joints characteristic of CPPD, regardless of symptoms in those joints (e.g., incidental findings in the shoulder on CT chest). Given the large number of imaging items that were highly rated as +3, the steering committee proposed including imaging of the knee and wrist, as these are commonly involved in CPPD and are most clinically impactful, as well as adding 1 additional affected joint. Dual-energy CT was consistently rated +3 for all joint locations, consistent with the limited literature suggesting high specificity for CPPD. Ultrasound evidence of CPPD was also rated highly, reflecting the published reports of high specificity in fibro- or hyaline cartilage (26,27). Plain radiograph calcification in regions of fibro- or hyaline cartilage in all proposed joint locations was also rated as strongly influencing a decision toward classifying as CPPD, while calcification in synovial membrane, capsule, or tendon generally was not as strongly rated because this feature was potentially more likely due to another form of calcium crystal deposition, e.g. calcium hydroxyapatite. Specific MRI sequences received mixed ratings, and notably one-third of respondents did not answer questions about MRI due to lack of experience with these sequences and lack of sufficient data in the published literature to guide decision-making. Ongoing work for imaging items includes refining definitions to include technical parameters and developing a set of representative CPPD images for reference (28).

The next phase of classification criteria development will involve de-identified prospective data collection from real-life patients to form a derivation cohort of patients with a range of probabilities for being classified as having CPPD. Collecting data on the 56 items from hundreds of patients will enable evaluation of the distribution of each item across a range of patient scenarios, ranging from "definite CPPD" to "not CPPD" and inform a subsequent multicriterion decision analytic exercise using 1000Minds software. Prior to the 1000Minds exercise, latent class analysis will be performed to identify which items cluster together to further reduce the number of items and identify potential domains. An expert consensus workshop and 1000Minds exercise will further reduce items to a manageable framework of domains and establish weights for individual items within each domain and across domains. The current item-reduction phase was successful in achieving a manageable number of candidate criteria for the data collection step and subsequent testing.

Classification criteria for CPPD will provide a validated definition of CPPD for use in clinical research, facilitating comparisons across studies and enrollment into clinical trials. An international CPPD combined expert committee developed a comprehensive list of candidate items, informed by the literature and by expert opinion. This list was then refined and narrowed to items that experts considered most influential in pushing them toward or against classifying an individual patient as having CPPD. Phase III, item weighting and threshold determination, is underway. The final phase, phase IV, will validate the CPPD classification criteria system.

ACKNOWLEDGMENTS

We would like to thank our patient research partners, Diane A. Sachs (US) and Patricia Brackley (UK), for their contributions.

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. Tedeschi 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. Tedeschi, Naden, Taylor, Terkeltaub, Choi, Abhishek.

Acquisition of data. Tedeschi, Pascart, Latourte, Godsave, Kundakci, Dalbeth, Neogi, Perez-Ruiz, Rosenthal, Becce, Pascual, Andres, Bardin, Doherty, Ea, Filippou, FitzGerald, Guitierrez, Iagnocco, Jansen, Kohler, Lioté, Matza, McCarthy, Ramonda, Reginato, Richette, Singh, Sivera, So, Stamp, Yinh, Yokose, Choi, Abhishek.

Analysis and interpretation of data. Tedeschi, Pascart, Latourte, Taylor, Dalbeth, Neogi, Perez-Ruiz, Rosenthal, Becce, Choi, Abhishek.

REFERENCES

1. Abhishek A, Neogi T, Choi H, Doherty M, Rosenthal AK, Terkeltaub T. Unmet needs and the path forward in joint disease associated with calcium pyrophosphate crystal deposition. Arthritis Rheumatol 2018; 70:1182–91.
2. Zhang W, Doherty M, Bardin T, Barskova V, Guerne Pa, Jansen TL, et al. European League Against Rheumatism recommendations for calcium pyrophosphate deposition. Part I: terminology and diagnosis. Ann Rheum Dis 2011;70:563–70.
3. Ramonda R, Musacchio E, Perissinotto E, Sartori L, Punzi L, Corti MC, et al. Prevalence of chondrocalcinosis in Italian subjects from northeastern Italy: the Pro.V.A. (PROgetto Veneto Anziani) study. Clin Exp Rheumatol 2009;27:981–4.
4. Maravic M, Ea HK. Hospital burden of gout, pseudogout and other crystal arthropathies in France. Joint Bone Spine 2015;82:326–9.
5. Felson DT, Anderson JJ, Naimark A, Kannel W, Meenan RF. The prevalence of chondrocalcinosis in the elderly and its association with knee osteoarthritis: the Framingham study. J Rheumatol 1989;16:1241–5.
6. Fuller A, Cai K, Diaz-Torne C, Filippou G, Pascart T, Hensey O, et al. Outcome domains reported by patients, caregivers, healthcare professionals and stakeholders for calcium pyrophosphate deposition (CPPD): a content analysis based on semi-structured qualitative interviews from the OMERACT CPPD working group. Semin Arthritis Rheum 2021;51:650–4.
7. Cai K, Fuller A, Hensey O, Grossberg D, Christensen R, Shea B, et al. Outcome domains reported in calcium pyrophosphate deposition studies: a scoping review by the OMERACT CPPD working group. Semin Arthritis Rheum 2020;50:719–27.

8. 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 Rheumatol 2015;67:2557–68.
9. 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.
10. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71:1400–12.
11. Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
12. McCarty DJ. Calcium pyrophosphate dihydrate crystal deposition disease: nomenclature and diagnostic criteria. Ann Intern Med 1977;87:241–2.
13. Ryan L, McCarty D. Calcium pyrophosphate crystal deposition disease; pseudogout; articular chondrocalcinosis. In: McCarty D, editor. Arthritis and allied conditions. 10th ed. Philadelphia: Lea and Febiger; 1985. p. 1515–46.
14. Utsinger PD, Resnick D, Zvaifler NJ. Wrist arthropathy in calcium pyrophosphate dihydrate deposition disease. Arthritis Rheum 1975; 18:485–91.
15. Tanikawa H, Ogawa R, Okuma K, Harato K, Niki Y, Kobayashi S, et al. Detection of calcium pyrophosphate dihydrate crystals in knee meniscus by dual-energy computed tomography. J Orthop Surg Res 2018;13:73.
16. Andres M, Vela P, Jovani V, Pascual E. Most needle-shaped calcium pyrophosphate crystals lack birefringence. Rheumatology (Oxford) 2019;58:1095–8.
17. Berendsen D, Neogi T, Taylor WJ, Dalbeth N, Jansen TL. Crystal identification of synovial fluid aspiration by polarized light microscopy: an online test suggesting that our traditional rheumatologic competence needs renewed attention and training. Clin Rheumatol 2017;36:641–7.
18. Pascual E, Tovar J, Ruiz MT. The ordinary light microscope: an appropriate tool for provisional detection and identification of crystals in synovial fluid. Ann Rheum Dis 1989;48:983–5.
19. Lumbreras B, Pascual E, Frasquet J, Gonzalez-Salinas J, Rodriguez E, Hernandez-Aguado I. Analysis for crystals in synovial fluid: training of the analysts results in high consistency. Ann Rheum Dis 2005;64:612–5.
20. Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res (Hoboken) 2015;67:891–7.
21. Classification and Response Criteria Subcommittee of the American College of Rheumatology Committee on Quality Measures. Development of classification and response criteria for rheumatic diseases. Arthritis Rheum 2006;55:348–52.
22. Fransen J, Johnson SR, van den Hoogen F, Baron M, Allanore Y, Carreira PE, et al. Items for developing revised classification criteria in systemic sclerosis: results of a consensus exercise. Arthritis Care Res (Hoboken) 2012;64:351–7.
23. Schmajuk G, Hoyer BF, Aringer M, Johnson SR, Daikh DI, Dörner T, et al. Multicenter Delphi exercise to identify important key items for classifying systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70:1488–94.
24. Ea HK, Gauffenic A, Nguyen QD, Pham NG, Oliver O, Frochot V, et al. Calcium pyrophosphate crystal deposition in gouty tophi. Arthritis Rheumatol 2021;73:324–9.
25. Rosenthal AK, Ryan LM. Calcium pyrophosphate deposition disease. N Engl J Med 2016;374:2575–84.
26. Filippou G, Adinolfi A, Iagnocco A, Filippucci E, Cimmino MA, Bertoldi I, et al. Ultrasound in the diagnosis of calcium pyrophosphate dihydrate deposition disease: a systematic literature review and a meta-analysis. Osteoarthritis Cartilage 2016;24:973–81.
27. Cipolletta E, Filippou G, Scire CA, Di Matteo A, Di Battista J, Salaffi F, et al. The diagnostic value of conventional radiography and musculoskeletal ultrasonography in calcium pyrophosphate deposition disease: a systematic literature review and meta-analysis. Osteoarthritis Cartilage 2021;29:619–32.
28. Tedeschi SK, Becce F, Pascart T, Guermazi A, Budzik JF, Dalbeth N, et al. Imaging features of calcium pyrophosphate deposition (CPPD) disease: consensus definitions from an international multidisciplinary working group. Arthritis Care Res (Hoboken) 2022. Epub ahead of print. doi: 10.1002/acr.24898.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1659–1666
DOI 10.1002/acr.24643

American College of Rheumatology
Empowering Rheumatology Professionals

Predictors of Lumbar Spine Degeneration and Low Back Pain in the Community: The Johnston County Osteoarthritis Project

Adam P. Goode,[1] Rebecca J. Cleveland,[2] Steven Z. George,[1] Todd A. Schwartz,[2] Virginia B. Kraus,[1] Jordan B. Renner,[2] Richard H. Gracely,[2] Louis E. DeFrate,[1] David Hu,[2] Joanne M. Jordan,[2] and Yvonne M. Golightly[2]

Objective. To determine the incidence and worsening of lumbar spine structure and low back pain (LBP) and whether they are predicted by demographic characteristics or clinical characteristics or appendicular joint osteoarthritis (OA).

Methods. Paired baseline (2003–2004) and follow-up (2006–2010) lumbar spine radiographs from the Johnston County Osteoarthritis Project were graded for osteophytes (OST), disc space narrowing (DSN), spondylolisthesis, and presence of facet joint OA (FOA). Spine OA was defined as at least mild OST and mild DSN at the same level for any level of the lumbar spine. LBP, comorbidities, and back injury were self-reported. Weibull models were used to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs) of spine phenotypes accounting for potential predictors including demographic characteristics, clinical characteristics, comorbidities, obesity, and appendicular OA.

Results. Obesity was a consistent and strong predictor of incidence of DSN (HR 1.80 [95% CI 1.09–2.98]), spine OA (HR 1.56 [95% CI 1.01–2.41]), FOA (HR 4.99 [95% CI 1.46–17.10]), spondylolisthesis (HR 1.87 [95% CI 1.02–3.43]), and LBP (HR 1.75 [95% CI 1.19–2.56]), and worsening of DSN (HR 1.51 [95% CI 1.09–2.09]) and LBP (HR 1.51 [95% CI 1.12–2.06]). Knee OA was a predictor of incident FOA (HR 4.18 [95% CI 1.44–12.2]). Spine OA (HR 1.80 [95% CI 1.24–2.63]) and OST (HR 1.85 [95% CI 1.02–3.36]) were predictors of incidence of LBP. Hip OA (HR 1.39 [95% CI 1.04–1.85]) and OST (HR 1.58 [95% CI 1.00–2.49]) were predictors of LBP worsening.

Conclusion. Among the multiple predictors of spine phenotypes, obesity was a common predictor for both incidence and worsening of lumbar spine degeneration and LBP.

INTRODUCTION

Chronic low back pain (LBP) impacts >31 million Americans at any given moment (1) and has increased 3-fold in prevalence over a 10-year period (2). The traditional gateway to interventions for LBP is diagnostic clinical imaging (3). This is especially true within the primary care setting, where although plain film radiographs are not generally recommended by guidelines (4), they are nonetheless frequently performed for examining whether lumbar spine structure is linked to LBP (5,6). Improved understanding of the relationship between lumbar spine imaging and LBP is critically important (7–13), as discordance between spine degeneration and LBP may lead to additional tests and referrals, some of which may have questionable benefit (14).

The findings and conclusions herein are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention, the National Institute on Aging, or the National Institute of Arthritis and Musculoskeletal and Skin Diseases.

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases [NIAMS] grant R01-AR-071440 to Drs. Goode, Cleveland, George, Schwartz, Kraus, Jordan, and Golightly, grant R01-AR-075399 to Drs. Goode, George, Kraus, and DeFrate, and grant R01-AR-07800 to Drs. DeFrate and Goode). The Johnston County Osteoarthritis Project is supported in part by the Centers for Disease Control and Prevention/Association of Schools of Public Health (cooperative agreements S043, S1734, and S3486) and by the NIH (NIAMS Multipurpose Arthritis and Musculoskeletal Disease Center grant 5-P60-AR-30701, NIAMS Multidisciplinary Clinical Research Center grant 5-P60-AR-49465-03, and NIAMS Core Centers for Clinical Research grant 1P30-AR-072580-01A1). Dr. Kraus' work was supported by the NIH (National Institute on Aging grant P30-AG-0287160). Dr. Golightly's work was supported by the NIH (NIAMS grant R01-AR-067743).

[1]Adam P. Goode, DPT, PhD, Steven Z. George, PT, PhD, Virginia B. Kraus, MD, PhD, Louis E. DeFrate, ScD: Duke University School of Medicine, Durham, North Carolina; [2]Rebecca J. Cleveland, PhD, Todd A. Schwartz, DrPH, Jordan B. Renner, MD, Richard H. Gracely, PhD, David Hu, MS, Joanne M. Jordan, MD, MPH, Yvonne M. Golightly, PT, PhD: University of North Carolina, Chapel Hill.

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

Address correspondence to Adam P. Goode, DPT, PhD, Duke University School of Medicine, Box 104002 DUMC, Durham, NC 27708. Email: Adam.goode@duke.edu.

Submitted for publication August 20, 2020; accepted in revised form May 4, 2021.

SIGNIFICANCE & INNOVATIONS

- This longitudinal study of 819 participants over an average of 5.5 years found that obesity and appendicular joint osteoarthritis were significant predictors of the incidence and worsening of both lumbar spine degeneration and low back pain.
- Since obesity is a modifiable risk factor, efforts to decrease it could lessen the development and worsening of lumbar spine degeneration and low back pain.

Disc space narrowing (DSN) from degeneration of the intervertebral disc, vertebral osteophytes (OST) formation, facet joint osteoarthritis (FOA), and spondylolisthesis are potential sources of nociceptive pain in the lower back. Cross-sectional studies have linked lumbar spine degeneration with demographic and clinical characteristics (7–9,15,16). However, longitudinal community-based studies are sparse, with most including only women, a considerable length of follow-up (~9 years), and limited LBP examination (10,12,13). Muraki et al (17), using data from a Japanese cohort, identified that sex was a significant predictor of incidence of lumbar spine degeneration; more severe spine degeneration was also a significant predictor of LBP. However, differences between Japanese and US lifestyles may result in different predictors of incidence and worsening of lumbar spine degeneration and LBP.

To our knowledge, no other community-based study within the US has examined the incidence, worsening, and the longitudinal relationship between demographic and clinical characteristics and appendicular joint OA as predictors of radiographic lumbar spine degeneration and LBP within the same cohort. Therefore, our objective was to: 1) describe the incidence and worsening of lumbar spine vertebral OST, DSN, FOA, spondylolisthesis, and LBP, and 2) determine demographic or clinical characteristics and appendicular joint OA predictors of incidence and worsening of lumbar spine degeneration and LBP. We hypothesized that there would be 1) multiple factors predictive of degeneration, 2) multiple factors predictive of LBP, and 3) few (if any) factors predictive of both degeneration and LBP. These hypotheses are driven by previous research that suggests that factors predictive of structural changes are not the same factors as those predictive of LBP (17–19). However, we intentionally posit a general hypothesis as we believe several factors are likely to be predictive of these outcomes.

PATIENTS AND METHODS

Participants. Details of the sampling strategy and recruitment methods used for the Johnston County Osteoarthritis (JoCo OA) Project are described elsewhere (7,20). This ongoing longitudinal study of OA includes African American (nearly one-third of the cohort) and White participants living in a largely rural county in North Carolina. Civilian, noninstitutionalized residents age ≥45 years from 6 townships in Johnston County were enrolled between 1991 and 1998 (n = 3,187, original cohort), and additional residents were enrolled in 2003–2004 (n = 1,015, enrichment cohort). Since the enrichment cohort aimed to supplement the sample for African American and younger participants, participants enrolled during 2003–2004 tended to be younger (mean age 59.3 versus 65.8 years) and were more likely to be African American (40% versus 28%) than the original cohort participants at first follow-up (1999–2003); the 2 groups did not differ according to sex (21). Participants in the JoCo OA Project completed follow-up clinical and interview data collection approximately every 5 years, with 1,695 participants seen during the 2006–2010 clinic visit (time point T2). All participants in the JoCo OA Project have provided informed consent for participation, and the JoCo OA Project has been continuously approved by the institutional review boards of the University of North Carolina and the Centers for Disease Control and Prevention in Atlanta, Georgia.

Radiographic spine structure. Lumbar spine radiographs were included in the JoCo OA study for participants at the T2 2006–2010 clinic visit (n = 1,685). There were 819 returning participants at the T3 time point (2013–2015) who completed lumbar spine radiographs. By protocol, women of reproductive age (<50 years) were excluded from having lumbar spine radiographs. Lumbar spine radiographs were performed with the participant lying on his/her left side, a common position for clinical radiographs, with the central beam centered at the lumbar spine. The Burnett Atlas (22) was used to grade lumbar spine radiographic features of FOA, DSN, and OST. FOA was graded as absent or present at each lumbar level, while DSN and OST were graded in a semiquantitative fashion (0 = none, 1 = mild, 2 = moderate, and 3 = severe). Spine OA was defined as the presence of at least mild OST and mild DSN at the same vertebral level (23,24), which is similar to studies that define spine degeneration with the Kellgren/Lawrence (K/L) atlas (13,17). Spondylolisthesis was graded based on the translation of the vertebrae relative to the diameter of the affected intervertebral disc space, with 0 = no listhesis, 1 = ≤25%, 2 = 26–50%, 3 = 51–75%, 4 = 76–100%, and 5 = >100% translation. All lateral lumbar spine radiographs were graded at each lumbar level by an experienced single bone and joint radiologist (JBR) with an intrarater reliability of $\kappa = 0.73$ for FOA, 0.89 for DSN, and 0.90 for OST (25).

LBP. LBP was ascertained at the clinical interview by asking participants to answer "yes" or "no" to the question, "On most days do you have pain, aching or stiffness in your lower back?" Those participants who reported "yes" were also asked to quantify the severity of their symptoms as "mild," "moderate," or "severe."

Demographic and clinical characteristics. Demographic data were collected by clinical interview and examination, including age, sex, and race (African American/White). Clinical characteristics included self-reports of diabetes mellitus, high blood pressure, back injury, and history of cigarette smoking, as well as body mass index (BMI) at the time of interview (calculated from height measured without shoes and weight measured with a balance beam scale).

Appendicular joint OA. The protocol for conducting appendicular joint OA radiographs has been described in detail elsewhere (26,27). All knee, hip, and hand radiographs were read for K/L (28) score by the same radiologist (JBR). Interrater and intrarater reliability have been reported previously with a κ of 0.86 and 0.89, respectively, for both the hip and knee (26). Hip and knee OA for these analyses was defined as a K/L score of 2–4 in at least 1 extremity. Hand OA was defined as K/L grade of 2–4 in at least 1 distal interphalangeal of 1 extremity, with at least 2 other interphalangeal joints or carpometacarpal joints affected (K/L grade 2–4) across both hands (27).

Statistical analysis. Descriptive statistics were generated for the total sample. Incidence was defined as the absence of a specific radiographic feature at all levels of the lumbar spine at baseline and the presence of that feature at any level of the lumbar spine at follow-up. Worsening was defined as a ≥1-unit increase in severity from baseline to follow-up for OST, DSN, and spondylolisthesis. Incidence was measured for all radiographic features, whereas worsening was not measured for FOA since this was measured on a dichotomous scale. LBP was considered incident if the participant reported "no" LBP at baseline but "yes" at follow-up. LBP was considered worsening if there was a ≥1-unit increase in severity from baseline status to follow-up. Those with baseline severe LBP were excluded, as they were unable to have a 1-unit increase in symptoms. Since the prevalence varied for each spine feature at baseline, the incidence and worsening sample sizes varied accordingly (Figure 1).

Our outcomes were interval-censored because the exact time of occurrence was not observed, but rather only whether the event occurred between time points. In addition, follow-up intervals were of varying length for JoCo OA Project participants (on average, 5.5 years). Due to these features of our data, we selected Weibull models to estimate hazard ratios (HRs) accounting for potential predictors with corresponding 95% confidence intervals (95% CIs). All models reported are multivariable and included demographic variables (age, race, and sex), clinical characteristics (diabetes mellitus, high blood pressure, smoking status, and BMI), and appendicular joint OA (knee, hip, and hand OA) predictors simultaneously. When FOA was the outcome, we also included spine OA; likewise, when spine OA, DSN, or OST was the outcome, we included FOA. We explored several pairwise interaction terms between each BMI interval and diabetes mellitus and demographic (sex, race, and age) and clinical characteristics (BMI, smoking, and diabetes mellitus), but we did not identify any significant interactions. Similar to other studies, we conducted post hoc stratified analyses of the relationships between upper (L1–L3) and lower (L4–L5) lumbar levels for spine OA and FOA. Those results are provided in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24643. In addition, we compared differences in outcomes and predictors between participants in this study and those who were lost to follow-up (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24643). Finally, we conducted a simple sensitivity analysis to determine if our definition of spine OA would be

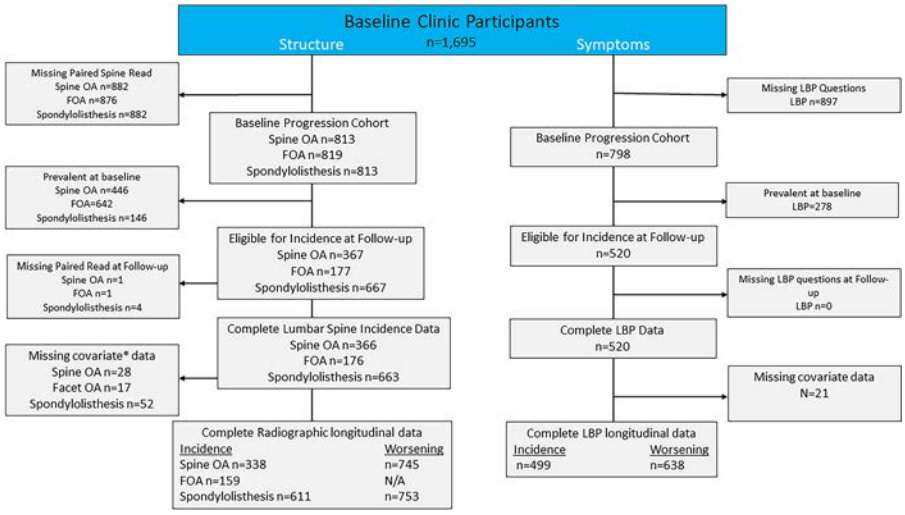

Figure 1. Flow diagram. FOA = facet joint osteoarthritis; LBP = low back pain; OA = osteoarthritis. * = Covariates were age, race, sex, cohort, body mass index, back pain, knee OA, hip OA, hand OA, diabetes mellitus, high blood pressure, and ever smoker.

influenced by the severity of OST and DSN. All analyses were conducted in SAS, version 9.4, and α was set at 0.05.

RESULTS

Figure 1 illustrates the selection of participants in the JoCo OA Project for both structure and symptoms for these analyses. For our structure outcomes, ~50% did not have lumbar spine radiographic data due to being lost to follow-up or failure to return for the clinic follow-up visit. A large proportion had baseline prevalent lumbar spine structural abnormality (54.9% with spine OA and 78.4% with FOA) or LBP (34.8%). Some were missing or refused lumbar spine radiographs at follow-up for lumbar spine structure (n = 1 for FOA, n = 1 for spine OA, and n = 6 for spondylolisthesis), and some had missing covariate data for lumbar spine structure (n = 28 for spine OA and spondylolisthesis, and n = 17 for FOA). For our symptom outcome of LBP, ~50% had missing LBP data due to being lost to follow-up or failure to return for the clinic follow-up visit. After missing covariate data (n = 21), 499 participants were eligible for incidence analysis, and 638 were eligible for analysis of worsening.

Table 1 describes the baseline demographic and clinical characteristics as well as appendicular joint OA and lumbar spine degenerative factors. Just over two-thirds (67.9%) were women, and 31.8% were African American. OST and FOA were common at baseline: 84.3% and 78.4%, respectively; DSN and spondylolisthesis occurred less frequently: 26.1% and 17.8%, respectively. A majority (54.7%) of the participants were obese, a small percentage (1.6%) reported a history of back injury, and ~35% reported the presence of LBP. The proportions of knee (39.6%), hip (35.5%), and hand OA (32.0%) were similar among participants.

The number of participants at risk for incident lumbar spine degeneration is described in Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24643. A large proportion of participants developed OST at any level of the lumbar spine (59.0%). The incidences of DSN and spine OA were similar at 38.1% and 31.4%, respectively. The incidences of any FOA and spondylolisthesis were 10.2% and 9.1%, respectively. Approximately 34% of participants had a worsening of OST, 24% a worsening of DSN, and 8% a worsening of spondylolisthesis.

Table 2 describes the demographic predictors for incidence of lumbar spine degeneration in multivariable models including demographic and clinical characteristics and OA variables as potential predictors simultaneously. Women were more likely to develop spondylolisthesis (HR 2.16 [95% CI 1.07–4.34]). Obesity was a risk factor for incidence of DSN (HR 1.80 [95% CI 1.09–2.98]), FOA (HR 4.99 [95% CI 1.46–17.10]), spine OA (HR 1.56 [95% CI 1.01–2.41]), and spondylolisthesis (HR 1.87 [95% CI 1.02–3.43]). The presence of diabetes mellitus was a protective factor for the incidence of DSN (HR 0.54 [95% CI 0.30–0.97]).

Table 1. Characteristics of participants with paired lumbar spine radiographs at baseline (n = 819)*

	No./ total no.	% (95% CI)
Radiographic spine outcomes		
OST grade 0 versus 1–3	691/819	84.3 (81.9–86.9)
Grade 0	128/819	15.6 (13.2–18.3)
Grade 1	511/819	62.4 (59.0–65.70
Grade 2	154/819	18.8 (16.2–21.7)
Grade 3	26/819	3.2 (2.1–4.6)
DSN grade 0 versus 1–3	579/819	26.1 (23.0–29.3)
Grade 0	240/819	29.3 (26.2–32.6)
Grade 1	364/819	44.4 (41.0–47.9)
Grade 2	214/819	26.1 (23.2–29.3)
Grade 3	1/819	0.12 (0.0–0.7)
Spondylolisthesis	146/813	17.8 (15.3–20.6)
Grade 0	673/813	82.2 (79.4–84.7)
Grade 1	139/813	17.0 (14.5–19.7)
Grade 2	5/813	0.61 (0.2–1.4)
Grade 3	2/813	0.24 (0.03–0.9)
Grade 4	0/813	NA
Grade 5	0/813	NA
Spine OA present	446/813	54.9 (51.4–58.3)
FOA present	642/819	78.4 (75.6–81.2)
Demographic predictors		
Age		
<55	34/819	4.2 (2.9–5.8)
55 to <65	371/819	45.3 (41.9–48.9)
65+	414/819	50.5 (47.1–54.0)
Sex, female	556/819	67.9 (65.0–71.0)
Race, African American	260/819	31.8 (28.6–35.1)
Clinical characteristic predictors		
Obesity (BMI ≥30)	448/819	54.7 (51.3–8.1)
Diabetes mellitus	160/795	20.1 (17.3–23.0)
High blood pressure	499/796	62.7 (59.3–66.1)
Smoking	410/812	50.5 (47.0–54.0)
Back injury	13/811	1.6 (1.0–2.0)
Back pain	278/798	34.8 (31.5–38.1)
None	520/798	65.2 (61.0–69.1)
Mild	96/798	12.0 (9.5–15.1)
Moderate/severe	182/798	22.8 (19.5–26.6)
Appendicular joint OA predictors		
Knee	318/803	39.6 (36.2–43.0)
Hip	287/809	35.5 (32.2–38.8)
Hand	260/813	32.0 (28.8–35.3)

* 95% CI = 95% confidence interval; BMI = body mass index; DSN = disc space narrowing; FOA = facet joint osteoarthritis; NA = not applicable; OA = osteoarthritis; OST = osteophytes.

Smokers were 39% less likely to develop DSN (HR 0.61 [95% CI 0.38–0.98]) and 40% less likely to develop spine OA (HR 0.60 [95% CI 0.39–0.91]). Knee OA was a risk factor for the incidence of FOA (HR 4.18 [95% CI 1.44–12.2]). However, hip and hand OA were not risk factors for radiographic changes in the lumbar spine. In our sensitivity analysis of spine OA coding, we did not identify that the severity of OST or DSN changed the results.

Table 3 describes the predictors for worsening of lumbar spine degeneration in multivariable models including demographic and clinical characteristics and OA variables simultaneously as potential predictors. Women had a 46% increased hazard of

Table 2. Baseline demographic and clinical characteristics and appendicular joint OA as predictors of incident lumbar spine degeneration in Weibull models*

	OST (n = 114)	DSN (n = 220)	Spine OA (n = 338)	FOA (n = 159)	Spondylolisthesis (n = 611)
Demographic characteristics					
Women versus men	2.31 (0.88–6.04)	1.33 (0.79–2.23)	1.26 (0.78–2.04)	0.50 (0.14–1.75)	2.16 (1.07–4.34)†
African American versus White	1.19 (0.60–2.37)	0.69 (0.41–1.17)	0.81 (0.52–1.28)	0.96 (0.29–3.23)	1.20 (0.66–2.18)
Age 55–65 versus <55 years	0.41 (0.11–1.58)	1.05 (0.34–3.23)	2.51 (0.75–8.42)	0.69 (0.03–14.9)	0.67 (0.19–2.41)
Age 65+ versus <55 years	0.38 (0.09–1.55)	1.24 (0.37–4.10)	2.17 (0.62–7.58)	1.09 (0.05–26.1)	0.96 (0.26–3.59)
Clinical characteristics					
Obesity	0.63 (0.31–1.27)	1.80 (1.09–2.98)†	1.56 (1.01–2.41)†	4.99 (1.46–17.1)†	1.87 (1.02–3.43)†
Diabetes mellitus	0.51 (0.17–1.57)	0.54 (0.30–0.97)†	0.71 (0.42–1.18)	0.82 (0.20–3.35)	0.69 (0.34–1.39)
High blood pressure	0.80 (0.42–1.50)	0.76 (0.48–1.21)	0.81 (0.54–1.22)	0.51 (0.16–1.56)	1.26 (0.69–2.29)
Smoking	0.87 (0.49–1.56)	0.61 (0.38–0.98)†	0.60 (0.39–0.91)†	1.04 (0.35–3.11)	0.79 (0.46–1.38)
Back injury	0.57 (0.07–4.59)	0.56 (0.07–4.43)	0.72 (0.17–3.02)	NR	1.82 (0.24–14.0)
Appendicular joint OA					
Knee OA	1.28 (0.62–2.65)	0.92 (0.54–1.57)	0.95 (0.62–1.47)	4.18 (1.44–12.2)†	0.88 (0.50–1.53)
Hip OA	0.61 (0.34–1.10)	0.80 (0.48–1.34)	0.91 (0.59–1.38)	0.41 (0.13–1.27)	1.28 (0.74–2.22)
Hand OA	1.02 (0.56–1.84)	1.16 (0.68–1.97)	1.31 (0.84–2.02)	0.58 (0.16–2.09)	1.20 (0.66–2.19)

* Values are the hazard ratio (95% confidence interval). Weibull models include sex, race, age, body mass index, diabetes mellitus, high blood pressure, smoking, and back injury, as well as knee, hip, and hand osteoarthritis (OA) predictors simultaneously. Additionally, disc space narrowing (DSN) and spine OA models included facet joint OA (FOA) presence, and FOA models included spine OA presence. NR = not reported (due to small sample size); OST = osteophytes.
† Significant.

worsening of DSN (HR 1.46 [95% CI 1.02–2.08]). Obesity was a risk factor for worsening of DSN (HR 1.51 [95% CI 1.09–2.09]) but a protective factor for vertebral OST worsening (HR 0.75 [95% CI 0.57–0.99]). The presence of diabetes mellitus was a risk factor for the worsening of vertebral OST (HR 1.38 [95% CI 1.00–1.92]).

Table 4 describes the predictors for incidence and worsening of LBP in multivariable models including demographic and clinical characteristics and OA variables simultaneously as potential predictors. Obesity was a strong risk factor of both LBP incidence (HR 1.75 [95% CI 1.19–2.56]) and worsening (HR 1.51 [95% CI 1.12–2.06]). Having mild or moderate LBP at baseline was a predictor for LBP worsening. Participants with hip OA at baseline were 39% more likely to progress in LBP severity (HR 1.39 [95% CI 1.04–1.85]). Baseline OST and spine OA were predictors of incident LBP (HR 1.85 [95% CI 1.02–3.36]; HR 1.80 [95% CI 1.24–2.63], respectively). The presence of vertebral OST at baseline was also a predictor of worsening LBP (HR 1.58 [95% CI 1.00–2.49]).

The results for the post hoc stratified analysis of the incidence of upper and lower lumbar spine OA and FOA are described in Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24643. Findings were similar to our primary analysis of the lumbar spine degeneration features. In addition, African

Table 3. Baseline demographic and clinical characteristics and appendicular joint OA as predictors of lumbar spine degeneration worsening in Weibull models*

	OST (n = 730)	DSN (n = 754)	Increased levels of spine OA (n = 745)	Spondylolisthesis (n = 753)
Demographic characteristics				
Women versus men	0.85 (0.63–1.13)	1.46 (1.02–2.08)†	1.22 (0.92–1.62)	1.93 (0.96–3.87)
African American versus White	1.04 (0.77–1.40)	0.79 (0.55–1.13)	0.91 (0.69–1.21)	1.19 (0.66–2.13)
Age 55–65 versus <55 years	0.84 (0.41–1.74)	1.40 (0.55–3.60)	2.95 (1.07–8.15)†	0.67 (0.19–2.40)
Age 65+ versus <55 years	0.89 (0.43–1.88)	1.29 (0.49–3.38)	3.01 (1.08–8.42)†	0.90 (0.24–3.33)
Clinical characteristics				
Obesity	0.75 (0.57–0.99)†	1.51 (1.09–2.09)†	1.24 (0.96–1.62)	1.77 (0.97–3.22)
Diabetes mellitus	1.38 (1.00–1.92)†	0.85 (0.56–1.28)	0.77 (0.56–1.08)	0.81 (0.40–1.62)
High blood pressure	0.88 (0.67–1.16)	0.88 (0.64–1.20)	1.13 (0.87–1.47)	1.15 (0.64–2.04)
Smoking	1.00 (0.76–1.30)	1.08 (0.79–1.47)	0.95 (0.74–1.23)	0.77 (0.44–1.35)
Back injury	0.58 (0.18–1.83)	0.61 (0.15–2.49)	0.59 (0.19–1.85)	1.41 (0.19–10.6)
Appendicular joint OA				
Knee OA	0.97 (0.74–1.28)	0.84 (0.61–1.16)	1.00 (0.78–1.30)	0.83 (0.48–1.45)
Hip OA	0.91 (0.69–1.19)	0.72 (0.52–1.00)	0.77 (0.60–1.01)	1.11 (0.64–1.91)
Hand OA	0.98 (0.73–1.32)	1.06 (0.75–1.48)	1.10 (0.84–1.43)	1.05 (0.58–1.90)

* Values are the hazard ratio (95% confidence interval). Weibull models include sex, race, age, body mass index, diabetes mellitus, high blood pressure, smoking, and back injury, as well as knee, hip, and hand osteoarthritis (OA) predictors simultaneously. DSN = disc space narrowing; OST = osteophytes.
† Significant.

Table 4. Baseline demographic and clinical characteristics, appendicular joint OA, and lumbar spine degeneration as predictors of incident and worsening low back pain in Weibull models*

	Low back pain incidence (n = 509)†	Low back pain worsening (n = 652)‡
Demographic characteristics		
Women versus men	1.19 (0.79–1.79)	1.35 (0.97–1.87)
African American versus White	0.66 (0.43–1.01)	0.81 (0.58–1.12)
Age 55–65 versus <55 years	0.70 (0.26–1.85)	0.68 (0.33–1.43)
Age 65+ versus <55 years	0.55 (0.20–1.50)	0.66 (0.31–1.42)
Clinical characteristics		
Obesity	1.75 (1.19–2.56)§	1.51 (1.12–2.06)§
Diabetes mellitus	0.68 (0.40–1.16)	0.92 (0.64–1.32)
High blood pressure	1.33 (0.91–1.95)	1.30 (0.96–1.76)
Smoking	0.69 (0.47–1.01)	0.83 (0.62–1.11)
Low back injury	1.85 (0.42–8.10)	2.39 (0.69–8.27)
Mild low back pain to moderate and severe	NR	5.35 (3.62–7.92)§
Moderate low back pain to severe	NR	2.26 (1.55–3.29)§
Appendicular joint OA		
Knee	1.40 (0.97–2.02)	1.15 (0.86–1.53)
Hip	1.08 (0.75–1.57)	1.39 (1.04–1.85)§
Hand	1.19 (0.79–1.77)	0.97 (0.71–1.33)
Spine degeneration		
OST	1.85 (1.02–3.36)§	1.58 (1.00–2.49)§
DSN	1.52 (0.99–2.33)	1.11 (0.80–1.53)
Spine OA	1.80 (1.24–2.63)§	1.27 (0.95–1.70)
FOA	1.06 (0.69–1.63)	1.05 (0.73–1.51)
Spondylolisthesis	1.14 (0.71–1.82)	1.12 (0.78–1.62)

* Values are the hazard ratio (95% confidence interval). Weibull models include sex, race, age, body mass index, diabetes mellitus, high blood pressure, smoking, and back injury, as well as knee, hip, and hand osteoarthritis (OA) predictors simultaneously. DSN = disc space narrowing; FOA = facet joint OA; NR = not reported (due to small sample size); OST = osteophytes.
† None to mild or greater.
‡ One-unit change from baseline; excludes severe at baseline.
§ Significant.

American participants were 59% less likely (HR 0.41 [95% CI 0.24–0.72]) to have incident upper lumbar spine OA with symptoms and 53% less likely (HR 0.47 [95% CI 0.28–0.79]) to have lower FOA. Those with self-reported high blood pressure were 59% more likely (HR 1.59 [95% CI 1.01–2.51]) to have incident upper spine OA.

DISCUSSION

We determined the incidence and worsening of radiographic lumbar spine features commonly used in clinical practice to help better delineate the relationship between spine structure and LBP. Obesity was a strong predictor for nearly all lumbar spine features and LBP. OST and spine OA had prognostic value as demonstrated by their prediction of LBP incidence and worsening. These findings may have important implications for clinical practice, especially where diagnostic imaging is being used for better understanding the relationship between lumbar spine degeneration and LBP.

We identified that obesity was a predictor of both the incidence and worsening of LBP. Muraki and colleagues (17) did not find obesity to be a significant predictor of incident LBP; however, their Japanese study population had a considerably lower average BMI compared to ours. Since our cohort had a large proportion of participants considered to be obese, these findings may not be generalizable to other community-based studies with a much lower proportion of obese participants. For radiographic features commonly examined during initial visits for LBP, we found that spine OA was a significant risk factor for incident LBP, similar to findings of others (17). The baseline presence of at least mild OST was a significant predictor of LBP incidence and worsening. Although OST is common among the population, the presence of OST among those with LBP may be an indicator of continued or worse mechanical LBP. We also identified hip OA as a significant predictor of LBP worsening. Some have supported a hip-spine syndrome whereby influences from the presence of hip symptoms or OA have resulted in LBP (29–34). To our knowledge, this is the first study to report a longitudinal relationship between baseline hip OA and LBP worsening.

Obesity was a strong and consistent predictor for the incidence of DSN, spine OA, FOA, and spondylolisthesis, and only for worsening of DSN. Muraki et al (17) found that BMI was a significant risk factor for lumbar spine degeneration. However, our study also includes lumbar spine FOA and spondylolisthesis, which were not included in their study. Our findings support cross-sectional studies from both our group and others that indicate an association between obesity and FOA (7,14,35). Due to the older age of our sample in the JoCo OA Project, >75% of participants already had FOA at baseline. As such, the small number of incident cases for FOA limits these findings. The finding that obesity was a protective factor for the worsening of OST is difficult to explain. One reason may be related to the large number of prevalent OST at baseline that limited the number at risk during follow-up. As such, this sample may be a more resistant group to OST development. Similar to our findings, a Framingham Study analysis (36) determined a higher proportion of women who developed spondylolisthesis.

We identified some interesting relationships between diabetes mellitus and smoking and the incidence and worsening of DSN. Smoking has been associated with intervertebral disc disease in twin studies (37), while other community-based studies (17–19) have found no significant association with spine degeneration. Our study, however, identified that smokers were significantly less likely to have incident DSN. It has been suggested that inadequate statistical control of BMI may be 1 factor that leads to this inverse association; however, this was not the case in our study. The inverse association of smoking with DSN is concordant with the inverse association of smoking with knee OA

found in many studies (38,39). Self-reported presence of diabetes mellitus was identified as a significant risk factor for worsening of OST. This may be related to impaired glucose tolerance, which has been linked to ankylosing hyperostosis in the spine (40). However, we also identified self-reported diabetes mellitus as inversely associated with incidence of DSN. We explored any potential differences between those participants self-reporting insulin use for diabetes mellitus treatment, based on the study by Shirinsky et al (41). In this prior study, they found that medication-treated diabetes mellitus was protective for knee OA progression. However, we did not identify any significant differences between these 2 groups. It is possible that other antidiabetic medications, such as metformin, which has both antioxidant and antisenolytic effects, might help explain the inverse association of diabetes mellitus and lumbar spine degeneration. The relationship between diabetes mellitus and DSN does appear to be focused more in the upper lumbar spine than the lower lumbar spine. A biologic rationale is not known for the BMI-independent inverse association of diabetes mellitus and smoking with DSN but may be related to potential underlying nicotine effects or cellular mechanisms (39). In addition, there may be other nonbiological reasons that could explain these findings. We did not adjust our findings for multiple comparisons, and several of these findings are nearly statistically significant by our established threshold. In addition, while our models included several potential predictors simultaneously, we did not include all potential confounders that might influence our findings, and therefore we cannot rule out the potential for residual confounding.

Our study has several strengths but is not without limitations. Lateral lumbar spine radiographs could lead to nondifferential misclassification of FOA status, since lateral views may underestimate the occurrence of FOA. However, prevalence estimates of FOA based on lateral spine radiography (7) are similar to those previously reported based on computed tomography scans (35). The JoCo OA Project protocol excluded women of childbearing age from having lumbar spine radiographs to prevent unnecessary radiation exposure; therefore, the results may not be generalizable to this subgroup. While the average length of follow-up of 5.5 years is ideal for examining degeneration, it limits the examination of incident and worsening of LBP since it may be quite variable over a long period. We measured the presence and severity of LBP but not how LBP interfered with daily activity. In addition, our question for LBP also includes pain, aching, and stiffness, which may overestimate the true incidence of LBP since stiffness may be present without pain. We did not include lower back–specific functional measures or account for previous LBP episodes, widespread pain, or psychosocial factors that are known to be associated with incident or progressive LBP (42). Our study had >40% loss to follow-up of cohort participants, which may limit generalizability since those lost to follow-up were more likely to be older and to have a BMI of <30, self-reported diabetes mellitus, high blood pressure, and appendicular joint OA (knee, hip, and hand OA). However, the primary reason for loss to follow-up was participant death. The loss to follow-up we experienced over this time frame may influence the direction and strength of our estimates relative to the true population values. This may be the case for some estimates without clear statistical significance such as the relationship we identified between DSN and diabetes mellitus. In addition, there was a high prevalence of OST and FOA among participants based upon the average age of the cohort. As such, future studies should consider cohorts with younger participants to enhance generalizability.

In conclusion, our study is unique in that it provides estimates of lumbar spine incidence and worsening of radiographic features commonly used in clinical practice to examine spine health. The predictors identified in this study are commonly used in routine primary care for participants with LBP. The finding that obesity, existing lumbar spine degeneration, and appendicular hip OA are predictive of future LBP may assist in the ongoing efforts to decrease LBP in the community.

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. Goode 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. Goode, Cleveland, George, Schwartz, Kraus, Gracely, Jordan, Golightly.

Acquisition of data. Goode, Cleveland, George, Schwartz, Kraus, Gracely, Jordan, Golightly.

Analysis and interpretation of data. Goode, Cleveland, George, Schwartz, Kraus, Renner, DeFrate, Hu, Jordan, Golightly.

REFERENCES

1. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 1994;331:69–73.
2. Freburger JK, Holmes GM, Agans RP, Jackman AM, Darter JD, Wallace AS, et al. The rising prevalence of chronic low back pain. Arch Intern Med 2009;169:251–8.
3. Jarvik JG, Comstock BA, James KT, Avins AL, Bresnahan BW, Deyo RA, et al. Lumbar imaging with reporting of epidemiology (lire): protocol for a pragmatic cluster randomized trial. Contemp Clin Trials 2015;45:157–63.
4. Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin CC, Chenot JF, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. Eur Spine J 2018;27:2791–803.
5. Chou R, Qaseem A, Snow V, Casey D, Cross JT Jr, Shekelle P, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med 2007;147:478–91.
6. Deyo RA, Jarvik JG, Chou R. Low back pain in primary care. BMJ 2014;349:g4266.
7. Goode AP, Marshall SW, Renner JB, Carey TS, Kraus VB, Irwin DE, et al. Lumbar spine radiographic features and demographic, clinical, and radiographic knee, hip, and hand osteoarthritis. Arthritis Care Res (Hoboken) 2012;64:1536–44.

8. Pye SR, Reid DM, Smith R, Adams JE, Nelson K, Silman AJ, et al. Radiographic features of lumbar disc degeneration and self-reported back pain. J Rheumatol 2004;31:753–8.
9. De Schepper EI, Damen J, van Meurs JB, Ginai AZ, Popham M, Hofman A, et al. The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. Spine (Phila Pa 1976) 2010;35:531–6.
10. Raastad J, Reiman M, Coeytaux R, Ledbetter L, Goode AP. The association between lumbar spine radiographic features and low back pain: a systematic review and meta-analysis. Semin Arthritis Rheum 2015;44:571–85.
11. Goode AP, Carey TS, Jordan JM. Low back pain and lumbar spine osteoarthritis: how are they related? Curr Rheumatol Rep 2013; 15:305.
12. Gellhorn AC, Katz JN, Suri P. Osteoarthritis of the spine: the facet joints. Nat Rev Rheumatol 2013;9:216–24.
13. Muraki S, Oka H, Akune T, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of radiographic lumbar spondylosis and its association with low back pain in elderly subjects of population-based cohorts: the ROAD study. Ann Rheum Dis 2009;68:1401–6.
14. Deyo RA. Cascade effects of medical technology. Annu Rev Public Health 2002;23:23–44.
15. Pye SR, Reid DM, Adams JE, Silman AJ, O'Neill TW. Influence of weight, body mass index and lifestyle factors on radiographic features of lumbar disc degeneration. Ann Rheum Dis 2007;66:426–7.
16. Pye SR, Reid DM, Lunt M, Adams JE, Silman AJ, O'Neill TW. Lumbar disc degeneration: association between osteophytes, end-plate sclerosis and disc space narrowing. Ann Rheum Dis 2007;66:330–3.
17. Muraki S, Akune T, Oka H, Ishimoto Y, Nagata K, Yoshida M, et al. Incidence and risk factors for radiographic lumbar spondylosis and lower back pain in Japanese men and women: the ROAD study. Osteoarthritis Cartilage 2012;20:712–8.
18. Hassett G, Hart DJ, Manek NJ, Doyle DV, Spector TD. Risk factors for progression of lumbar spine disc degeneration: the Chingford Study. Arthritis Rheum 2003;48:3112–7.
19. Symmons DP, van Hemert AM, Vandenbroucke JP, Valkenburg HA. A longitudinal study of back pain and radiological changes in the lumbar spines of middle aged women. I. Clinical findings. Ann Rheum Dis 1991;50:158–61.
20. Jordan JM, Helmick CG, Renner JB, Luta G, Dragomir AD, Woodard J, et al. Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol 2007; 34:172–80.
21. Allen KD, Chen JC, Callahan LF, Golightly YM, Helmick CG, Renner JB, et al. Associations of occupational tasks with knee and hip osteoarthritis: the Johnston County Osteoarthritis Project. J Rheumatol 2010;37:842–50.
22. Burnett S, Hart DJ, Cooper C, Spector TD. A Radiographic atlas of osteoarthritis. London: Springer-Verlag; 1994.
23. Goode AP, Cleveland RJ, George SZ, Kraus VB, Schwartz TA, Gracely RH, et al. Different phenotypes of osteoarthritis in the lumbar spine reflected by demographic and clinical characteristics: the Johnston County Osteoarthritis Project. Arthritis Care Res (Hoboken) 2020;72:974–81.
24. Nelson AE, Renner JB, Schwartz TA, Kraus VB, Helmick CG, Jordan JM. Differences in multijoint radiographic osteoarthritis phenotypes among African Americans and Caucasians: the Johnston County Osteoarthritis project. Arthritis Rheum 2011;63:3843–52.
25. Goode AP, Nelson AE, Kraus VB, Renner JB, Jordan JM. Biomarkers reflect differences in osteoarthritis phenotypes of the lumbar spine: the Johnston County Osteoarthritis Project. Osteoarthritis Cartilage 2017;25:1672–9.
26. Jordan JM, Linder GF, Renner JB, Fryer JG. The impact of arthritis in rural populations. Arthritis Care Res 1995;8:242–50.
27. Kraus VB, Jordan JM, Doherty M, Wilson AG, Moskowitz R, Hochberg M, et al. The Genetics of Generalized Osteoarthritis (GOGO) study: study design and evaluation of osteoarthritis phenotypes. Osteoarthritis Cartilage 2007;15:120–7.
28. Kellgren JH. The epidemiology of rheumatic diseases. Ann Rheum Dis 1964;23:109–22.
29. Hicks GE, Sions JM, Velasco TO. Hip symptoms, physical performance, and health status in older adults with chronic low back pain: a preliminary investigation. Arch Phys Med Rehabil 2018;99:1273–8.
30. Weiner DK, Fang M, Gentili A, Kochersberger G, Marcum ZA, Rossi MI, et al. Deconstructing chronic low back pain in the older adult: step by step evidence and expert-based recommendations for evaluation and treatment. Part I. Hip osteoarthritis. Pain Med 2015; 16:886–97.
31. Sembrano JN, Polly DW Jr. How often is low back pain not coming from the back? Spine (Phila Pa 1976) 2009;34:E27–32.
32. Ben-Galim P, Ben-Galim T, Rand N, Haim A, Hipp J, Dekel S, et al. Hip-spine syndrome: the effect of total hip replacement surgery on low back pain in severe osteoarthritis of the hip. Spine (Phila Pa 1976) 2007;32:2099–102.
33. Hsieh PH, Chang Y, Chen DW, Lee MS, Shih HN, Ueng SW. Pain distribution and response to total hip arthroplasty: a prospective observational study in 113 patients with end-stage hip disease. J Orthop Sci 2012;17:213–8.
34. Chimenti PC, Drinkwater CJ, Li W, Lemay CA, Franklin PD, O'Keefe RJ. Factors associated with early improvement in low back pain after total hip arthroplasty: a multi-center prospective cohort analyses. J Arthroplasty 2016;31:176–9.
35. Kalichman L, Li L, Kim DH, Guermazi A, Berkin V, O'Donnell CJ, et al. Facet joint osteoarthritis and low back pain in the community-based population. Spine (Phila Pa 1976) 2008;33:2560–5.
36. Kauppila LI, Eustace S, Kiel DP, Felson DT, Wright AM. Degenerative displacement of lumbar vertebrae: a 25-year follow-up study in Framingham. Spine (Phila Pa 1976) 1998;23:1868–73.
37. Battie MC, Videman T, Gill K, Moneta GB, Nyman R, Kaprio J, et al. 1991 Volvo Award in clinical sciences. Smoking and lumbar intervertebral disc degeneration: an MRI study of identical twins. Spine (Phila Pa 1976) 1991;16:1015–21.
38. Kong L, Wang L, Meng F, Cao J, Shen Y. Association between smoking and risk of knee osteoarthritis: a systematic review and meta-analysis. Osteoarthritis Cartilage 2017;25:809–16.
39. Felson DT, Zhang Y. Smoking and osteoarthritis: a review of the evidence and its implications. Osteoarthritis Cartilage 2015;23:331–3.
40. Crisp AJ, Heathcote JG. Connective tissue abnormalities in diabetes mellitus. J R Coll Physicians Lond 1984;18:132–41.
41. Shirinsky IV, Shirinsky VS. Effects of medication-treated diabetes on incidence and progression of knee osteoarthritis: a longitudinal analysis of the Osteoarthritis Initiative data. Rheumatol Int 2017; 37:983–91.
42. Van den Berg R, Chiarotto A, Enthoven WT, de Schepper E, Oei EH, Koes BW, et al. Clinical and radiographic features of spinal osteoarthritis predict long-term persistence and severity of back pain in older adults. Ann Phys Rehabil Med 2022;65:101427.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1667–1675
DOI 10.1002/acr.24580

American College of Rheumatology
Empowering Rheumatology Professionals

Coexistence of Low Back Pain and Lumbar Kyphosis and the Association With Increased Functional Disability in Knee Osteoarthritis: Results From a Population-Based Cohort

Masashi Taniguchi,[1] Tome Ikezoe,[1] Mitsuhiro Masaki,[2] Tsukasa Kamitani,[1] Tadao Tsuboyama,[3] Hiromu Ito,[1] Shuichi Matsuda,[1] Yasuharu Tabara,[1] Fumihiko Matsuda,[1] and Noriaki Ichihashi,[1] on behalf of the Nagahama Study Group

Objective. To examine the association of low back pain (LBP) and lumbar kyphosis with functional disabilities and knee symptoms in patients with knee osteoarthritis (OA).

Methods. We analyzed 586 participants (80.1% female; mean ± SD age 68.8 ± 5.2 years) from the Nagahama Study who were age ≥60 years and had radiographically confirmed knee OA. The Knee Society Knee Scoring System (KSS) was used to assess functional disabilities and knee symptoms. LBP was defined as the presence of any persistent back pain for more than 3 months. Lumbar kyphosis was determined by skin-surface methods using a computer-aided electronic device called the Spinal Mouse. Multiple linear regression analysis was used for assessing the association of LBP and lumbar kyphosis with the KSS scores. Subgroup analyses based on sex were also performed.

Results. LBP and lumbar kyphosis were independently associated with a lower KSS function score after adjustment for covariates (mean difference –4.96 [95% confidence interval (95% CI) –7.56, –2.36] points for LBP alone, mean difference –4.47 [95% CI –8.51, –0.43] points for lumbar kyphosis alone, and mean difference –13.86 [95% CI –18.86, –8.86] points for the coexistence of LBP and lumbar kyphosis, respectively). The coexistence of LBP and lumbar kyphosis in women was associated with a lower KSS symptom score (mean difference –4.49 [95% CI –6.42, –2.55] points).

Conclusion. These findings suggest that both LBP and lumbar kyphosis are useful clinical signals indicating functional disability and knee symptoms in patients with knee OA.

INTRODUCTION

Knee osteoarthritis (OA) is a common musculoskeletal disorder and is the leading cause of knee pain and disability (1). A recent systematic review indicated that a greater burden of comorbidity was associated with functional disabilities and severe knee symptoms in patients with knee OA than in those without (2). In particular, low back pain (LBP) is a major comorbid condition in patients with knee OA (3). Both knee pain and LBP have potential deteriorative effects on functional disability (4).

Standing posture with decreased lumbar lordosis (i.e., lumbar kyphosis) and increased knee flexion are known as the knee-spine syndrome (5), and lumbar kyphosis is mostly linked to spinal stenosis (6). This malalignment may cause not only issues with the spine, but also knee joint overload via the kinetic chain. In community-dwelling older adults, age-related kyphosis, especially the lack of lumbar lordosis, reportedly causes impaired physical function, including reduced gait speed and quadriceps strength (7). Thus, increased mechanical stress of knee and spine joints and its related

Supported by a university grant, the Center of Innovation Program, the Global University Project, and a Grant-in-Aid for Scientific Research (grants 25293141, 26670313, 26293198, 17H04182, 17H04126, 17H04123, 18K18450, and 19K17634) from the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Practical Research Project for Rare/Intractable Diseases (grants ek0109070, ek0109070, ek0109196, and ek0109348); the Comprehensive Research on Aging and Health Science Research Grants for Dementia R&D (grants dk0207006 and dk0207027); the Program for an Integrated Database of Clinical and Genomic Information (grant kk0205008); the Practical Research Project for Lifestyle-Related Diseases including cardiovascular diseases and diabetes mellitus (grants 17ek0210066, 18ek0210096, and 19ek0210116); the Research Program for Health Behavior Modification by Utilizing IoT (grants le0110005 and le0110013) from the Japan Agency for Medical Research and Development (AMED); the Takeda Medical Research Foundation, the Mitsubishi Foundation, the Daiwa Securities Health Foundation, and the Sumitomo Foundation; the Comprehensive Research on Aging and Health Science Research Grants for Dementia R&D from Japan AMED (grants H26-Choju-Ippan-001, 15dk0107007h0003, and 16dk0110007h0003); and the JSPS KAKENHI Grant in Aid for Research Activity Start-Up (grant 19K21493).

[1]Masashi Taniguchi, RPT, PhD, Tome Ikezoe, RPT, PhD, Tsukasa Kamitani, RPT, DrPH, Hiromu Ito, MD, PhD, Shuichi Matsuda, MD, PhD, Yasuharu Tabara, PhD, Fumihiko Matsuda, PhD, Noriaki Ichihashi, RPT, PhD: Kyoto University, Kyoto, Japan; [2]Mitsuhiro Masaki, RPT, PhD: Takasaki University of Health and Welfare, Kyoto, Japan; [3]Tadao Tsuboyama, MD, PhD: Kyoto University and Bukkyo University, Kyoto, Japan.

Address correspondence to Masashi Taniguchi, RPT, PhD, Department of Physical Therapy, Human Health Sciences, Graduate School of Medicine, Kyoto University, 53-Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan. Email: taniguchi.masashi.7a@kyoto-u.ac.jp.

Submitted for publication July 6, 2020; accepted in revised form February 12, 2021.

SIGNIFICANCE & INNOVATIONS

- Lumbar kyphosis in participants with knee osteoarthritis (OA) was associated with functional disabilities in patients with knee OA.
- The coexistence of low back pain (LBP) and lumbar kyphosis was remarkably associated with functional disabilities in patients with knee OA.
- The coexistence of LBP and lumbar kyphosis had adverse associations on knee symptoms in women.
- These findings suggest that both LBP and lumbar kyphosis are useful clinical signs indicating functional disability and knee symptoms in patients with knee OA.

dysfunction may also be associated with functional disabilities in knee OA. Additionally, previous studies (8,9) reported that degenerative changes with decreasing lumbar lordosis and the burden of restriction in activities were more common in women than in men. However, sex differences in the association of lumbar kyphosis and/or LBP with functional disability and knee symptoms have not been well investigated in knee OA patients.

Generally, decreased lumbar lordosis is related to the presence of LBP in middle-aged and older adults (10). In contrast, Wang et al (11) reported no differences in lumbar lordosis in patients with knee OA with or without LBP. Decreased lumbar lordosis in patients with knee OA could be caused by a compensatory strategy against knee flexion contracture and knee symptoms, regardless of the presence of LBP. Therefore, we assumed that LBP and lumbar kyphosis independently contributed to functional disabilities and severe knee symptoms in patients with knee OA. However, in recent literature, individual or coexistent associations of LBP and lumbar kyphosis with functional disabilities and knee symptoms in patients with knee OA remains unclear.

The primary objective of this study was to determine the association of LBP and lumbar kyphosis with functional disability and knee symptoms in patients with knee OA. The secondary objective was to determine the difference in these associations between the sexes. We hypothesized that the coexistence of LBP and lumbar kyphosis could be associated with functional disabilities and worse knee symptoms. We also hypothesized that those associations were stronger in women than in men. These findings will provide valuable information for the effective management of functional disability and knee symptoms in patients with knee OA who also have spinal problems.

PATIENTS AND METHODS

Study participants and selection. We analyzed the data set of the Nagahama Prospective Cohort for Comprehensive Human Bioscience (Nagahama Study), which was a population-based cohort. Our study was a cross-sectional analysis of the baseline measurements obtained between 2013 and 2016 from the general population of Nagahama City, which comprises 125,000 inhabitants located in a predominantly rural area of the Shiga Prefecture of central Japan (12). Community residents ages 30–74 years at recruitment, living independently without serious health problems, were recruited via mass communications in the local community, such as public-relations magazines and newspapers, and personal solicitations. From a total of 9,850 individuals in this cohort, we selected those older than 60 years who participated in the optional physical assessment, which included knee radiography and a sagittal spinal alignment evaluation. Among these individuals, we selected participants who had unilateral or bilateral knee OA on radiographs and could walk for more than 10 meters with or without a cane, after excluding those who meet the following exclusion criteria: acute LBP, rheumatoid arthritis, central nervous system impairments, and surgical history of the herniated intervertebral disk, spinal canal stenosis, and other lower-extremity joint diseases.

All study procedures were approved by the Ethics Committee of the Kyoto University Graduate School of Medicine and by the Nagahama Municipal Review Board. The study was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent for the use of data was obtained from all participants in the Nagahama Study.

Definition of knee OA. Participants who had unilateral or bilateral knee OA on radiographs were included in the present study. The tibiofemoral joints of both knees were evaluated using weight-bearing anteroposterior radiographs. Two experienced orthopedists, who were blinded regarding each patient's clinical status, evaluated each knee based on the Kellgren/Lawrence (K/L) grading system, and radiographic knee OA was defined as a K/L grade of ≥2 (13). If the grades were different between the two examiners, a third examiner evaluated and determined the final grade. This study classified the severity of radiographic knee OA as follows: either knee with a K/L grade of 2, both knees with a K/L grade of 2, either knee with a K/L grade of 2 and the other ≥3, and both knees with a K/L grade of ≥3.

Knee Society Knee Scoring System (KSS) functions and symptoms. Self-reported functional disability and knee symptoms were evaluated using the new KSS scoring system (Japanese edition, 2011) (14). The 2011 KSS is a self-administered outcome measurement tool that consists of 4 subcategories, of which 2, namely "functional activities" (0–100 points) and "symptoms" (0–25 points), were used in this study. The functional activities category further consists of 4 components, including walking and standing (30 points), standard activities (30 points), advanced activities (25 points), and discretionary activities (15 points). The KSS symptom score relies on the degree of knee pain during walking and going up and down stairs, and stiffness. For these 2 subcategories, lower scores indicate

functional disabilities and poorer symptoms. The validity of the 2011 KSS in the Japanese population has previously been established (14).

Lumbar lordosis and LBP. Sagittal spinal alignment during quiet standing was measured using a Spinal Mouse (Index). The Spinal Mouse is a computer-aided electronic device that noninvasively measures intersegmental angles by tracing along the midline of the spine between the spinous process from C7 to S3. The measurement of spinal alignment by the Spinal Mouse has been validated with radiographic measurements, indicating a high correlation between these measurements for lumbar lordosis angle (r = 0.794) (15). The same examiner performed all measurements of lumbar lordosis angle. The lumbar lordosis angle was measured using the Spinal Mouse, which demonstrated high intra-rater reliability (intraclass correlation coefficient: 0.97). The lumbar lordosis was calculated as the sum of 6 segmental angles between Th12/L1 and L5/S1, and a negative value indicated lumbar kyphosis. Based on 1 SD of mean lumbar lordosis angle (reference: mean ± SD 13.4 ± 12.4 degrees) reported using the Spinal Mouse in a previous study (12), lumbar kyphosis was defined in our study as decreased lumbar lordosis (lumbar lordosis angle ≤1.0 degrees). Assessment test for lumbar lordosis angle was performed thrice, and the mean value was used for the analysis.

The presence of current LBP was identified using the self-reported questionnaire with the question, "Have you had any back pain continuously for more than three months until the present?" We simultaneously obtained information on the

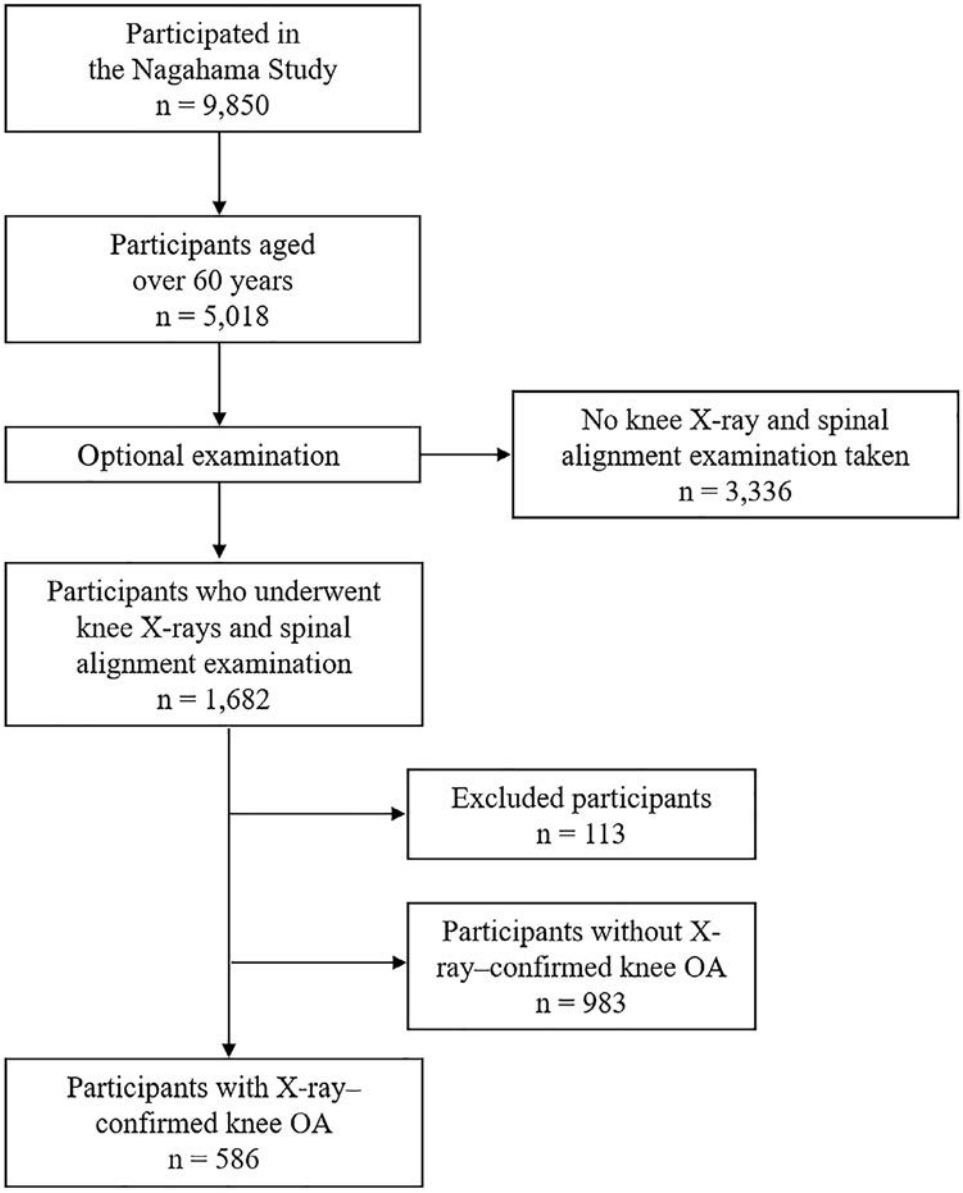

Figure 1. Flow chart for the selection of study participants. OA = osteoarthritis.

Table 1. Clinical characteristics of the participants, stratified by sex*

	Total (n = 586)	Women (n = 470)	Men (n = 116)	*P*	95% CI
KSS function (0–100 points)	84.8 ± 16.2	84.5 ± 16.5	86.1 ± 15.1	0.361	−4.84, 1.76
KSS symptom (0–25 points)	19.8 ± 5.6	19.9 ± 5.5	19.2 ± 6.1	0.235	0.45, 1.84
Gait speed, meters/second	1.6 ± 0.2	1.6 ± 0.2	1.7 ± 0.3	0.123	−0.09, 0.01
Lumbar lordosis, degrees	13.1 ± 13.1	13.6 ± 13.6	11.0 ± 10.9	0.049†	−5.34, −0.01†
LBP, no. (%)	195 (34.3)	156 (33.2)	39 (33.6)	0.93	0.66, 1.57
Age, years	68.8 ± 5.2	68.5 ± 5.1	70.2 ± 5.3	0.002†	−2.74, −0.63†
BMI, kg/m^2	23.0 ± 3.3	22.8 ± 3.3	23.9 ± 3.3	0.001†	−1.78, −0.46†
OA severity, no. (%)					
Either knee with K/L grade 2	171 (29.2)	127 (27.0)	44 (37.9)	0.197	0.69, 1.08
Both knees with K/L grade 2	290 (49.5)	243 (51.7)	47 (40.5)	–	–
Either knee with K/L grade 2 and the other ≥3	44 (7.5)	33 (7.0)	11 (9.5)	–	–
Both knees with K/L grade ≥3	81 (13.8)	67 (14.3)	14 (12.1)	–	–
Diabetes mellitus, no. (%)	62 (10.6)	43 (9.1)	19 (16.4)	0.023†	1.09, 3.49†
Osteoporosis, no. (%)	99 (17.4)	98 (20.9)	1 (0.8)	<0.001†	0.01, 0.24†
Depression, no. (%)	159 (28.0)	128 (27.2)	31 (26.7)	0.912	0.62, 1.54

* Values are the mean ± SD unless otherwise indicated. Sex differences in outcome measures and covariates were tested using unpaired *t*-test or chi-square test. 95% CI = 95% confidence interval; BMI = body mass index; K/L = Kellgren/Lawrence; KSS = Knee Society Knee Scoring System; LBP = low back pain; OA = osteoarthritis.
† Significant.

surgical history for any spinal disorders, which was used as part of the exclusion criteria.

Gait speed. Maximum gait speed was measured using a wireless phototube (Brower Timing Systems) as the objective measure of physical function. The phototube was set at 4 and 10 meters on a 12-meter gait path, and the time taken to walk past the phototube was measured in 0.01-second units. The participant was instructed to walk as fast as possible on the gait path. The maximum gait speed (meters/second) was calculated from the taken-time of the 6-meter distance.

Covariates. The participants' body height and weight were measured to the nearest 0.1 cm and 0.1 kg and converted to body mass index (BMI; kg/m^2). History of diabetes mellitus and osteoporosis was retrieved from the cohort data. Moreover, depressive symptomatology in the previous week was evaluated using the 20-item version of the Center for Epidemiological Studies Depression Scale (CES-D; range 0–60 points) (16,17). A higher score indicates a more depressive status, and the presence of depressive symptoms was defined as a CES-D score of ≥16 points. We used the Japanese version of the CES-D with its validity previously established in the Japanese population (18).

Statistical analysis. Continuous variables are expressed as means and SDs and categorical variables as frequencies (%). Outcome measures and covariates were compared between women and men using unpaired *t*-test or chi-square test. To investigate the associations of lumbar kyphosis with the KSS function score, KSS symptom score, and gait speed in the total sample, univariate analysis of variance (ANOVA) was performed. The adjusted mean difference between the groups was also estimated, with adjustments for age, BMI, sex, OA severity, and the presence of LBP, diabetes mellitus, osteoporosis, and depression. Similarly, a univariate ANOVA was conducted in the subgroups for sex, and the adjusted mean difference was calculated, adjusting for the above covariates except for sex. For the KSS scores and gait speed, 3-way ANOVAs were conducted to assess the interaction among sex, lumbar kyphosis, and LBP, with adjustments for age, BMI, OA severity, and the presence of diabetes mellitus, osteoporosis, and depression. Then, we classified the participants into 4 subgroups based on the combinations of LBP and lumbar kyphosis as follows: the absence of both LBP and lumbar kyphosis (reference category), LBP alone, lumbar kyphosis alone, and coexisting LBP and lumbar kyphosis. A multiple linear regression analysis was conducted using the KSS function score, KSS symptom score, and gait speed as the dependent variables and the 4 subgroup categories as the independent variables, with adjustments for the covariates including age, BMI, sex, OA severity, and the presence of LBP, diabetes mellitus, osteoporosis, and depression. As a secondary analysis, multiple linear regression analyses for each sex were conducted while adjusting for covariates except for sex. All statistical analyses were performed with SPSS software, version 25.0. The level of significance was set at a *P* value of less than 0.05.

RESULTS

Of a total of 9,850 individuals in this cohort, 5,018 were age >60 years. Among these, 1,682 participated in the optional physical assessment. After excluding participants who did not meet the inclusion criteria (participants without radiographic-confirmed

Table 2. Clinical characteristics of subgroups classified by the absence or presence of low back pain (LBP) and lumbar kyphosis (LK)*

	Women				Men			
	Absence of LBP and LK (n = 273)	LBP alone (n = 127)	LK alone (n = 41)	Both LBP and LK (n = 29)	Absence of LBP and LK (n = 66)	LBP alone (n = 34)	LK alone (n = 11)	Both LBP and LK (n = 5)
KSS function (0–100 points)	87.6 ± 15.1	81.8 ± 16.1	82.7 ± 17.2	69.7 ± 19.5	89.8 ± 10.6	82.6 ± 18.3	83.7 ± 14.6	65.4 ± 24.3
KSS symptom (0–25 points)	20.8 ± 5.1	18.7 ± 5.5	20.8 ± 5.0	15.2 ± 6.3	19.8 ± 5.5	18.6 ± 6.6	17.8 ± 6.5	19.6 ± 9.1
Gait speed, meters/second	1.7 ± 0.2	1.6 ± 0.2	1.6 ± 0.2	1.5 ± 0.2	1.7 ± 0.3	1.6 ± 0.3	1.7 ± 0.2	1.6 ± 0.4
Age, years	67.6 ± 4.9	68.5 ± 4.9	70.7 ± 5.2	73.7 ± 4.9	70.0 ± 5.5	70.2 ± 5.0	70.7 ± 5.8	71.2 ± 6.5
BMI, kg/m^2	22.9 ± 3.4	22.8 ± 3.1	22.5 ± 3.2	22.5 ± 2.8	23.6 ± 2.8	24.3 ± 3.7	24.1 ± 3.9	26.2 ± 4.4
OA severity, no. (%)								
Either knee with K/L grade 2	75 (27.5)	31 (24.4)	15 (36.6)	6 (20.7)	25 (37.9)	13 (38.2)	5 (45.5)	1 (20.0)
Both knees with K/L grade 2	141 (51.6)	71 (55.9)	17 (41.5)	14 (48.3)	27 (40.9)	13 (38.2)	5 (45.5)	2 (40.0)
Either knee with K/L grade 2 and the other ≥3	20 (7.3)	11 (8.7)	2 (4.9)	0 (0)	7 (10.6)	4 (11.8)	0 (0)	0 (0)
Both knees with K/L grade ≥3	37 (13.6)	14 (11.0)	7 (17.1)	9 (31.0)	7 (10.6)	4 (11.8)	1 (9.1)	2 (40.0)
Diabetes mellitus, no. (%)	25 (9.2)	13 (10.2)	3 (7.3)	2 (6.9)	14 (21.2)	4 (11.8)	1 (9.1)	0 (0)
Osteoporosis, no. (%)	37 (13.6)	39 (30.7)	11 (26.8)	11 (37.9)	1 (1.5)	0 (0)	0 (0)	0 (0)
Depression, no. (%)	68 (24.9)	40 (31.5)	9 (22.0)	11 (37.9)	17 (25.8)	8 (23.5)	4 (36.4)	2 (40.0)

* Values are the mean ± SD unless otherwise indicated. Outcome measures and physical characteristics in each subgroup based on the combinations of LBP and LK are represented. BMI = body mass index; K/L = Kellgren/Lawrence; KSS = Knee Society Knee Scoring System; OA = osteoarthritis.

knee OA, n = 983; participants who could not walk for more than 10 meters, n = 113), 586 participants with radiographic knee OA were included in the analysis (Figure 1). The clinical characteristics of those included are summarized in Table 1. The mean age was 68.8 ± 5.2 years (range 60–81 years), and women accounted for 80.1% of the included participants with knee OA. There were

Table 3. Differences in functional disabilities and symptoms between participants with and without lumbar kyphosis (LK)*

	Without LK, mean ± SD	With LK, mean ± SD	F	*P*	With LK vs. without LK, adjusted mean difference (95% CI)
Total sample (n = 500 without LK, n = 86 with LK)					
KSS function (0–100 points)	86.1 ± 15.3	77.4 ± 19.1	13.95	<0.001†	−6.14 (−9.36, −2.91)†
KSS symptom (0–25 points)	20.0 ± 5.5	18.5 ± 6.4	2.24	0.135	−0.90 (−2.08, 0.28)
Gait speed, meters/second	1.7 ± 0.2	1.6 ± 0.2	5.32	0.021†	−0.06 (−0.11, −0.01)†
Women (n = 400 without LK, n = 70 with LK)					
KSS function (0–100 points)	85.8 ± 15.7	77.3 ± 19.2	11.34	0.001†	−6.32 (−10.01, −2.63)†
KSS symptom (0–25 points)	20.2 ± 5.3	18.5 ± 6.2	2.69	0.102	−1.07 (−2.35, 0.21)
Gait speed, meters/second	1.7 ± 0.2	1.5 ± 0.2	8.74	0.003†	−0.09 (−0.14, −0.03)†
Men (n = 100 without LK, n = 16 with LK)					
KSS function (0–100 points)	87.4 ± 14.0	78.0 ± 19.4	4.10	0.045†	−6.92 (−13.70, −0.15)†
KSS symptom (0–25 points)	19.4 ± 5.9	18.4 ± 7.2	0.11	0.746	−0.51 (−3.59, 2.58)
Gait speed, meters/second	1.7 ± 0.3	1.7 ± 0.3	0.00	0.992	0.00 (−0.12, 0.12)

* The adjusted mean differences between participants with and without LK were estimated, with adjustments for age, body mass index, sex, osteoarthritis grade, and the presence of low back pain, diabetes mellitus, osteoporosis, and depression. 95% CI = 95% confidence interval; KSS = Knee Society Knee Scoring System.
† Significant.

Table 4. Associations by multiple linear regression analysis of the subcategories, stratified by the absence or presence of LBP and LK, with KSS scores and gait speed*

	Total				Women				Men			
	Absence of LBP and LK (n = 339)	LBP alone (n = 161)	LK alone (n = 52)	Coexisting LBP and LK (n = 34)	Absence of LBP and LK (n = 273)	LBP alone (n = 127)	LK alone (n = 41)	Coexisting LBP and LK (n = 29)	Absence of LBP and LK (n = 66)	LBP alone (n = 34)	LK alone (n = 11)	Coexisting LBP and LK (n = 5)
KSS function												
Beta (95% CI)	Ref.	−4.96 (−7.56, −2.36)†	−4.47 (−8.51, −0.43)†	−13.86 (−18.86, −8.86)†	Ref.	−4.47 (−7.46, −1.48)†	−4.63 (−9.28, 0.02)	−13.45 (−19.04, −7.86)†	Ref.	−6.09 (−11.38, −0.80)†	−4.93 (−13.05, 3.19)	−17.47 (−29.36, −5.59)†
Standard beta		−0.137†	−0.078†	−0.200†		−0.121†	−0.079	−0.196†		−0.184†	−0.096	−0.236†
P		<0.001†	0.03†	<0.001†		0.003†	0.051	<0.001†		0.024†	0.232	0.004†
KSS symptom												
Beta (95% CI)	Ref.	−1.72 (−2.67, −0.77)†	−0.41 (−1.89, 1.07)	−3.43 (−5.25, −1.60)†	Ref.	−1.87 (−2.90, −0.84)†	−0.08 (−1.69, 1.52)	−4.49 (−6.42, −2.55)†	Ref.	−1.02 (−3.41, 1.37)	−2.08 (−5.75, 1.59)	1.99 (−3.38, 7.37)
Standard beta		−0.136†	−0.021	−0.142†		−0.151†	−0.004	−0.196†		−0.077	−0.101	0.067
P		<0.001†	0.584	<0.001†		<0.001†	0.919	<0.001†		0.399	0.264	0.464
Gait speed												
Beta (95% CI)	Ref.	−0.03 (−0.07, 0.10)	−0.07 (−0.13, 0.00)†	−0.08 (−0.16, 0.00)†	Ref.	−0.02 (−0.06, 0.03)	−0.09 (−0.16, −0.02)†	−0.10 (−0.19, −0.01)†	Ref.	−0.08 (−0.18, 0.01)	−0.02 (−0.17, 0.12)	−0.04 (−0.26, 0.17)
Standard beta		−0.060	−0.080†	−0.084†		−0.031	−0.108†	−0.105†		−0.146	−0.021	−0.034
P		0.137	0.046†	0.039†		0.494	0.017†	0.024†		0.088	0.802	0.688

* A multiple linear regression analysis was conducted using the Knee Society Knee Scoring System (KSS) scores and gait speed as a dependent variable, and the 4 subgroup categories as independent variables, with adjustments for the covariates including age, body mass index, sex, osteoarthritis severity, and the presence of diabetes mellitus, osteoporosis, and depression. In the subanalysis by sex, multiple linear regression analyses were conducted with adjusting covariates except for sex. 95% CI = 95% confidence interval; LBP = low back pain; LK = lumbar kyphosis; Ref. = reference.
† Significant.

no significant differences between both sexes with respect to the KSS function and symptom score. However, the lumbar lordosis angle and comorbidities (diabetes mellitis and osteoporosis) were higher in women than in men, but women had a lower age and BMI. The clinical characteristics of the 4 subgroups classified by the absence or presence of LBP and/or lumbar kyphosis are shown in Table 2.

Associations between lumbar kyphosis and KSS scores or gait speed. In total, 86 participants (14.7%) had lumbar kyphosis (women: 70 [14.9%]; men: 16 [13.8%]). Although the KSS symptom score was not significantly different between the 2 groups, the KSS function scores in those with lumbar kyphosis were significantly lower than those without lumbar kyphosis, and the gait speed was significantly slower (Table 3). The adjusted mean difference in women with lumbar kyphosis was −6.32 points (95% confidence interval [95% CI] −10.01, −2.63) against those without lumbar kyphosis, and in men was −6.92 points (95% CI −13.70, −0.15).

Association of individual or coexisting LBP and lumbar kyphosis with KSS scores and gait speed. Results of the 3-way ANOVA showed a significant interaction for the KSS symptom score ($F = 5.315$, $P = 0.021$), but not for the KSS function score and gait speed ($F = 0.070$, $P = 0.792$ and $F = 0.114$, $P = 0.735$, respectively). Associations by multiple linear regression analysis of the subcategories, stratified by the absence or presence of LBP and lumbar kyphosis, with KSS scores and gait speed are shown in Table 4. Multiple linear regression analysis showed that the presence of LBP and lumbar kyphosis in participants with knee OA were independently associated with the KSS function score after adjusting for the covariates (beta = −4.96 [95% CI −7.56, −2.36], $P < 0.001$; beta = −4.47 [95% CI −8.51, −0.43], $P = 0.030$, respectively). LBP coexisting with lumbar kyphosis was significantly associated with decreased KSS function scores (beta = −13.86 [95% CI −18.86, −8.86], $P < 0.001$) in participants with knee OA.

In addition, LBP alone or LBP coexisting with lumbar kyphosis was significantly associated with the KSS symptom score (beta = −1.72 [95% CI −2.67, −0.77], $P < 0.001$; beta = −3.43 [95% CI −5.25, −1.60], $P < 0.001$, respectively). The presence of LBP in women was associated with a lower KSS symptom score, after adjustment for covariates (beta = −1.87 [95% CI −2.90, −0.84], $P < 0.001$). The coexistence of LBP and lumbar kyphosis also reinforced the decrease in the KSS symptom score (beta = −4.49 [95% CI −6.42, −2.55], $P < 0.001$) in women. In contrast, these relationships were not confirmed in men.

Moreover, we found that lumbar kyphosis alone, or its coexistence with LBP was significantly associated with a slow gait speed (beta = −0.07 [95% CI −0.13, 0.00], $P = 0.046$; beta = −0.08 [95% CI −0.16, 0.00]; $P = 0.039$, respectively).

DISCUSSION

The primary objective of this study was to determine the association of lumbar kyphosis and LBP with KSS function and symptoms in patients with knee OA. Results of the univariate ANOVA showed that lumbar kyphosis in participants with knee OA was associated with a reduction of KSS function scores and slow gait speed, but not with the KSS symptom score. Additionally, the association of the coexistence of LBP and lumbar kyphosis with functional disabilities was more remarkable than the individual association of these factors, which supported our hypothesis. A secondary objective was to determine the difference in these associations between the sexes. The coexistence of LBP and lumbar kyphosis was associated with a lower KSS symptom score in only women, and hence, this finding partially supports our secondary hypothesis. To the best of our knowledge, our study could be the first to determine the associations of both LBP and lumbar kyphosis with functional disabilities and knee symptoms in individuals with knee OA.

Concurring with the results of previous reports, which included community-dwelling older adults (7), our findings indicated that lumbar kyphosis was associated with a reduction of KSS function score and slow gait speed in patients with knee OA. In the knee-spine syndrome, a backward-tilting pelvis and knee-flexed posture produce dysfunction in the antigravity muscles (19). Miyazaki et al (7) have reported that decreased lumbar lordosis significantly correlated with poor knee extensor strength and slow gait speed. It is also well known that muscle weakness in patients with knee OA is one of the important factors causing functional disabilities (20). Therefore, lumbar kyphosis could induce muscle weakness and functional disabilities in patients with knee OA.

In contrast, lumbar kyphosis was not associated with a lower KSS symptom score in this study. We hypothesized that there is a significant relationship between lumbar kyphosis and knee symptoms because the postural changes caused by knee-spine syndrome could lead to increased knee shearing forces (21). One of the potential factors regarding this disparity could be the individual differences in compensatory patterns found in sagittal spinal malalignment. Since the compensatory strategy of the pelvis tilt in patients with knee OA varies according to the degree of femoral inclination, i.e., the knee flexion angle (11), this difference may have canceled the possible relationship between lumbar kyphosis and knee-related symptoms.

Our results showed no sex differences between lumbar kyphosis and functional disabilities, and the adjusted mean difference between participants with lumbar kyphosis and without lumbar kyphosis with respect to the KSS function score was nearly equal in women and men (women: −6.32 points; men: −6.92 points). Similar to the results of a previous report (8), we found a significant difference in the lumbar lordosis angles between women and men, but it was quite small. Therefore, although there

was a small difference between sex in the lumbar lordosis angle, this difference had minimal effects on functional disabilities.

The results of our study showed the individual association as well as the coexisting association of LBP and lumbar kyphosis in participants with knee OA. A recent systematic review (2) concluded that the presence of LBP in patients with knee OA was associated with functional disabilities and severe knee symptoms; however, its effect size was small. Our study also showed that the individual association of LBP was small, but coexisting LBP and lumbar kyphosis was associated with a remarkable reduction in KSS function scores of 13.86 points. Given that the minimum clinically important difference for the KSS on functional activities after total knee arthroplasty was 4.1 points (22), the association of the coexistence of LBP and lumbar kyphosis with functional activities had a sufficient clinical impact. Decreased lumbar lordosis was one of the major risk factors predictive of severe LBP (10); hence, participants with LBP coexisting with lumbar kyphosis in this study could have more severe LBP than participants with LBP alone. Additionally, since patients with spinal stenosis often exhibit a reduced lumbar lordosis (6), an accompanying muscle weakness in the lower extremity may also be associated with functional disabilities (23). These results suggest the necessity for a therapeutic approach focusing on LBP and lumbar kyphosis in patients with knee OA. Based on our findings, both symptom relief for LBP and back extensor strengthening (24) for preventing deterioration of lumbar kyphosis is needed for efficient management of functional disabilities in patients with knee OA with respect to the clinical scenario.

Our results confirmed the association of lower KSS symptom scores with LBP alone, as well as with LBP coexisting with lumbar kyphosis in only women. Several previous studies (3,25,26) have reported the association between severe knee symptoms and LBP in patients with knee OA. One study (27) suggested that women with LBP have a higher risk of developing knee pain than men with LBP, implying a sex-related difference in the association of LBP and knee pain aggravation. In addition, women with knee-spine syndrome demonstrated a larger knee flexion angle than did men with the same degree of lumbar kyphosis (5). Therefore, increased mechanical stress caused by an increased knee flexion angle, together with LBP, possibly affected the knee symptoms in women in our study. Although the pathologic mechanism of the observed sex difference is unclear, the association between severe knee symptoms and the coexistence of LBP and lumbar kyphosis needs attention from a clinical point of view.

This study had several limitations. First, although the sample size for data analysis in the total sample who had knee OA was large, the number of men was relatively small when classified into subgroups. Thus, we could not reach a definite conclusion about the association of the coexistence of LBP and lumbar kyphosis with functional disability and knee symptoms in men. Second, we did not perform radiographic examinations for spinal stenosis. In participants who had both knee OA and spinal stenosis, it should be noted that the effects of these diseases demonstrated mixed functional disabilities since spinal stenosis is also associated with physical dysfunction (23). Moreover, there was a lack of information regarding the severity of LBP. Although the severity of LBP was higher in women than in age-matched men (28), the present study did not consider this factor. Finally, the cross-sectional design of this study does not allow us to determine any causal relationship between LBP and/or lumbar kyphosis and dysfunction in patients with knee OA. Future longitudinal studies are warranted to determine how LBP and/or lumbar kyphosis could influence functional disability and symptoms in patients with knee OA.

In conclusion, the presence of LBP and lumbar kyphosis are associated with functional disabilities in patients with knee OA, and the association is more remarkable in patients who had both LBP and lumbar kyphosis. In addition, the coexistence of LBP and lumbar kyphosis had adverse associations on knee symptoms in women. These findings suggest that both LBP and lumbar kyphosis are useful clinical signs indicating functional disability and knee symptoms in patients with knee OA.

ACKNOWLEDGMENTS

We are extremely grateful to the Nagahama City Office and nonprofit organization Zeroji Club for their help in conducting the Nagahama 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. Taniguchi 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. Taniguchi, Ikezoe, Masaki, Kamitani, Tsuboyama, Ito, S. Matsuda, Tabara, F. Matsuda, Ichihashi.

Acquisition of data. Taniguchi, Ikezoe, Masaki, Kamitani, Tsuboyama, Ito, S. Matsuda, Tabara, F. Matsuda, Ichihashi.

Analysis and interpretation of data. Taniguchi, Ikezoe, Kamitani, Tsuboyama, Ichihashi.

REFERENCES

1. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. Ann Rheum Dis 2001;60:91–7.
2. Calders P, van Ginckel A. Presence of comorbidities and prognosis of clinical symptoms in knee and/or hip osteoarthritis: a systematic review and meta-analysis. Semin Arthritis Rheum 2018;47:805–13.
3. Suri P, Morgenroth DC, Kwoh CK, Bean JF, Kalichman L, Hunter DJ. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2010;62:1715–23.

4. Iijima H, Suzuki Y, Aoyama T, Takahashi M. Interaction between low back pain and knee pain contributes to disability level in individuals with knee osteoarthritis: a cross-sectional study. Osteoarthritis Cartilage 2018;26:1319–25.
5. Murata Y, Takahashi K, Yamagata M, Hanaoka E, Moriya H. The knee-spine syndrome: association between lumbar lordosis and extension of the knee. J Bone Joint Surg Br 2003;85:95–9.
6. Abbas J, Hamoud K, May H, Hay O, Medlej B, Masharawi Y, et al. Degenerative lumbar spinal stenosis and lumbar spine configuration. Eur Spine J 2010;19:1865–73.
7. Miyazaki J, Murata S, Horie J, Uematsu A, Hortobagyi T, Suzuki S. Lumbar lordosis angle (LLA) and leg strength predict walking ability in elderly males. Arch Gerontol Geriatr 2013;56:141–7.
8. Asai Y, Tsutsui S, Oka H, Yoshimura N, Hashizume H, Yamada H, et al. Sagittal spino-pelvic alignment in adults: the Wakayama Spine Study. PLoS One 2017;12:e0178697.
9. Makris UE, Fraenkel L, Han L, Leo-Summers L, Gill TM. Epidemiology of restricting back pain in community-living older persons. J Am Geriatr Soc 2011;59:610–4.
10. Chun SW, Lim CY, Kim K, Hwang J, Chung SG. The relationships between low back pain and lumbar lordosis: a systematic review and meta-analysis. Spine J 2017;17:1180–91.
11. Wang WJ, Liu F, Zhu YW, Sun MH, Qiu Y, Weng WJ. Sagittal alignment of the spine-pelvis-lower extremity axis in patients with severe knee osteoarthritis: a radiographic study. Bone Joint Res 2016;5:198–205.
12. Tabara Y, Masaki M, Ikezoe T, Setoh K, Kato T, Kawaguchi T, et al. Small degree of lumbar lordosis as an overlooked determinant for orthostatic increases in blood pressure in the elderly: the Nagahama study. Am J Hypertens 2019;32:61–9.
13. Lee S, Kim TN, Kim SH. Sarcopenic obesity is more closely associated with knee osteoarthritis than is nonsarcopenic obesity: a cross-sectional study. Arthritis Rheum 2012;64:3947–54.
14. Taniguchi N, Matsuda S, Kawaguchi T, Tabara Y, Ikezoe T, Tsuboyama T, et al. The KSS 2011 reflects symptoms, physical activities, and radiographic grades in a Japanese population. Clin Orthop Relat Res 2015;473:70–5.
15. Imagama S, Ito Z, Wakao N, Seki T, Hirano K, Muramoto A, et al. Influence of spinal sagittal alignment, body balance, muscle strength, and physical ability on falling of middle-aged and elderly males. Eur Spine J 2013;22:1346–53.
16. Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D (Center for Epidemiological Studies Depression) depression symptoms index. J Aging Health 1993;5:179–93.
17. Rushton JL, Forcier M, Schectman RM. Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. J Am Acad Child Adolesc Psychiatry 2002;41:199–205.
18. Shima S, Shikano T, Kitamura T, Asai M. New self-rating scales for depression. Seishin-Igaku 1985;27:717–23.
19. Barrey C, Roussouly P, Perrin G, Le Huec JC. Sagittal balance disorders in severe degenerative spine: can we identify the compensatory mechanisms? Eur Spine J 2011;20 Suppl 5:626–33.
20. Ruhdorfer A, Wirth W, Eckstein F. Relationship between isometric thigh muscle strength and minimum clinically important differences in knee function in osteoarthritis: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2015;67:509–18.
21. Harato K, Nagura T, Matsumoto H, Otani T, Toyama Y, Suda Y. Knee flexion contracture will lead to mechanical overload in both limbs: a simulation study using gait analysis. Knee 2008;15:467–72.
22. Nishitani K, Yamamoto Y, Furu M, Kuriyama S, Nakamura S, Ito H, et al. The minimum clinically important difference for the Japanese version of the new Knee Society Score (2011KSS) after total knee arthroplasty. J Orthop Sci 2019;24:1053–7.
23. Kasukawa Y, Miyakoshi N, Hongo M, Ishikawa Y, Kudo D, Kijima H, et al. Lumbar spinal stenosis associated with progression of locomotive syndrome and lower extremity muscle weakness. Clin Interv Aging 2019;14:1399–405.
24. Kasukawa Y, Miyakoshi N, Hongo M, Ishikawa Y, Kudo D, Suzuki M, et al. Age-related changes in muscle strength and spinal kyphosis angles in an elderly Japanese population. Clin Interv Aging 2017;12:413–20.
25. Reeuwijk KG, de Rooij M, van Dijk GM, Veenhof C, Steultjens MP, Dekker J. Osteoarthritis of the hip or knee: which coexisting disorders are disabling? Clin Rheumatol 2010;29:739–47.
26. Zullig LL, Bosworth HB, Jeffreys AS, Corsino L, Coffman CJ, Oddone EZ, et al. The association of comorbid conditions with patient-reported outcomes in Veterans with hip and knee osteoarthritis. Clin Rheumatol 2015;34:1435–41.
27. Ito H, Tominari S, Tabara Y, Nakayama T, Furu M, Kawata T, et al. Low back pain precedes the development of new knee pain in the elderly population; a novel predictive score from a longitudinal cohort study. Arthritis Res Ther 2019;21:98.
28. Cecchi F, Debolini P, Lova RM, Macchi C, Bandinelli S, Bartali B, et al. Epidemiology of back pain in a representative cohort of Italian persons 65 years of age and older: the InCHIANTI study. Spine (Phila Pa 1976) 2006;31:1149–55.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1676–1688
DOI 10.1002/acr.24603

Effectiveness of Nonsurgical Interventions for Hallux Valgus: A Systematic Review and Meta-Analysis

Sheree E. Hurn,[1] Barry G. Matthews,[2] Shannon E. Munteanu,[3] and Hylton B. Menz[3]

Objective. To conduct a systematic review and meta-analysis investigating the effectiveness of nonsurgical interventions for hallux valgus (HV).

Methods. Medline, CINAHL, Embase, and the Cochrane Library were searched to April 2020, including parallel-group and crossover studies investigating nonsurgical interventions for HV. Two reviewers independently screened articles for inclusion, extracted data, determined risk of bias, and made assessments using the Grading of Recommendations, Assessment, Development, and Evaluation methodology. Risk of bias was assessed using version 2 of the Cochrane risk-of-bias tool. Effect sizes (mean differences or risk ratios, and 95% confidence intervals) were calculated and pooled where possible for the primary outcomes, foot pain, and HV angle.

Results. Eighteen included studies investigated a wide range of nonsurgical interventions for HV. Most studies had small sample sizes and concerns regarding risk of bias. Five separate meta-analyses for foot orthoses, splints, manual therapy, and taping added to foot exercises showed no significant effects on primary outcomes. However, results from 8 studies showed a significant pain reduction with the use of foot orthoses, night splints, dynamic splints, manual therapy, taping added to foot exercises, a multifaceted physical therapy program, and Botox injections. Four studies reported a clinically significant reduction in HV angle with night splints, foot exercises, multifaceted physical therapy, and Botox injections.

Conclusion. There is a low level of certainty surrounding the effectiveness of nonsurgical interventions for HV, but a reduction in pain appears more likely than improvement in HV angle.

INTRODUCTION

Hallux valgus (HV) is a progressive and disabling foot deformity in which the hallux deviates laterally toward the lesser toes, disrupting the alignment of the first metatarsophalangeal (MTP) joint. HV is prevalent and estimated to affect 23% of adults ages 18–65 years and 36% of adults age >65 years (1). HV is associated with both degenerative and inflammatory arthritic conditions. In older adults, a higher frequency of osteoarthritic change in the first MTP joint is associated with increasing HV severity (2). Furthermore, HV is a common finding in rheumatoid arthritis (RA), with research linking HV to painful plantar callosities in the forefoot in RA (3).

As the first MTP joint becomes progressively subluxed, foot function is disturbed, leading to postural instability (4,5) and an increased risk of falls in older adults (6,7). People with HV compared to controls experience more disabling foot pain, difficulty with footwear, and concerns about appearance (8), which may lead them to seek corrective orthopedic surgery. HV correction procedures are among the top 10 most common foot and ankle procedures performed by foot and ankle surgeons in the US, causing an estimated economic burden of 325.1 million US dollars in 2011 (9).

Conservative treatment is recommended prior to surgical intervention, and a range of nonsurgical interventions is available, including change in footwear, foot orthoses, various types of toe splints, joint mobilization, taping, and stretching or strengthening exercises (10). However, the effectiveness of nonsurgical interventions for HV is uncertain. A systematic review published in 2014 (11) identified 2 randomized trials investigating foot orthoses for HV, but other conservative interventions had not been evaluated. Therefore, the aim of this study was to conduct an updated

Supported by the La Trobe University Research Focus Area in Sport, Exercise, and Rehabilitation.

[1]Sheree E. Hurn, PhD: Queensland University of Technology, Brisbane, Queensland, Australia; [2]Barry G. Matthews, DipAppSc(Pod): Queensland University of Technology, Brisbane, Queensland, and La Trobe University, Melbourne, Victoria, Australia; [3]Shannon E. Munteanu, PhD, Hylton B. Menz, DSc: La Trobe University, Melbourne, Victoria, Australia.

Dr. Hurn has received consulting fees from FRANKiE4 footwear (less than $10,000). No other disclosures relevant to this article were reported.

Address correspondence to Hylton B. Menz, DSc, La Trobe University, School of Allied Health, Human Services and Sport, Bundoora, Victoria 3086, Australia. Email: h.menz@latrobe.edu.au.

Submitted for publication October 8, 2020; accepted in revised form March 23, 2021.

SIGNIFICANCE & INNOVATIONS

- This systematic review and meta-analysis provides an up-to-date synthesis of clinical trials investigating nonsurgical interventions for hallux valgus (HV).
- A wide range of nonsurgical interventions has been tested in trials for HV, with very low to moderate certainty regarding effectiveness.
- Results indicate that nonsurgical interventions are unlikely to improve HV angle.
- Nonsurgical interventions may improve pain outcomes in the short to intermediate term.

systematic review of the literature surrounding nonsurgical interventions for HV.

MATERIALS AND METHODS

Review registration and search strategy. This review was registered with PROSPERO (CRD42019111711) and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (12). A systematic search strategy was developed using Medline, CINAHL, Embase, and the Cochrane Library. Two strings of search terms were developed, including medical subject heading (MeSH) terms, key words, and synonyms: 1) "hallux valgus" or "bunion" and 2) "intervention" or "treatment" or "therapy" or "management" or "effects." Truncation, proximity, and Boolean operators were used as appropriate, and limiters applied for human studies and English language (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24603). The first search was conducted without date restriction up to January 2019. An updated search was conducted in April 2020, with date restrictions from January 2018 to April 2020. In addition to the electronic database search, reference lists of included studies and systematic literature reviews were handsearched, and 1 expert in the field independent to the research team was consulted. Duplicates were removed using Endnote X8 (2019 search) and Covidence (2020 search).

Study inclusion. Predetermined inclusion criteria were based on the PICO framework (Population, Intervention, Comparison, Outcome), where: P = population of human participants diagnosed with HV (clinical or radiographic diagnosis); I = any nonsurgical intervention aiming to improve outcomes in HV; C = any other comparative treatment or no treatment (e.g., waitlist or placebo); O = any measure of treatment effects, including self-reported pain and function, or quantitative measures involving the foot. The following study designs were considered for inclusion, based on the Australian National Health and Medical Research Council Hierarchy of Evidence levels II to III: randomized controlled trials, pseudo-randomized controlled trials, and comparative studies with and without concurrent controls (including crossover studies) (13). Crossover trials were considered appropriate since HV is a relatively stable, chronic condition, and nonsurgical interventions are expected to have temporary effects. Exclusion criteria included populations with recent foot surgery or major trauma. Two reviewers (SEH and BGM) screened all titles and abstracts for potential relevance, warranting full-text retrieval. The same 2 reviewers performed full-text review using Covidence, and reasons for exclusion were noted. Figure 1 outlines a flow chart of study selection.

Data extraction. Data extraction was performed by 2 reviewers (SEH and BGM) independently, using a review template created in Covidence. If >1 published manuscript described the same study, these were treated as a single study for the purpose of data extraction. Details about experimental and comparison interventions or placebo were extracted using the Template for Intervention Description and Replication framework (14). For the purpose of combining similar interventions, the following broad definitions were used: foot orthoses (shoe inserts designed to provide arch support), night splints (any corrective device for hallux alignment worn at rest), dynamic splints (any corrective device for hallux alignment worn during activities of daily living), manual therapy (any type of foot mobilization or manipulation), and foot exercises (strengthening exercises targeting the intrinsic foot muscles). Primary outcome measures for this review were HV angle and self-reported pain. If >1 measure of pain was used, visual analog scales (VAS) or numeric rating scales (NRS) were chosen over other pain questionnaires, resting pain was chosen before walking pain, and average pain was chosen over maximum pain. Means ± SDs and sample size for experimental and control groups were extracted for all outcome measures at baseline and the following periods of follow-up: short-term (≤3 months); intermediate (>3 months and <12 months); and long-term (≥12 months). If 2 follow-up assessments were completed within 1 of the defined time points, the longer follow-up period was selected.

Risk-of-bias assessment. Assessment of risk of bias was performed independently by 2 reviewers (SEH and BGM) using version 2 of the Cochrane risk-of-bias tool for randomized trials (RoB 2) (15). The RoB 2 assesses the risk of bias across 5 domains, and calls for an assessor judgment on the level of bias within each domain and for the outcome estimate overall (low risk = low risk of bias for all domains; some concerns = some concerns for at least 1 domain, and not high risk for any domain; high risk = high risk of bias in at least 1 domain). For crossover study designs, signaling questions from the archived version of RoB 2 for crossover trials were used to inform judgments for the second (deviations from intended interventions) and third (missing outcome data) domains. Reviewers used the Microsoft Excel tool to implement RoB 2 (available at www.riskofbias.info) (15).

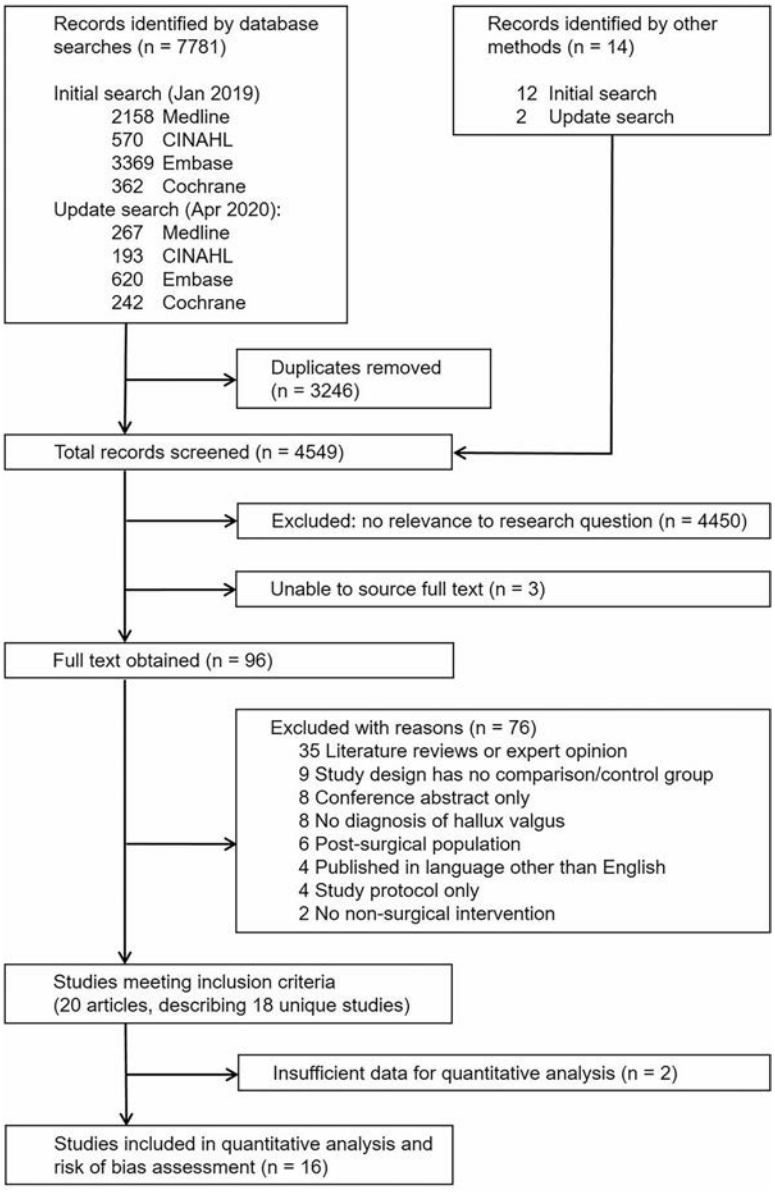

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart of study selection.

If consensus could not be reached, other authors (SEM and HBM) were consulted to achieve consensus.

Data synthesis and analysis. In studies where bilateral measures were taken (e.g., HV angle) and extracted means ± SDs represented feet rather than participants, the bilateral measures were treated as clustered data and the following additional information was collated: the number of clusters (participants), the average size of each cluster (e.g., 2 feet), and a conservative estimate of the intracluster correlation coefficient was predetermined to be 0.05. The following formula was then applied to determine the effective sample size: $1 + (M–1) \times ICC$, where M is the average cluster size and ICC is the intracluster correlation coefficient (16). For crossover study designs, within-person differences were extracted if possible (mean ± SD) for paired analysis (17).

All extracted data were imported from Covidence into Review Manager 5.3 for analysis (18). For continuous outcomes, mean differences and 95% confidence intervals (95% CIs) were calculated using an inverse variance method and random-effects model. We expected that most studies would report pain on a 100-point scale with a direction of lower = better. Where pain outcomes were reported on a 10-point scale, data were multiplied by 10 to convert to a 100-point scale. Where the direction of a pain scale was higher = better, data were converted to negative values so that the direction of mean differences would align across studies. Categorical variables were converted into dichotomous data based on a consensus approach, and risk ratios and 95% CIs

Table 1. Selected characteristics of included studies*

Author, year (ref.)	Study design	Follow-up	Sample size (no.)	Age, years	Sex (M:F)	Experimental group	Comparison groups	Primary outcomes	Secondary outcomes
Orthoses									
Kilmartin et al, 1994 (35)	RCT†	3 years	122 (139 feet)	9–10	16:106	Foot orthoses	Control (no treatment)	HV angle‡	IM angle
Reina et al, 2013 (39)	RCT§	12 months	54 (83 feet)	30.6 ± 11.9	0:54	Foot orthoses	Control (no treatment)	HV angle‡	IM angle
Torkki et al, 2001/2003 (41,42)	RCT	2 years	209	48 ± 9.7	16:193	Foot orthoses	Control (waiting list), surgery	Pain VAS	Satisfaction, ability to work, AOFAS, cosmetic disturbance, footwear problems, QoL, global foot assessment
Splints									
Chadchavalpanichaya and Chueluecha, 2011 (31)	RCT	12 months	47	44.1 ± 15.3	3:44	Night splint	Control (general foot care, shoes)	HV angle‡	Satisfaction NRS
Chadchavalpanichaya et al, 2018 (32)	RCT	12 months	90	60.6 ± 6.6	5:85	Night splint (toe separator)	Control (general foot care, shoes)	HV angle, pain NRS‡	Satisfaction NRS
Mirzshahi et al, 2012 (26)¶	RCT	12 months	30	8–60	NR	Dynamic splint (slipper)	Night splint	HV angle‡	IM angle, comfort
Moulodi et al, 2019 (37)	Cross-over	1 month	24	22.8 ± 1.4	12:12	Dynamic splint	Night splint	HV angle, pain (FAOS)#	First MTP joint range, FAOS subscales (symptoms, ADL, sport, QoL)
Plaass et al, 2020 (38)	RCT	2–10 months	70	50.9 ± 13.6	4:66	Night splint	Control (no treatment)	HV angle, pain (FAOS)‡	IM angle, pain (rest, start-up, walking, running), ADL reduction, AOFAS, first MTP joint range, FAOS subscales (symptoms, ADL, sport, QoL), SF-36
Tehrraninasr et al, 2008 (40)	RCT†	3 months	30 (60 feet)	27 ± 8.9	0:30	Dynamic splint (insole)	Night splint	HV angle, pain VAS‡	IM angle
Physical therapies									
Abdalbary, 2018 (27)	RCT	12 months	56	45.6 ± 6.5	0:56	Manual therapy, exercises, toe separator, footwear advice	Control (footwear advice)	HV angle, pain VAS‡	AOFAS scale, IM angle, ankle dorsiflexion, hallux plantarflexion, abduction, and toe grip strength
Bayar et al, 2011 (28)	RCT†	8 weeks	20 (40 feet)	NR	0:20	Taping and foot exercises	Foot exercises	HV angle, pain VAS#	Walking pain, walking ability scale
Brantingham et al, 2005 (29)	RCT	3 weeks	60	50.2	0:60	Manual therapy	Sham APT (placebo)	Pain NRS	Satisfaction, FFI, AOFAS, pressure pain threshold
Broodryk, 2000 (30)	RCT	3 weeks	60	40.8 (15–65)	10:50	Manual therapy	Sham laser (placebo)	Pain NRS	Pressure-pain threshold, FFI

(Continued)

Table 1. (*Cont'd*)

Author, year (ref.)	Study design	Follow-up	Sample size (no.)	Age, years	Sex (M:F)	Experimental group	Comparison groups	Primary outcomes	Secondary outcomes
Choi, 2017 (33)	RCT	6 weeks	24	20–29	0:24	Foot exercises	Taping; exercises and taping	HV angle‡	Postural sway and limits of stability test
Du Plessis et al, 2011 (34)	RCT	6 weeks	30	42 (25–65)	15:15	Manual therapy	Night splint	Pain VAS	FFI, first MTP joint range
Kim et al, 2015 (36)	RCT	8 weeks	24	22.5 ± 2.4	13:11	Foot exercises and dynamic splint	Dynamic splint	HV angle‡	Abductor hallucis CSA, active abduction HV angle
Other									
Khan, 1996 (25)¶	RCT§	8 weeks	60 (80 feet)	NR	NR	*Tagetes patula* paste (marigold therapy) with felt pad	Placebo paste with felt pad	HV angle, pain VAS**	Width, satisfaction
Wu et al, 2015 (43)	RCT†	2 months	16 (26 feet)	42.9 ± 12	1:15	Intramuscular botulinum toxin type A injection	Placebo intramuscular saline injection	HV angle, pain (FFI)‡	FFI disability subscale

* ADL = activities of daily living; AOFAS = hallux-metatarsophalangeal scale of the American Orthopaedic Foot and Ankle Society; APT = action potential therapy; CSA = cross-sectional area; F = female; FAOS = Foot and Ankle Outcomes Scale; FFI = Foot Function Index; HV = hallux valgus; IM = intermetatarsal; M = male; MTP = metatarsophalangeal; NR = not reported; NRS = numeric rating scale; QoL = quality of life; RCT = randomized controlled trial; SF-36 = Short Form 36 health survey; VAS = visual analog scale.
† Cluster randomized controlled trial (unit of analysis feet rather than participants).
‡ Measured on radiograph.
§ Cluster randomized controlled trial (unit of analysis feet rather than participants). Not all participants were randomly allocated to groups.
¶ Insufficient data for quantitative analysis.
# Measured using goniometer.
** Measured on radiograph or photograph.

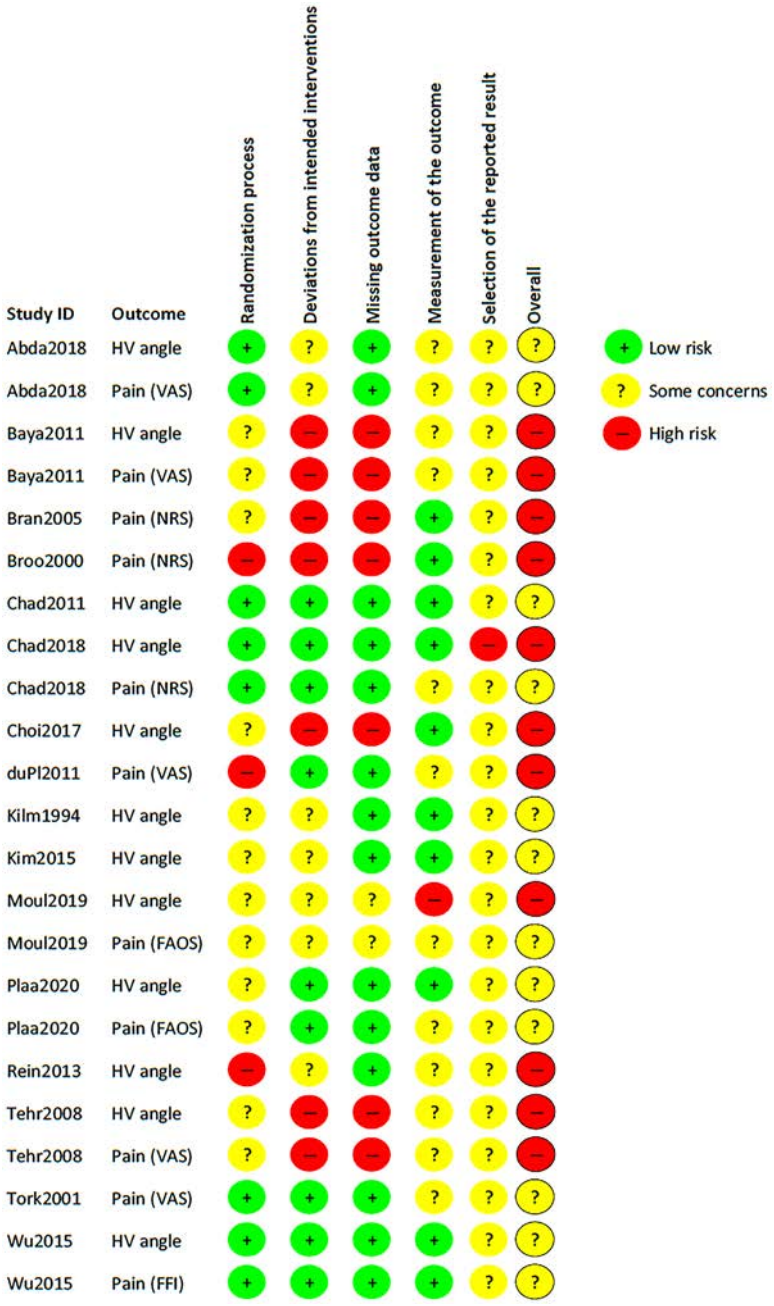

Figure 2. Summary of risk-of-bias assessment (using version 2 of the Cochrane risk-of-bias tool for randomized trials). FAOS = Foot and Ankle Outcomes Scale; FFI = Foot Function Index; HV = hallux valgus; NRS = numeric rating scale; VAS = visual analog scale.

were calculated using an inverse variance method and random-effects model. Meta-analysis was performed where possible, whenever studies had a similar type of intervention and comparator, and had reported 1 of the same primary outcomes (HV angle or pain). Subgroup analysis by study population (adult versus juvenile HV) was considered if possible, based on the included studies.

Certainty of the evidence was assessed using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach (19). Two authors (SEH and BGM) independently conducted the GRADE assessments using the online GRADEpro tool (available at www.gradepro.org) (19), and other authors (SEM and HBM) were contacted if required to achieve consensus. GRADE considerations included risk of bias, inconsistency, imprecision, indirectness, and publication bias. When considering imprecision of effect size estimates and CIs, a clinically important effect for HV angle was considered to be a mean difference of ≥2.5 degrees based on published minimum detectable change values (20), and a clinically important effect for pain was considered to be a mean difference of ≥10 points

Table 2. Summary of findings with GRADE certainty of the evidence*

Outcomes	Mean difference (95% confidence interval) between groups at follow-up	Total participants/ studies (refs.)	GRADE certainty of the evidence
Orthoses compared to control			
HV angle assessed with radiograph (degrees); follow-up long-term	No statistically significant difference; the mean HV angle in the orthoses group was on average 1.87 degrees higher (−0.32, 4.06) compared to control at follow-up	214/2 (35,39)	⊕⊕⊖⊖ low†
Pain assessed with VAS (0–100 scale); follow-up intermediate- and long-term	At intermediate follow-up, mean pain score was 9 points lower (−16.8, −1.2) on a 100-point pain scale in the orthoses group, but there was no difference between groups at long-term follow-up	138/1 (41)	⊕⊕⊕⊖ moderate‡
Orthoses compared to surgery			
Pain assessed with VAS (0–100 scale); follow-up intermediate- and long-term	At intermediate follow-up, mean pain score was 10 points higher (2.2, 17.8) in the orthoses group compared to surgery; at long-term follow-up, mean pain score was 17 points higher (9.4, 24.6) in the orthoses group compared to surgery	140/1 (41)	⊕⊕⊕⊖ moderate‡
Night splints compared to control			
HV angle assessed with radiographs (degrees); follow-up short- and long-term	No statistically significant difference; the mean HV angle in the night splints group was on average 0.3 degrees (−3.4, 2.8) lower	207/3 (31,32,38)	⊕⊕⊖⊖ low§
Pain assessed with NRS or FAOS (0–100 scale); follow-up short- and long-term	No statistically significant difference; the mean pain score in the night splints group was on average 7.2 points lower (−16.3, 1.9) compared to control	160/2 (32,38)	⊕⊕⊖⊖ low†
Dynamic splints compared to night splints			
HV angle assessed with radiograph or goniometer (degrees); follow-up short-term	No statistically significant difference; the mean HV angle was 0.3 to 1.2 degrees higher in the dynamic splints group compared to night splints	106/2 (37,40)	⊕⊕⊖⊖ low¶
Pain assessed with VAS or FAOS (0–100 scale); follow-up short-term	Inconsistent study results; 1 study reported pain scores 13.4 points lower (−22.3, −4.5), and another study reported pain scores 1.6 points higher in the dynamic splints group (−6.0, 9.2)	78/2 (37,40)	⊕⊖⊖⊖ very low#
Manual therapy compared to placebo			
Pain assessed with NRS (0–100 scale); follow-up short-term	No statistically significant difference; the mean pain score in the manual therapy group was on average 13.5 points lower (−51.7, 24.8) compared to placebo	120/2 (29,30)	⊕⊖⊖⊖ very low**
Manual therapy compared to night splints			
Pain assessed with VAS (0–100 scale); follow-up short-term	The mean pain score in the manual therapy group was 18.3 points lower (−26.0, −10.8) compared to night splints	30/1 (34)	⊕⊖⊖⊖ very low††
Foot exercises and dynamic splints compared to dynamic splints			
HV angle assessed with radiographs; follow-up short-term	The mean HV angle was 3.8 degrees lower (−6.6, −1.0) in the exercise group compared to control	24/1 (36)	⊕⊖⊖⊖ very low††
Taping and foot exercises compared to foot exercises alone			
HV angle assessed with radiograph or goniometer; follow-up short-term	No statistically significant difference; the mean HV angle was 0.6 degrees lower (−2.5, 1.2) in the taping and exercise group compared to exercise alone	54/2 (28,33)	⊕⊖⊖⊖ very low‡‡
Pain assessed with VAS (0–100 scale); follow-up short-term	The mean pain score was 18.8 points lower (−30.6, 7.0) in the taping and exercise group compared to exercise alone.	20/1 (28)	⊕⊖⊖⊖ very low§§
Multifaceted physical therapy compared to control			
HV angle assessed with radiographs; follow-up short- and long-term	The mean HV angle was 8.1 degrees lower (−9.8, 6.5) compared to control at short-term follow-up and 7.1 degrees lower (−8.5, −5.7) compared to control at long-term follow-up	56/1 (27)	⊕⊕⊖⊖ low¶¶
Pain assessed with VAS (0–100 scale); follow-up short- and long-term	The mean pain score was 30 points lower (−36.1, 23.9) compared to control at short-term follow-up and 35 points lower (−41.1, −28.9) compared to control at long-term follow-up	56/1 (27)	⊕⊕⊖⊖ low¶¶
Botox injections compared to saline injections			
HV angle assessed with radiographs; follow-up short-term	The mean HV angle was 4.4 degrees lower (−5.8, 3.0) in the Botox group compared to placebo (saline)	26/1 (43)	⊕⊕⊕⊖ moderate‡

(Continued)

Table 2. (*Cont'd*)

Outcomes	Mean difference (95% confidence interval) between groups at follow-up	Total participants/ studies (refs.)	GRADE certainty of the evidence
Pain assessed with FFI pain subscale (0–100 scale); follow-up short-term	The mean pain score was 12.6 points lower (−16.1, 9.1) in the Botox group compared to placebo (saline)	26/1 (43)	⊕⊕⊕⊖ moderate‡

* FAOS = Foot and Ankle Outcomes Scale; FFI = Foot Function Index; GRADE = Grading of Recommendations, Assessment, Development, and Evaluation; HV = hallux valgus; NRS = numeric rating scale; VAS = visual analog scale.
† Downgraded 1 level due to imprecision and risk of bias.
‡ Downgraded due to imprecision.
§ Downgraded due to imprecision and inconsistency.
¶ Downgraded 2 levels due to risk of bias.
# Downgraded due to imprecision and inconsistency, and downgraded 2 levels due to risk of bias.
** Downgraded 1 level due to risk of bias, imprecision, inconsistency, and publication bias.
†† Downgraded 1 level due to risk of bias, imprecision, and publication bias.
‡‡ Downgraded 2 levels due to risk of bias and publication bias.
§§ Downgraded 2 levels due to risk of bias and due to publication bias and imprecision.
¶¶ Downgraded 1 level due to risk of bias and publication bias.

(21–23). When considering inconsistency of effects, forest plots were visually inspected and heterogeneity was assessed using the I^2 statistic. Considerable heterogeneity was considered to be indicated by an I^2 value >75% (24). Publication bias was assessed via visual inspection of funnel plots, with SE plotted against effect sizes for primary outcomes. For each intervention category, conference abstracts and trial registrations were reviewed and compared to published trials.

RESULTS

Search. The database search yielded a total of 7,781 records, and after removal of 3,246 duplicates, 4,535 records remained. Reference list searches revealed another 14 potentially relevant records, and expert consultation also noted 1 of these studies. Thus, a total of 4,549 records were screened. After title and abstract screening, 4,450 studies were determined to be ineligible, and 3 articles could not be sourced via library services or direct contact with authors. Full text was reviewed against inclusion criteria for 96 studies, and exclusion reasons noted (Figure 1). Twenty articles (18 unique studies) met eligibility criteria. Two studies did not report sufficient data for quantitative analysis, and the required data could not be sourced after contacting the authors (25,26). Therefore, 16 unique studies were included in the meta-analysis and risk-of-bias assessment (27–43).

Included studies. Table 1 summarizes selected characteristics of 18 included studies. Studies were conducted in 11 countries, including Egypt, Finland, Germany, Iran, the Republic of Korea, South Africa, Spain, Taiwan, Thailand, Turkey, and the UK. Seventeen studies used parallel-group experimental designs, and 1 crossover study was included (37). A total of 1,026 participants were included (95 men and 841 women, based on 16 studies reporting sex). Sample sizes ranged from 16 to 209 participants, with only 4 studies having a sample size larger than 60 participants (32,35,38,41). Seven studies recruited patients with HV from outpatient clinics (26,27,31,32,34,38,41), 6 studies used convenience sampling (29,30,33,35,37,40), 1 study reported a combination (39), and 4 studies did not report their recruitment methods (25,28,36,43). Fourteen studies confirmed a diagnosis of HV using radiographs (26,27,29,31–36,38–41,43). Chronic conditions (such as RA) were not considered to be an exclusion criterion, but no eligible studies were found that focused on these populations. Follow-up periods ranged from 3 weeks to 3 years. Three included studies investigated the effects of foot orthoses compared to no treatment (35,39,41), and 1 of these studies included a third study group that underwent surgery (41). Six studies investigated the effectiveness of other devices, including various designs of dynamic splints (26,37,40) and night splints (31,32,38). Seven studies investigated physical therapies, including manual therapy (29,30,34), foot exercises (33,36), taping (33), or a combination (27,28,33). One study investigated intramuscular botulinum toxin type A (Botox) injections (43), and another study reported the effects of marigold therapy (*Tagetes patula* paste with a felt pad) (25).

In addition to the primary outcomes (HV angle or pain), other physical outcome measures included intermetatarsal angle, abductor hallucis cross-sectional area, first MTP joint range of motion, pressure-pain threshold, postural sway, toe grip strength, and hallux plantarflexion or abduction strength. A wide range of self-reported outcome scales was used, including VAS and NRS, the Foot Function Index, the hallux-metatarsophalangeal scale of the American Orthopaedic Foot and Ankle Society, the Foot and Ankle Outcome Scale, and the Quality of Life Scale. Two categorical variables reported by 1 study (41) were converted into dichotomous data for extraction and analysis (footwear problems: none = no, and moderate/severe = yes; global foot assessment: better = yes, and as good as/worse = no).

Risk of bias. Results of the risk-of-bias assessment are outlined in Figure 2. The most common domains raising serious

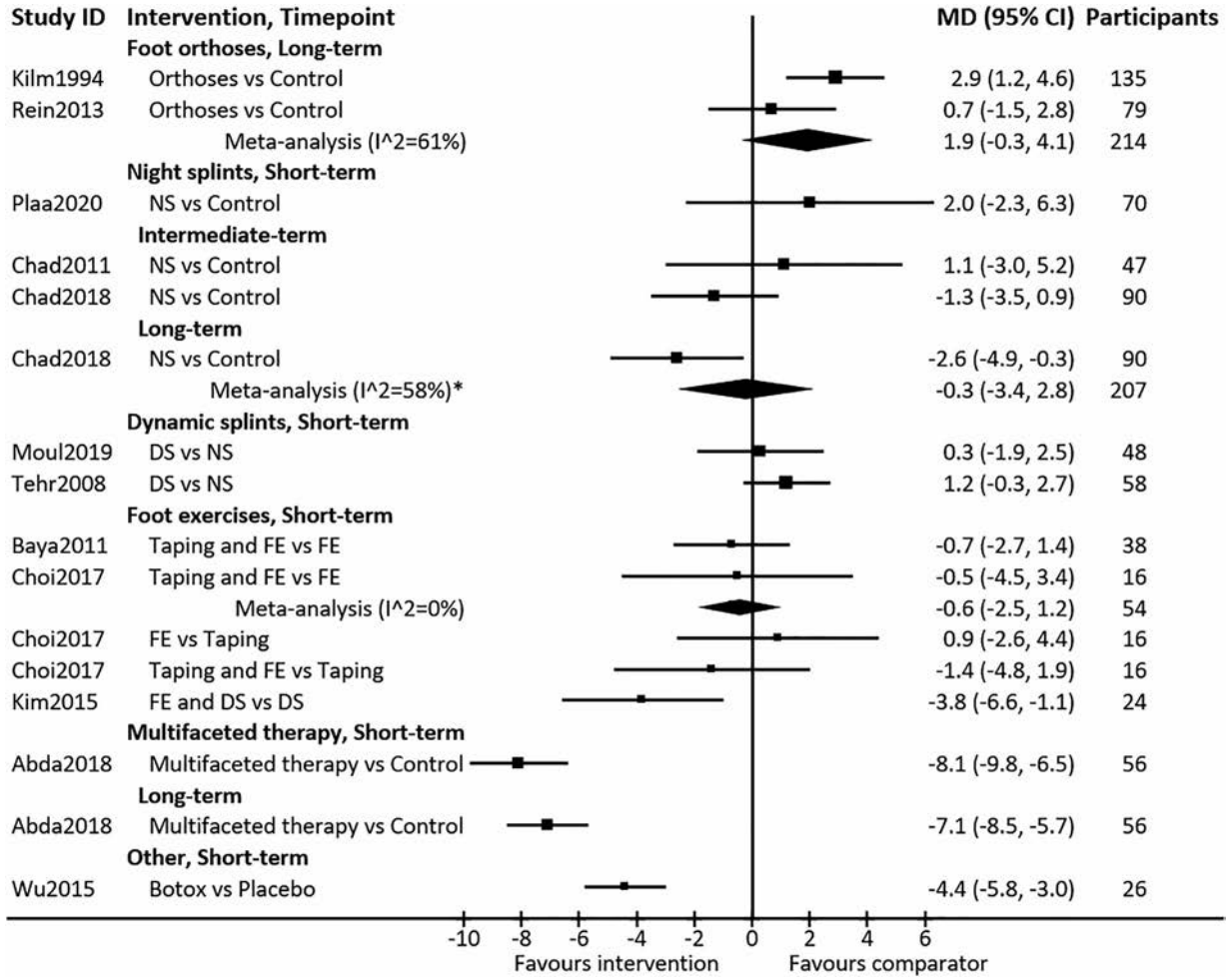

Figure 3. Forest plot showing individual study-effect estimates and meta-analyses for hallux valgus angle. 95% CI = 95% confidence interval; DS = dynamic splints; FE = foot exercises; MD = mean difference; NS = night splints. * = meta-analysis based on 3 studies (Plaass et al, 2020 [38], Chadchavalpanichaya and Chueluecha, 2011 [31], and Chadchavalpanichaya et al, 2018 long-term [32]).

concerns were deviations from intended intervention, where 5 studies (31%) (28–30,33,40) did not report adherence, and missing outcome data, where 5 studies (31%) (28–30,33,40) did not report how many participants were included in the final analysis and/or there was unclear reporting of study dropouts. Only 3 studies (19%) had a registered clinical trial protocol (32,38,43). No studies had a prespecified statistical analysis plan with sufficient detail, and thus all studies were judged as having at least "some concerns" in the fifth domain of "selection of the reported result."

Visual inspection of funnel plots by intervention category revealed asymmetry in the funnel plot for physical therapies, indicating bias toward publication of studies showing a positive effect of the experimental intervention. Therefore publication bias was strongly suspected and was considered a factor in the GRADE assessment.

Effectiveness of nonsurgical interventions. Table 2 shows a summary of findings and GRADE certainty of the evidence across 10 comparisons for the primary outcomes. Forest plots in Figures 3 and 4 display individual study results (mean differences and 95% CIs) at different follow-up points for HV angle (Figure 3) and pain (Figure 4). For both primary outcomes, lower values indicate improvement, and therefore negative effect sizes demonstrate an effect in favor of the experimental intervention compared to the control group or comparison intervention. The crossover trial (37) did not provide sufficient statistical information for a paired analysis, thus the post-intervention mean was taken from each period and analyzed in a similar manner to parallel-group trials. This approach was considered preferable to excluding the trial, but it was not considered appropriate for inclusion in meta-analysis (16). Meta-analyses were performed for foot orthoses versus control (HV angle) (35,39), night splints versus control (HV angle and pain) (31,32,38), manual therapy versus placebo (pain) (29,30), and taping and foot exercises versus foot exercises (HV angle) (28,33). None of these meta-analyses showed significant differences between groups following intervention. However, significantly reduced HV angles were reported by 4 of 12 studies

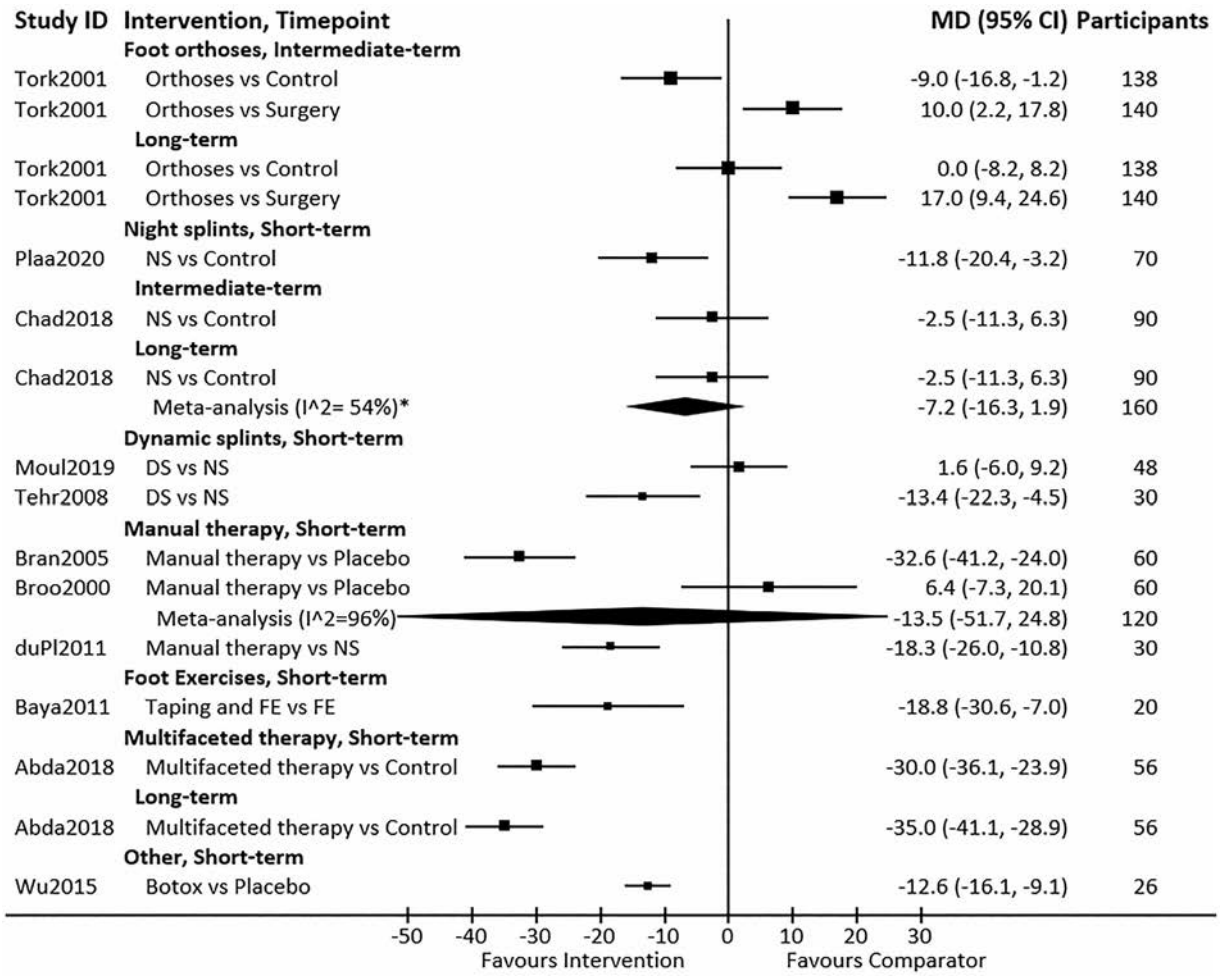

Figure 4. Forest plot showing individual study-effect estimates and meta-analyses for foot pain. 95% CI = 95% confidence interval; DS = dynamic splints; FE = foot exercises; MD = mean difference; NS = night splints. * = meta-analysis based on 2 studies (Plaass et al, 2020 [38] and Chadchavalpanichaya et al, 2018 long-term [32]).

that measured this outcome (27,32,36,43), and significantly reduced pain scores were reported by 8 of 11 studies that measured this outcome (27–29,34,38,40,41,43). Significant effects in favor of the experimental intervention were found across studies involving convenience samples (29,40), those recruited from outpatient clinics (27,32,34,38,41), samples of young adults (36,40), older adults (32), and mixed ages (27,29,34,38,41,43) (Table 1). A subgroup analysis of treatment effects according to age of study population could not be performed due to the lack of studies in juvenile populations.

Studies reported a wide range of secondary outcome measures (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24603). Self-reported function and satisfaction outcomes showed similar trends to self-reported pain outcomes. However, Torkki et al (41,42) reported a unique self-report variable, "footwear problems," reporting that those in the foot orthoses group experienced more footwear problems compared to the surgery group. One study reported that manual therapy increased first MTP joint range of motion compared to night splints (34). Foot exercises were shown to increase the cross-sectional area of the abductor hallucis muscle (36). Participants undertaking a multifaceted physical therapy program compared to controls showed improvements in toe grip, hallux plantarflexion, and hallux abduction strength (27).

DISCUSSION

This systematic review and meta-analysis aimed to synthesize the current state of the evidence for effectiveness of nonsurgical treatments for HV. Overall, results indicate a low level of certainty surrounding the effectiveness of nonsurgical interventions for HV, but reduction in pain appears more likely than improvement in HV angle. This synthesis will aid clinicians in evidence-based decision-making, given the vast range of nonsurgical treatments available for HV.

Meta-analyses showed that foot orthoses do not improve HV angle, and splints do not improve either HV angle or pain. Other

meta-analyses showed no significant short-term pain reduction with manual therapy, and no significant improvement in HV angle when taping was added to foot exercises. However, results from 8 studies showed significant pain reduction with foot orthoses (41), night splints (38), dynamic splints (40), manual therapy (29,34), taping added to foot exercises (28), multifaceted physical therapy (27), and Botox injections (43). Differences between groups at follow-up were clinically significant in each of these studies, with the exception of foot orthoses (41). A clinically significant reduction in HV angle was reported by 4 studies using night splints (32), foot exercises (36), multifaceted physical therapy (27), and Botox injections (43).

Conflicting results may be explained by the diversity of nonsurgical interventions employed, for example different types of splints, manual therapy techniques, or foot exercises. Study samples differed in terms of recruitment method, and a sample of patients seeking treatment for symptomatic moderate-to-severe HV may respond differently compared to a convenience sample of young adults or those with mild HV. Results should be interpreted in light of study characteristics outlined in Table 1.

This analysis builds on a previous systematic review of 2 randomized trials published in 2014 (11) by employing broader inclusion criteria and including 7 new studies published since 2014. This synthesis will inform patient-centered management for HV, including setting expectations for outcomes that are likely to be modifiable. While HV angle may not change significantly, pain is a potentially modifiable outcome, with other outcomes such as range of motion and muscle strength also potential indicators of improvement. Multifaceted interventions should be considered, incorporating a combination of footwear advice, foot orthoses, manual therapy, exercises, or Botox injections, depending upon examiner experience and patient preference, which are important pillars of evidence-based medicine, along with the best available evidence from the literature (44).

Limitations of this review should be considered when interpreting the findings. First, the search was limited to studies published in English, and despite this search filter, a small number of studies were found published in other languages and thus were excluded (45–48). Systematic bias is unlikely to be introduced by English-language restriction, but inclusion of more studies may have improved precision (49). The issue of publication bias must also be considered, as this bias was strongly suspected in the physical therapies. This review did include gray literature wherever possible, with 1 included thesis (30), but 2 other theses were unable to be sourced (50,51). Finally, there was a substantial amount of statistical heterogeneity across studies combined in the meta-analyses.

Acknowledging the overall low quality of the studies included in this review is important, as the low quality reduces the confidence in the reported effects. Many studies had significant risk of bias due to not reporting adherence, study dropouts, or missing data. There was poor reporting of group randomization and sample characteristics, and the high proportion of female participants may suggest recruitment bias. While participant blinding was not always possible, blinding of examiners was often unclear or not performed. Sample sizes were small, leading to imprecision of effect-size estimates and the potential for type 2 error. Over half of the studies (10 of 16) used follow-up periods of 3 months or less, and thus long-term effects were only reported for some interventions. Finally, investigation of potential harms or adverse outcomes were not reported, and thus could not be evaluated.

After reviewing the current state of the evidence, some recommendations for future research are presented below. First, several studies were limited by only measuring HV angle and not reporting on symptoms or self-reported function. Second, longer follow-up periods would be advisable, as only 6 studies followed participants for ≥12 months. Third, examining how different populations respond to interventions would be worthwhile. Only 1 study has been conducted in a juvenile HV population (35), another 3 studies included young adults (33,36,37), and some studies included both juvenile and adult patients (26,30). HV severity worsens with increasing age (1), and some interventions may be more effective in the early stages of deformity (10). Future trials should clearly report age and severity of HV of their participants. Finally, future trials should document adherence and adverse events to provide insights into the practicality, acceptability, and safety of these interventions.

In conclusion, this review provides guidance to clinicians regarding nonsurgical options for treatment of HV and provides recommendations for planning future clinical trials in this area. There is a very low to low level of certainty around most effect estimates presented, but there is moderate certainty that foot orthoses may reduce pain in the intermediate term, and moderate certainty that pain and HV angle may improve with Botox injections. Further high-quality randomized trials with adequate sample sizes and robust methodology are needed to investigate nonsurgical treatments for HV given the high cost of surgery to the health care system (9) and the potential benefits of early intervention (36).

ACKNOWLEDGMENTS

We would like to thank Professor Robert Ware for his statistical advice in relation to the meta-analysis and Dr. Jill Ferrari for her expert consultation during the initial search phase.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Hurn, Matthews, Munteanu, Menz.
Acquisition of data. Hurn, Matthews.
Analysis and interpretation of data. Hurn, Matthews, Munteanu, Menz.

REFERENCES

1. Nix S, Smith M, Vicenzino B. Prevalence of hallux valgus in the general population: a systematic review and meta-analysis. J Foot Ankle Res 2010;3:21.
2. D'Arcangelo P, Landorf KB, Munteanu SE, Zammit GV, Menz HB. Radiographic correlates of hallux valgus severity in older people. J Foot Ankle Res 2010;3:20.
3. Mochizuki T, Yano K, Ikari K, Hiroshima R, Ishibashi M, Okazaki K. Relationship of callosities of the forefoot with foot deformity, Health Assessment Questionnaire disability index, and joint damage score in patients with rheumatoid arthritis. Mod Rheumatol 2020;30:287–92.
4. Hurn SE, Vicenzino B, Smith MD. Functional impairments characterizing mild, moderate, and severe hallux valgus. Arthritis Care Res (Hoboken) 2015;67:80–8.
5. Menz HB, Morris ME, Lord SR. Foot and ankle characteristics associated with impaired balance and functional ability in older people. J Gerontol A Biol Sci Med Sci 2005;60:1546–52.
6. Menz HB, Morris ME, Lord SR. Foot and ankle risk factors for falls in older people: a prospective study. J Gerontol A Biol Sci Med Sci 2006;61:866–70.
7. Mickle KJ, Munro BJ, Lord SR, Menz HB, Steele JR. ISB Clinical Biomechanics Award 2009: toe weakness and deformity increase the risk of falls in older people. Clin Biomech (Bristol, Avon) 2009;24:787–91.
8. Nix S, Vicenzino B, Smith M. Foot pain and functional limitation in healthy adults with hallux valgus: a cross-sectional study. BMC Musculoskelet Disord 2012;13:197.
9. Belatti DA, Phisitkul P. Economic burden of foot and ankle surgery in the US Medicare population. Foot Ankle Int 2014;35:334–40.
10. Hurn SE, Vicenzino BT, Smith MD. Nonsurgical treatment of hallux valgus: a current practice survey of Australian podiatrists. J Foot Ankle Res 2016;9:16.
11. Ferrari J. Hallux valgus (bunions). BMJ Clin Evid 2014;2014:1112.
12. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Br Med J 2009;339:b2535.
13. Merlin T, Weston A, Tooher R. Extending an evidence hierarchy to include topics other than treatment: revising the Australian 'levels of evidence'. BMC Med Res Methodol 2009;9:34.
14. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. Br Med J 2014;348:g1687.
15. Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. Br Med J 2019 366:l4898.
16. Higgins J, Eldridge S, Li T. Including variants on randomized trials. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V, editors. Cochrane handbook for systematic reviews of interventions. URL: www.training.cochrane.org/handbook.
17. Elbourne DR, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. Int J Epidemiol 2002;31:140–9.
18. Review Manager (RevMan). URL: training.cochrane.org/online-learning/core-software-cochrane-reviews/revman.
19. Schünemann H, Brożek J, Guyatt G, Oxman A. GRADE handbook: introduction to GRADE handbook. URL: https://gdt.gradepro.org/app/handbook/handbook.html.
20. Nix S, Russell T, Vicenzino B, Smith M. Validity and reliability of hallux valgus angle measured on digital photographs. J Orthop Sports Phys Ther 2012;42:642–8.
21. Desai S, Peterson AC, Wing K, Younger A, Crump T, Liu G, et al. Minimally important difference in the Foot and Ankle Outcome Score among patients undergoing hallux valgus surgery. Foot Ankle Int 2019;40:694–701.
22. Landorf KB, Radford JA. Minimal important difference: values for the Foot Health Status Questionnaire, Foot Function Index and Visual Analogue Scale. Foot 2008;18:15–9.
23. Sutton RM, McDonald EL, Shakked RJ, Fuchs D, Raikin SM. Determination of minimum clinically important difference (MCID) in visual analog scale (VAS) pain and Foot and Ankle Ability Measure (FAAM) scores after hallux valgus surgery. Foot Ankle Int 2019;40:687–93.
24. Deeks JJ, Higgins JP, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane handbook for systematic reviews of interventions. URL: www.training.cochrane.org/handbook.
25. Khan MT. The podiatric treatment of hallux abducto valgus and its associated condition, bunion, with Tagetes patula. J Pharm Pharmacol 1996;48:768–70.
26. Mirzashahi B, Ahmadifar M, Birjandi M, Pournia Y. Comparison of designed slippers splints with the splints available on the market in the treatment of hallux valgus. Acta Med Iran 2012;50:107–112.
27. Abdalbary SA. Foot mobilization and exercise program combined with toe separator improves outcomes in women with moderate hallux valgus at 1-year follow-up: a randomized clinical trial. J Am Podiatr Med Assoc 2018;108:478–86.
28. Bayar B, Erel S, Simsek IE, Sumer E, Bayar K. The effects of taping and foot exercises on patients with hallux valgus: a preliminary study. Turk J Med Sci 2011;41:403–9.
29. Brantingham JW, Guiry S, Kretzmann HH, Kite VJ, Globe G. A pilot study of the efficacy of a conservative chiropractic protocol using graded mobilization, manipulation and ice in the treatment of symptomatic hallux abductovalgus bunion. Clin Chiropr 2005;8:117–33.
30. Broodryk M. The efficacy of strain counterstrain mobilization in patients with painful hallux abducto valgus bunions [masters thesis]. Durban (South Africa): Technikon Natal; 2000.
31. Chadchavalpanichaya N, Chueluecha C. Effectiveness of hallux valgus strap: a prospective, randomized single-blinded controlled trial. Siriraj Med J 2011;63:42–6.
32. Chadchavalpanichaya N, Prakotmongkol V, Polhan N, Rayothee P, Seng-lad S. Effectiveness of the custom-mold room temperature vulcanizing silicone toe separator on hallux valgus: a prospective, randomized single-blinded controlled trial. Prosthet Orthot Int 2018;42:163–70.
33. Choi JH. Effects of kinesio taping and stretching on hallux valgus angle and balance in female hallux valgus patients. Res J Pharm Tech 2017;10:2926–30.
34. Du Plessis M, Zipfel B, Brantingham JW, Parkin-Smith GF, Birdsey P, Globe G, et al. Manual and manipulative therapy compared to night splint for symptomatic hallux abducto valgus: an exploratory randomised clinical trial. Foot 2011;21:71–8.
35. Kilmartin TE, Barrington RL, Wallace WA. A controlled prospective trial of a foot orthosis for juvenile hallux valgus. J Bone Joint Surg Br 1994;76:210–4.
36. Kim MH, Yi CH, Weon JH, Cynn HS, Jung DY, Kwon OY. Effect of toe-spread-out exercise on hallux valgus angle and cross-sectional area of abductor hallucis muscle in subjects with hallux valgus. J Phys Ther Sci 2015;27:1019–22.
37. Moulodi N, Kamyab M, Farzadi M. A comparison of the hallux valgus angle, range of motion, and patient satisfaction after use of dynamic and static orthoses. Foot 2019;41:6–11.
38. Plaass C, Karch A, Koch A, Wiederhoeft V, Ettinger S, Claassen L, et al. Short term results of dynamic splinting for hallux valgus: a prospective randomized study. Foot Ankle Surg 2020;26:146–50.

39. Reina M, Lafuente G, Munuera PV. Effect of custom-made foot orthoses in female hallux valgus after one-year follow up. Prosthet Orthot Int 2013;37:113–9.
40. Tehraninasr A, Saeedi H, Forogh B, Bahramizadeh M, Keyhani MR. Effects of insole with toe-separator and night splint on patients with painful hallux valgus: a comparative study. Prosthet Orthot Int 2008;32:79–83.
41. Torkki M, Malmivaara A, Seitsalo S, Hoikka V, Laippala P, Paavolainen P. Surgery vs orthosis vs watchful waiting for hallux valgus: a randomized controlled trial. JAMA 2001;285:2474–80.
42. Torkki M, Malmivaara A, Seitsalo S, Hoikka V, Laippala P, Paavolainen P. Hallux valgus: immediate operation versus 1 year of waiting with or without orthoses: a randomized controlled trial of 209 patients. Acta Orthop Scand 2003;74:209–15.
43. Wu KP, Chen CK, Lin SC, Pei YC, Lin RH, Tsai WC, et al. Botulinum toxin type A injections for patients with painful hallux valgus: a double-blind, randomized controlled study. Clin Neurol Neurosurg 2015;129 Suppl 1:S58–62.
44. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. Br Med J 1996;312:71.
45. Bek N, Kurklu B. Comparison of different conservative treatment approachs in patients with hallux valgus. Artroplasti Artroskopik Cerrahi 2002;13:90–3.
46. Jeon MY, Jeong HC, Jeong MS, Lee YJ, Kim JO, Lee ST, et al. Effects of taping therapy on the deformed angle of the foot and pain in hallux valgus patients. J Korean Acad Nurs 2004;34:685–92.
47. Lee H, Kim E, Park I, Bae M, Shin J, Lee J, et al. The effects of combined exercises of elastic-band and short foot exercise on plantar foot pressure, toe angle and balance for patients with low to moderate hallux valgus. J Korean Soc Integrat Med 2015;3:73–88.
48. Milachowski KA, Krauss AJ. Comparing radiological examinations between hallux valgus night brace and a new dynamic orthosis for correction of the hallux valgus. Fuss Sprunggelenk 2008; 61:14–8.
49. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care 2012;28:138–44.
50. Juriansz A. Conservative treatment of hallux valgus: a randomised controlled clinical trial of a hallux valgus night splint [dissertation]. London: King's College London; 1996.
51. Khan MT. Clinical study to investigate the effect of marigold therapy in the treatment of hallux abducto valgus and its associated condition, bunion [thesis]. London: University of East London; 1995.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1689–1695
DOI 10.1002/acr.24717

American College of Rheumatology
Empowering Rheumatology Professionals

Risk of Comorbidities Following Physician-Diagnosed Knee or Hip Osteoarthritis: A Register-Based Cohort Study

Andrea Dell'Isola,[1] Kenneth Pihl,[1] Aleksandra Turkiewicz,[1] Velocity Hughes,[1] Weiya Zhang,[2] Sita Bierma-Zeinstra,[3] Daniel Prieto-Alhambra,[4] and Martin Englund[1]

Objective. To estimate the risk of developing comorbidities in patients after physician-diagnosed knee or hip osteoarthritis (OA).

Methods. This was a cohort study using Swedish longitudinal health care register data; we studied residents in the Skåne region age ≥35 years on January 1, 2010 who were free from diagnosed hip or knee OA (n = 548,681). We then identified subjects with at least 1 new diagnosis of knee or hip OA (incident OA) between 2010 and 2017 (n = 50,942 considered exposed). Subjects without diagnosed OA were considered unexposed. From January 2010 both unexposed and exposed subjects were observed for the occurrence of 18 different predefined comorbidities until either relocation outside of the region, death, occurrence of the comorbidity, or December 2017, whichever came first. We calculated unadjusted hazard ratios (HRs) and adjusted HRs of comorbidities using Cox models with knee and hip OA as time-varying exposures.

Results. Subjects with incident knee or hip OA had 7% to 60% higher adjusted HRs (range 1.07–1.60) of depression, cardiovascular diseases, back pain, and osteoporosis than individuals without an OA diagnosis. An increased risk of diabetes mellitus was found only for knee OA (adjusted HR 1.19 [95% confidence interval 1.13–1.26]). For the rest of the diagnoses, we found either no increased risk or estimates with wide confidence intervals, excluding clear interpretations of the direction or size of effects.

Conclusion. Incident physician-diagnosed knee and hip OA is associated with an increased risk of depression, cardiovascular diseases, back pain, osteoporosis, and diabetes mellitus. However, the latter was only found for knee OA.

INTRODUCTION

Osteoarthritis (OA) is one of the most common chronic conditions and ranks 12th among 359 specific diagnoses contributing the most to global disability (1). Knee and hip OA are responsible for the largest burden caused by OA (2) and are associated with substantial joint pain and reduced function and quality of life in populations worldwide (2). Historically, OA has been considered a joint-specific "wear and tear" degenerative disease; however, recent research has revealed that it is a complex disorder with multiple genetic, constitutional, and environmental risk factors (3), which may also increase the risk of other chronic conditions. In this regard, a recent systematic review reported a pooled comorbidity prevalence of 67% in people with OA; approximately 20% higher than age and sex-matched controls without OA (4).

Common comorbidities among people with OA include conditions of the cardiovascular, neurologic, endocrine, and psychological systems (4,5). Several mechanisms may explain these relations, including obesity and the often reported low-grade inflammation associated with OA, which is hypothesized to increase the risk of cardiovascular disease and diabetes mellitus (6), as well as pain and disability that may limit physical activity, influencing other risk factors for chronic conditions, e.g., further weight gain (6). Furthermore, apart from a few studies on cardiovascular diseases and diabetes mellitus (7), research on the association between OA and specific comorbidities has often been limited to cross-sectional studies (4,5,8), restricting any

Dr. Bierma-Zeinstra's work was supported by The Netherlands Organization for Health Research and Development, the European Union, FOREUM, and the Dutch Arthritis Association.

[1]Andrea Dell'Isola, PhD, Kenneth Pihl, PhD, Aleksandra Turkiewicz, PhD, Velocity Hughes, PhD, Martin Englund, PhD: Lund University, Lund, Sweden; [2]Weiya Zhang, PhD: School of Medicine, University of Nottingham, Nottingham, UK; [3]Sita Bierma-Zeinstra, PhD: Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; [4]Daniel Prieto-Alhambra, PhD: University of Oxford, Oxford, UK.

Drs. Dell'Isola and Pihl contributed equally to this work.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24717&file=acr24717-sup-0001-Disclosureform.pdf.

Address correspondence to Andrea Dell'Isola, PhD, Clinical Epidemiology Unit, Orthopedics, Department of Clinical Sciences, Lund, Wigerthuset, Remissgatan 4, Lund 22185, Sweden. Email: andrea.dellisola@med.lu.se.

Submitted for publication January 7, 2021; accepted in revised form May 25, 2021.

SIGNIFICANCE & INNOVATIONS

- In this longitudinal study of approximately half a million Swedish residents, we investigated the temporality between incident knee or hip osteoarthritis (OA) and comorbidity. We found that people with incident physician-diagnosed knee or hip OA are at increased risks of subsequent diagnoses of depression, cardiovascular diseases, diabetes mellitus, and back pain.
- Our results highlight the importance of considering knee and hip OA as clinically relevant and potentially modifiable risk factors in the prevention of other chronic conditions, including cardiovascular diseases, diabetes mellitus, depression, and back pain.

interpretation of the temporal or potentially causal relationship. New knowledge, taking into account the time sequence between OA and other comorbidities may be an important initial step to shed light on any temporality between OA and comorbidities. Therefore, the main aim of this explorative study was to estimate the hazard of developing a number of specific comorbidities in people with incident physician-diagnosed knee or hip OA compared to people without OA using a longitudinal study design.

MATERIALS AND METHODS

Data sources. We conducted a cohort study using data from 3 registers that comprise the entire population of Skåne, the southernmost region in Sweden, with approximately 1.23 million inhabitants (one-eighth of the Swedish population) in the year 2009. From the Swedish Population Register we retrieved data on age, sex, residential addresses, and deaths, while individual-level data on income, education, marital status, and country of birth were retrieved from the Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA by the Swedish acronym). Last, from the Skåne Healthcare Register (SHR), we extracted information about diagnoses at any health care visit. SHR is a regional mandatory register that contains the publicly practicing physicians' diagnostic codes according to the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10). These codes are assigned at the time of the health care visit by the physicians themselves and are automatically transferred to the register from the electronic medical records. The positive predictive value of a knee OA diagnosis in SHR has previously been reported to be 88% (9). All data from the different registers were linked through the coded personal unique identification number that is assigned to all residents in Sweden by the Swedish Tax Agency. The study was approved by the Regional Ethical Review Board in Lund, Sweden, and is reported according to the Strengthening the Reporting of Observational studies in Epidemiology guideline (10).

Study design, exposures, and outcomes. The cohort consisted of individuals age ≥35 years on January 1, 2010 who were residents in the Skåne region between January 1, 1998 and January 1, 2010. Only people with at least 1 health care visit with any diagnosis registered during this period (96% of eligible persons) were included, to minimize potential confounding due to propensity to seek care. People had to be at risk for both exposures (knee/hip OA, i.e., individuals were excluded if they had prevalent knee or hip OA) and for an outcome of interest (one of the 18 comorbidities) on January 1, 2010, and they were followed up until relocation outside of the region, death, a diagnosis of the comorbidity of interest, or December 31, 2017, whichever occurred first.

The exposure of interest was incident physician-diagnosed knee and hip OA (ICD-10 codes M17 and M16). We defined individuals as exposed if no diagnosis was recorded between January 1, 1998 and December 31, 2009 and a new diagnosis was received between January 1, 2010 and December 31, 2017 (Figure 1). Individuals with both knee and hip OA diagnoses were classified according to the first diagnostic code they received. Those with no record of either knee or hip OA diagnosis were defined as not exposed. The time from the start of follow-up to the date of the first knee or hip OA diagnosis (index date) was treated as unexposed, while the time after the OA diagnosis was treated as exposed (Figure 1).

The outcome of interest was a new diagnosis (ICD-10 code) of any of the following 18 conditions between January 1, 2010 and December 31, 2017: depression, Alzheimer's disease, other dementia, hypertension, ischemic heart diseases, heart failure,

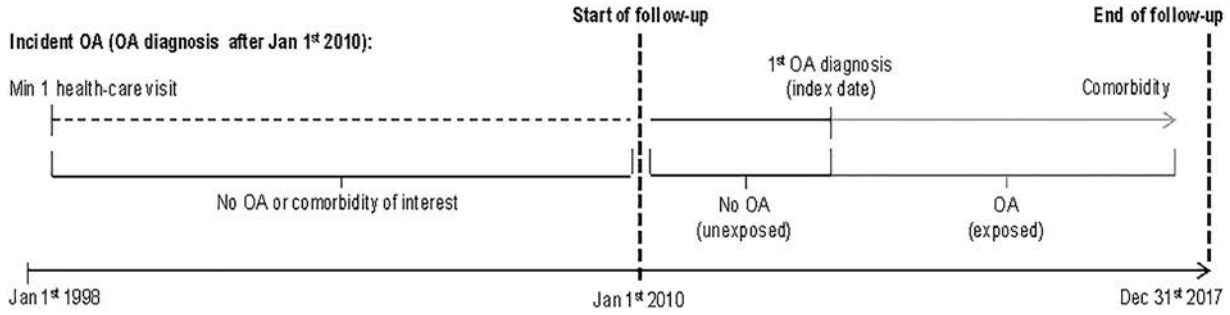

Figure 1. Study design for incident knee or hip osteoarthritis (OA). Min = minimum.

Table 1. Descriptive characteristics of the study cohort at start of follow-up (January 1, 2010)*

Characteristic (ICD-10 code)	No OA (n = 497,739)	Incident knee OA (n = 36,465)	Incident hip OA (n = 14,477)
Age at beginning of follow-up, mean ± SD years	57.3 ± 14.6	62.2 ± 12.2	65.3 ± 11.7
Women	254,593 (51)	21,553 (59)	8,306 (57)
Married†	350,172 (70)	27,955 (77)	10,937 (76)
Born in Sweden‡	422,713 (85)	31,292 (86)	12,949 (89)
Education up to 9 years§	128,738 (26)	11,364 (31)	4,745 (33)
Education, 10–12 years§	222,147 (45)	16,390 (45)	6,050 (42)
Education, 13–14 years§	59,950 (12)	3,908 (11)	1,556 (11)
Education ≥15 years§	83,620 (17)	4,590 (13)	2,065 (14)
Income in 100,000 SEK, median (interquartile range)¶	2.0 (1.3–2.7)	1.8 (1.3–2.5)	1.7 (1.3–2.5)
Alcohol-related disorders (F10)	11,694 (2)	667 (2)	322 (2)
Depression (F32.0–F33.9)	43,413 (9)	3,554 (10)	1,350 (9)
Alzheimer's disease (F00.0–F00.9, G30.0–F30.9)	2,425 (0)	91 (0)	30 (0)
Other dementia (F01.0–F03.0, G31.0–G31.1, G31.8–G32.8)	428 (0)	12 (0)	4 (0)
Hypertension (I10.0–I15.9)	99,937 (20)	10,440 (29)	4,569 (32)
Ischemic heart diseases (I20.0–I25.9)	39,246 (8)	3,491 (10)	1,733 (12)
Heart failure (I50.0–I50.9)	17,509 (4)	1,171 (3)	582 (4)
Cerebrovascular disease (I60.0–I69.8)	22,072 (4)	1,550 (4)	780 (5)
Diabetes mellitus (E10.0–E14.9)	34,616 (7)	3,054 (8)	1,255 (9)
Lung cancer (tracheal, bronchus, and lung: C33.0–C34.9)	1,206 (0)	43 (0)	33 (0)
Colorectal cancer (C18.0–C21.8)	2,906 (1)	230 (1)	111 (1)
Breast cancer, women only (C50.0–C50.9)	7,508 (3)	754 (4)	340 (4)
Prostate cancer, men only (C61)	7,634 (3)	644 (4)	351 (6)
Hip fracture (S72.0–S72.2)	7,406 (1)	327 (1)	334 (2)
Forearm fracture (S52.0–S52.9)	17,082 (3)	1,574 (4)	672 (5)
Ankle fracture (S82.3, S82.5–S82.6, S82.8)	8,370 (2)	699 (2)	307 (2)
Back pain (G54.2, G54.4, M47–M49, M49.2–M51.9, M53, M54.9, M99–M99.04, M99.1–M99.14, M99.2–M99.24, M99.3–M99.34, M99.4–M99.44, M99.5–M99.54, M99.6, M99.64, M99.7–M99.74, M99.8–M99.84)	88,552 (18)	8,827 (24)	3,672 (25)
Osteoporosis (M80.0–M82.8)	14,621 (3)	1,467 (4)	716 (5)
Chronic lower respiratory diseases (COPD, bronchitis, emphysema: J40–J44.9)	5,526 (1)	463 (1)	219 (2)

* Values are the number (%) unless indicated otherwise. Comorbidity codes are from the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10), at the start of follow-up as registered in the Skåne Healthcare Register during 1998–2009. COPD = chronic obstructive pulmonary disease; OA = osteoarthritis; SEK = Swedish Krona.
† Missing: n = 29.
‡ Missing: n = 44.
§ Missing: n = 3,558.
¶ Missing: n = 29.

cerebrovascular disease, diabetes mellitus, lung cancer (tracheal, bronchus, and lung), colorectal cancer, breast (in women only) and prostate cancer (in men only), fracture to the hip (neck of femur, pertrochanteric, and subtrochanteric), fracture to the forearm (radius and ulna), fracture to the ankle (lower end of tibia, and medial and lateral malleolus), back pain (neck and low back), osteoporosis, and chronic lower respiratory diseases (chronic obstructive pulmonary disease, bronchitis, and emphysema) (Table 1). These are among the most frequent conditions coexisting with OA (4,5,11), the most common cancer types (12), and the most common fractures among the elderly (13,14). In the analysis of the incidence of each comorbidity, individuals with a diagnostic code of that specific condition between January 1, 1998 and December 31, 2009 were excluded from the calculation of the hazard ratio (HR) for that comorbidity. In a sensitivity analysis, we analyzed people with prevalent OA, which was defined as having at least 1 diagnostic code for knee OA (ICD-10 code M17) or hip OA (M16) in the period between January 1, 1998 and December 31, 2009. In this case, the exposure started at the beginning of the follow-up time (January 1, 2010).

Confounders. Sex, age, alcohol-related disorders, marital status, if born in Sweden, residential area, income, education in years, and comorbidities were considered confounders, as they can potentially influence both exposure (incident OA) and outcome (incident comorbidity). For included individuals, information on income, education, marital status, and country of birth, as reported in the year 2009, was retrieved from the LISA register held by Statistics Sweden. We categorized education according to its length: <10 years, 10–12 years, 13–14 years, and ≥15 years. Marital status (married/registered partner or other) and country of birth (Sweden or outside Sweden) were binary, while income was continuous. Residential area was extracted

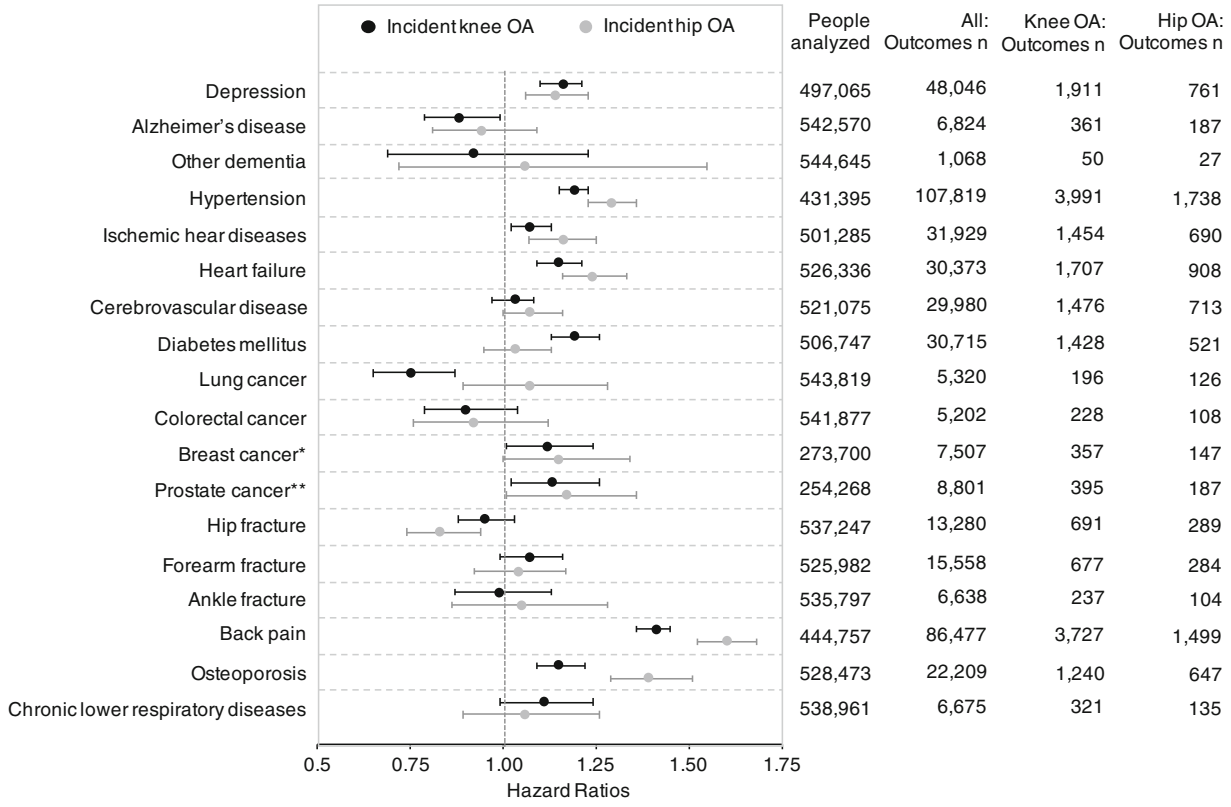

Figure 2. Adjusted hazard ratios of consultation for diseases occurring in persons with incident doctor-diagnosed knee or hip osteoarthritis (OA) compared to persons without OA. * = only women included in analysis; ** = only men included in analysis. Error bars show 95% confidence intervals. Complete crude and adjusted hazard ratios are available in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24717.

from the Swedish Population Register and was included as municipality. Information on diagnosed alcohol-related disorders (ICD-10 code F10) and all outcome conditions between January 1, 1998 and December 31, 2009 was retrieved from the SHR. Information on all covariates was collected up to the beginning of follow-up (January 1, 2010) and was not updated afterward to avoid adjusting for intermediates. Marital status and income were missing for 0.006% of all included individuals, while the country of birth was missing for 0.009% and education was missing for 0.7%. Those with missing data were a very low proportion and were thus excluded from the adjusted analyses.

Statistical analysis. Descriptive baseline data by exposure status are reported as means ± SDs, medians with interquartile ranges, or numbers with percentages as appropriate. We used the Cox proportional hazards model with calendar years as the time scale to estimate the HR of a diagnosis of each of the conditions of interest. Separate models were used for each outcome condition in which individuals diagnosed with the specific comorbidity of interest before January 1, 2010 were excluded. We adjusted all the analyses for baseline age, sex, socioeconomic status (residential area, income, and education), birth outside of Sweden, marital status, and the presence of a diagnosis of alcohol-related disorders or any of the outcome conditions (comorbidities) apart from the one analyzed. The proportional hazards assumption was evaluated using plots of Schoenfeld residuals, and no violations were detected (15). We repeated all analyses in sensitivity analyses using a stricter definition of knee and hip OA exposure (i.e., at least 2 diagnostic codes of knee or hip OA to be classified as exposed).

RESULTS

Study cohort. We included 548,681 Skåne residents with no prior records of knee or hip OA and age ≥35 years at the start of follow-up. Of these, 36,465 and 14,477 individuals were registered with at least 1 physician-recorded diagnostic code of knee or hip OA, respectively, during follow-up, i.e., considered incident OA. Those with incident OA were on average older at the start of follow-up than individuals without OA, were a higher proportion of women, and generally had a higher prevalence of comorbidities (Table 1).

Risk of comorbidities. The most common comorbidities during follow-up were depression, cardiovascular diseases, and back pain (Figure 2). In the adjusted analyses, individuals with newly diagnosed knee or hip OA had 7% to 60% higher hazards

of depression (knee OA adjusted HR 1.16 [95% confidence interval (95% CI) 1.10–1.21]; hip OA adjusted HR 1.14 [95% CI 1.06–1.23]), hypertension (knee OA adjusted HR 1.19 [95% CI 1.15–1.23]; hip OA adjusted HR 1.29 [95% CI 1.23–1.36]), ischemic heart diseases (knee OA adjusted HR 1.07 [95% CI 1.02–1.13]; hip OA adjusted HR 1.16 [95% CI 1.07–1.25]), heart failure (knee OA adjusted HR 1.15 [95% CI 1.09–1.21]; hip OA adjusted HR 1.24 [95% CI 1.16–1.33]), back pain (knee OA adjusted HR 1.41 [95% CI 1.36–1.45]; hip OA adjusted HR 1.60 [95% CI 1.52–1.68]), and osteoporosis (knee OA adjusted HR 1.15 [95% CI 1.09–1.22]; hip OA adjusted HR 1.39 [95% CI 1.29–1.51]) than individuals without an OA diagnosis (Figure 2 and Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24717). Only for knee OA, there was an association with diabetes mellitus (adjusted HR 1.19 [95% CI 1.13–1.26]), Alzheimer's disease (adjusted HR 0.88 [95% CI 0.79–0.99]), and lung cancer (adjusted HR 0.75 [95% CI 0.65–0.87]). Individuals with incident hip OA had lower hazards of hip fracture than those without OA (adjusted HR 0.83 [95% CI 0.74–0.94]), while no association was found for knee OA. For the rest of the diagnoses, we found either no association for individuals with knee or hip OA (i.e., HR estimates and their CIs close to 1.00) or estimates with wide CIs, excluding any clear interpretations of the direction or size of the effects (Figure 2).

In the sensitivity analysis for prevalent knee or hip OA, the observed associations were generally lower (i.e., HR closer to 1.00) than those observed among incident OA cases; however, for most comorbidities, the interpretations were similar (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24717). In the sensitivity analyses requiring 2 diagnostic OA codes (favoring high specificity of exposed), HRs were generally closer to 1.00 and had wider CIs than in the primary analyses (see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24717). Nevertheless, the interpretation did not change for any of the comorbidities in the main analysis.

DISCUSSION

Our findings of this explorative population-based study of approximately half a million residents in Sweden show that people with physician-diagnosed OA are at increased risks of subsequent diagnoses of depression, cardiovascular diseases, diabetes mellitus, and back pain. Despite the large sample size, estimates for most other conditions were inconclusive due to wide confidence intervals.

Several studies have previously found that knee and hip OA are associated with an increased risk of cardiovascular diseases (7,16,17). Our results support these findings and add to the increasing confidence that a diagnosis of knee or hip OA is associated with a higher risk of developing cardiovascular diseases. In contrast, the association of increased risk of diabetes mellitus among individuals with knee and hip OA is less supported by our findings (18). Kendzerska et al report a larger risk of diabetes mellitus for hip OA than knee OA (18), as opposed to our study, where only knee OA was associated with diabetes mellitus. Similarly, Swain et al reported a higher risk of diabetes mellitus only in people with knee OA (17). Several factors may explain these diverse results, including difference in exposure (incident OA versus prevalent OA), difference in adjustment (we could not adjust for body mass index [BMI]), and potential residual confounders.

Other surprising findings are that people with incident knee OA appear to have a lower risk of lung cancer than people without OA. This finding might be because knee OA, to a larger extent than hip OA, is a result of a prior knee injury (19) possibly reflecting a group being physically active and in general having a healthy lifestyle. The finding is supported by 2 previous studies suggesting a reduced risk of cancer mortality in physician-diagnosed and radiographic knee OA (20,21). However, we found people with knee OA to be at higher risk of developing other types of cancer such as breast and prostate cancer but not colorectal cancer, while people with hip OA had an increased risk of developing prostate cancer only. The positive association between cancer and OA has been previously reported in other cohorts and may hypothetically be explained by the presence of low-grade inflammation in people with OA (17,22). However, the reason why only certain types of cancers appear associated with OA is not clear.

Another interesting finding is the reduced risk of consulting for Alzheimer's disease in people with knee OA. The current evidence on OA and dementia is inconsistent but appears to point toward a negative effect of OA on cognitive health (17,23,24). However, a recent meta-analysis suggested that a diagnosis of OA is associated with higher cognitive scores at baseline and with a delayed cognitive decline (25). Persons with early symptoms of dementia (on the road to later becoming diagnosed with Alzheimer's disease) may consult for their knee problems to a lesser extent than persons mentally fully alert. Thus, those with a lack of consultation do not get their incipient OA diagnosed to the same extent, leading to a biased estimate of association. Finally, people with hip OA appear to have a reduced risk of hip fracture. This finding is in line with previous literature showing no or protective association between hip OA and hip fracture (26,27).

All in all, there is evidence to suggest that the associations found in our study are biologically plausible. Low-grade inflammation, obesity, and reduced physical activity are part of the pathogenesis of OA (28) and can increase the risk of other comorbidities (6,29). Low-grade inflammation has been reported to be associated with a series of other factors, including depression symptoms (e.g., sleep problems, low energy level, and widespread pain), high BMI, insulin resistance, cancer, Alzheimer's disease, and diabetes mellitus development (30,31). Despite the fact that most of these associations have no clear directionality

and are still debated in the scientific community, the presence of low-grade inflammation may partially explain the association between OA and comorbidities (32).

Obesity has been associated with OA through several pathophysiologic pathways, including low-grade inflammation due to the production of proinflammatory cytokines by the adipose tissue (33). Furthermore, strong evidence from a meta-analysis of longitudinal studies shows the existence of a bidirectional association between obesity and depression (34). This association becomes stronger when abdominal fat rather than BMI is used to define obesity and when metabolic dysregulation (e.g., hypertension, dyslipidemia, insulin resistance) is taken into account (35,36). All the evidence appears to point toward the existence of a vicious cycle around OA, reduced physical activity, increased weight, and low-grade inflammation, which can favor the development of some of the diseases analyzed here as a result of shared pathophysiology (37,38).

Our results can have important clinical implications. On the one hand, the co-occurrence of these diseases may represent a major obstacle in the treatment of each condition separately and calls for a more comprehensive approach. On the other hand, our results highlight the importance of considering knee and hip OA as clinically relevant and potentially modifiable risk factors in the prevention of other chronic conditions, including cardiovascular diseases, diabetes mellitus, depression, and back pain. For instance, treatment strategies targeting shared disease mechanisms might be beneficial for more than a single condition. For example, lifestyle interventions aimed at modifying dietary habits while promoting increased physical activity may potentially improve OA symptoms while preventing subsequent development of other diseases (39). Identifying interventions able to improve OA symptoms and reduce the risk of developing commodities is particularly important given the fact that people with OA and other coexisting conditions report worse pain and more restricted participation in social and domestic life than individuals with OA and no comorbidities (40,41). However, to identify appropriate strategies to prevent the onset of commodities, a better understanding of the causal mechanisms by which OA may cause other conditions is needed.

This study has important limitations. First, because of the observational nature of the study, we could not determine any causal relationships between OA and comorbidities, which should be acknowledged when considering the findings. Second, we were not able to adjust for BMI, which is a major risk factor for OA, diabetes mellitus, and cardiovascular diseases, and thus lack of adjustment may have biased our estimates upward. Third, as always, there may be a certain degree of misclassification of disease in the register, which may lead to either under- or overestimation of the risk of developing comorbidities. However, the validity of the diagnostic coding in the register has been reported to be high and is expected to be largely nondifferential (9,42). Finally, we focused on the effect of incident OA and we followed up the participants for a maximum of 8 years, and thus we cannot draw conclusions on the risk of developing comorbidity in people with longer OA duration. The strengths of our study include the large population-based cohort of an entire region, which supports the generalizability of our findings to the broader population of adults having knee or hip OA, as well as our sensitivity analyses yielding similar results to the primary analyses. Nevertheless, several of the associations found have not previously been investigated and thus need to be confirmed in future studies and different countries and populations.

In conclusion, physician-diagnosed knee and hip OA seem to be associated with an increased risk of depression, cardiovascular diseases, back pain, and osteoporosis. Only knee OA seems to be associated with an increased risk of diabetes mellitus, while hip OA, in general, is associated with larger risks of most comorbidities than is knee OA. The findings highlight the fact that prevention and early and effective treatments of OA may be important to avoid the development of other chronic conditions.

ACKNOWLEDGMENT

We would like to thank the registers used in this study for providing the data.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Dell'Isola, Pihl, Turkiewicz, Hughes, Zhang, Bierma-Zeinstra, Prieto-Alhambra, Englund.

Acquisition of data. Dell'Isola, Pihl, Turkiewicz, Englund.

Analysis and interpretation of data. Dell'Isola, Pihl, Turkiewicz.

REFERENCES

1. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392:1859–922.
2. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014;73:1323–30.
3. Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. Br Med Bull 2013;105:185–99.
4. Swain S, Sarmanova A, Coupland C, Doherty M, Zhang W. Comorbidities in osteoarthritis: a systematic review and meta-analysis of observational studies. Arthritis Care Res (Hoboken) 2020;72:991–1000.
5. Kadam UT, Jordan K, Croft PR. Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales. Ann Rheum Dis 2004;63:408–14.
6. Nuesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Juni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. BMJ 2011;342:d1165.

7. Williams A, Kamper SJ, Wiggers JH, O'Brien KM, Lee H, Wolfenden L, et al. Musculoskeletal conditions may increase the risk of chronic disease: a systematic review and meta-analysis of cohort studies. BMC Med 2018;16:167.
8. Van Dijk GM, Veenhof C, Schellevis F, Hulsmans H, Bakker JP, Arwert H, et al. Comorbidity, limitations in activities and pain in patients with osteoarthritis of the hip or knee. BMC Musculoskelet Disord 2008;9:95.
9. Turkiewicz A, Petersson IF, Bjork J, Hawker G, Dahlberg LE, Lohmander LS, et al. Current and future impact of osteoarthritis on health care: a population-based study with projections to year 2032. Osteoarthritis Cartilage 2014;22:1826–32.
10. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Int J Surg 2014;12:1500–24.
11. Muckelt PE, Roos E, Stokes M, McDonough S, Grønne D, Ewings S, et al. Comorbidities and their link with individual health status: a cross-sectional analysis of 23,892 people with knee and hip osteoarthritis from primary care. J Comorb 2020;10:2235042X20920456.
12. Fitzmaurice C, Akinyemiju TF, Al Lami FH, Alam T, Alizadeh-Navaei R, Allen C, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016: a systematic analysis for the Global Burden of Disease Study. JAMA Oncol 2018;4:1553–68.
13. Amin S, Achenbach SJ, Atkinson EJ, Khosla S, Melton LJ III. Trends in fracture incidence: a population-based study over 20 years. J Bone Miner Res 2014;29:581–9.
14. Curtis EM, van der Velde R, Moon RJ, van den Bergh JP, Geusens P, de Vries F, et al. Epidemiology of fractures in the United Kingdom 1988-2012: variation with age, sex, geography, ethnicity and socioeconomic status. Bone 2016;87:19–26.
15. Therneau TM. Modeling survival data: extending the Cox model. 2nd ed. Berlin: Springer; 2000.
16. Hall AJ, Stubbs B, Mamas MA, Myint PK, Smith TO. Association between osteoarthritis and cardiovascular disease: systematic review and meta-analysis. Eur J Prev Cardiol 2016;23:938–46.
17. Swain S, Coupland C, Mallen C, Kuo CF, Sarmanova A, Bierma-Zeinstra SM, et al. Temporal relationship between osteoarthritis and comorbidities: a combined case control and cohort study in the UK primary care setting. Rheumatology (Oxford) 2021;60:4327–39.
18. Kendzerska T, King LK, Lipscombe L, Croxford R, Stanaitis I, Hawker GA. The impact of hip and knee osteoarthritis on the subsequent risk of incident diabetes: a population-based cohort study. Diabetologia 2018;61:2290–9.
19. Poulsen E, Goncalves GH, Bricca A, Roos EM, Thorlund JB, Juhl CB. Knee osteoarthritis risk is increased 4-6 fold after knee injury: a systematic review and meta-analysis. Br J Sports Med 2019;53:1454–63.
20. Turkiewicz A, Kiadaliri AA, Englund M. Cause-specific mortality in osteoarthritis of peripheral joints. Osteoarthritis Cartilage 2019;27:848–54.
21. Mendy A, Park J, Vieira ER. Osteoarthritis and risk of mortality in the USA: a population-based cohort study. Int J Epidemiol 2018;47:1821–9.
22. Ziegler J. Cancer and arthritis share underlying processes. J Natl Cancer Inst 1998; 90: 802–3.
23. Ikram M, Innes K, Sambamoorthi U. Association of osteoarthritis and pain with Alzheimer's diseases and related dementias among older adults in the United States. Osteoarthritis Cartilage 2019;27:1470–80.
24. Innes KE, Sambamoorthi U. The association of osteoarthritis and related pain burden to incident Alzheimer's disease and related dementias: a retrospective cohort study of U.S. Medicare beneficiaries. J Alzheimers Dis 2020;75:789–805.
25. Mayburd AL, Baranova A. Increased lifespan, decreased mortality, and delayed cognitive decline in osteoarthritis. Sci Rep 2019;9:18639.
26. Chudyk AM, Ashe MC, Gorman E, Al Tunaiji HO, Crossley KM. Risk of hip fracture with hip or knee osteoarthritis: a systematic review. Clin Rheumatol 2012;31:749–57.
27. Yamamoto Y, Turkiewicz A, Wingstrand H, Englund M. Fragility fractures in patients with rheumatoid arthritis and osteoarthritis compared with the general population. J Rheumatol 2015;42:2055–8.
28. Pedersen BK, Saltin B. Exercise as medicine: evidence for prescribing exercise as therapy in 26 different chronic diseases. Scand J Med Sci Sports 2015;25 Suppl 3:1–72.
29. Schmidt M, Christiansen CF, Mehnert F, Rothman KJ, Sørensen HT. Non-steroidal anti-inflammatory drug use and risk of atrial fibrillation or flutter: population based case-control study. BMJ 2011;343:d3450.
30. Duncan BB, Schmidt MI, Pankow JS, Ballantyne CM, Couper D, Vigo A, et al. Low-grade systemic inflammation and the development of type 2 diabetes: the atherosclerosis risk in communities study. Diabetes 2003;52:1799–805.
31. Fried EI, von Stockert S, Haslbeck JM, Lamers F, Schoevers RA, Penninx BW. Using network analysis to examine links between individual depressive symptoms, inflammatory markers, and covariates. Psychol Med 2020;50:2682–90.
32. Robinson WH, Lepus CM, Wang Q, Raghu H, Mao R, Lindstrom TM, et al. Low-grade inflammation as a key mediator of the pathogenesis of osteoarthritis. Nat Rev Rheumatol 2016;12:580–92.
33. Osborn O, Olefsky JM. The cellular and signaling networks linking the immune system and metabolism in disease. Nat Med 2012;18:363–74.
34. Milaneschi Y, Simmons WK, van Rossum EF, Penninx BW. Depression and obesity: evidence of shared biological mechanisms. Mol psychiatry 2019;24:18–33.
35. Jokela M, Hamer M, Singh-Manoux A, Batty GD, Kivimäki M. Association of metabolically healthy obesity with depressive symptoms: pooled analysis of eight studies. Mol Psychiatry 2014;19:910–4.
36. Xu Q, Anderson D, Lurie-Beck J. The relationship between abdominal obesity and depression in the general population: a systematic review and meta-analysis. Obes Res Clin Pract 2011;5:e267–78.
37. Khan SS, Ning H, Wilkins JT, Allen N, Carnethon M, Berry JD, et al. Association of body mass index with lifetime risk of cardiovascular disease and compression of morbidity. JAMA Cardiol 2018;3:280–7.
38. Colditz GA, Willett WC, Rotnitzky A, Manson JE. Weight gain as a risk factor for clinical diabetes mellitus in women. Ann Intern Med 1995;122:481–6.
39. Skou ST, Pedersen BK, Abbott JH, Patterson B, Barton C. Physical activity and exercise therapy benefit more than just symptoms and impairments in people with hip and knee osteoarthritis. J Orthop Sports Phys Ther 2018;48:439–47.
40. Calders P, Van Ginckel A. Presence of comorbidities and prognosis of clinical symptoms in knee and/or hip osteoarthritis: a systematic review and meta-analysis. Semin Arthritis Rheum 2018;47:805–13.
41. Wilkie R, Blagojevic-Bucknall M, Jordan KP, Lacey R, McBeth J. Reasons why multimorbidity increases the risk of participation restriction in older adults with lower extremity osteoarthritis: a prospective cohort study in primary care. Arthritis Care Res (Hoboken) 2013;65:910–9.
42. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health 2011;11:450.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1696–1703
DOI 10.1002/acr.24642

Metacarpophalangeal Joint Impairment in Hand Osteoarthritis and Its Association With Mechanical Factors: Results From the Digital Cohort Osteoarthritis Design Hand Osteoarthritis Cohort

Inès Kouki,[1] Sophie Tuffet,[2] Michel D. Crema,[3] Alexandra Rousseau,[2] Pascal Richette,[4] Maxime Dougados,[5] Francis Berenbaum,[1] Jérémie Sellam,[1] and Alice Courties[1]

Objective. To determine the prevalence, distribution, and characteristics associated with radiographic metacarpophalangeal (MCP) joint osteoarthritis (OA).

Methods. This was a cross-sectional study of baseline data from the Digital Cohort Osteoarthritis Design, a French monocentric cohort including patients with symptomatic hand OA. We evaluated the prevalence of radiographic MCP joint OA, defined as ≥2 MCP joints with a Kellgren/Lawrence score of ≥2. We compared the prevalence of MCP joint OA in the dominant and nondominant hands. Associations between radiographic MCP joint OA and patient characteristics were studied using univariable and multivariable logistic regression.

Results. Radiographic MCP joint OA was present in 138 of the 425 patients (32.5%) but was not severe. Patients with MCP joint OA had a mean age of 69.2 ± 6.9 years, a body mass index of 25 ± 4.2 kg/m^2, and 86.2% were women. MCP joint OA was more frequent in the dominant hand and predominated at the first and second MCP joints. In the multivariable analysis, MCP joint OA was associated with older age (odds ratio [OR] 1.05 [95% confidence interval (95% CI) 1.01, 1.10] for each year), manual occupation (OR 3.74 [95% CI 1.21, 11.54]), scaphotrapezial OA (OR 2.18 [95% CI 1.27, 3.72]), and a high number of proximal interphalangeal joints with radiographic OA. MCP joint OA was not associated with metabolic syndrome or hand OA symptoms.

Conclusion. In this cross-sectional study using a hospital-based hand OA cohort, radiographic MCP joint OA was frequent and associated with structural hand OA features rather than with symptom severity. Our results suggest that the involvement of MCP joints in hand OA is predominantly related to mechanical rather than systemic factors in this population.

INTRODUCTION

Hand osteoarthritis (OA) affects a large part of the population. The prevalence of radiographic hand OA is estimated to be between 27% and 50% in the general population (1,2). Moreover, symptomatic hand OA affects approximately 13% of men and 26% of women age >70 years (3). The burden of hand OA is important for individual-level pain and disability as

DIGICOD is sponsored by Assistance Publique-Hôpitaux de Paris and by an unrestricted grant from TRB Chemedica.

[1]Inès Kouki, MD, Francis Berenbaum, MD, PhD, Jérémie Sellam, MD, PhD, Alice Courties, MD, PhD: Sorbonne University, Saint-Antoine Hospital, Assistance Publique–Hôpitaux de Paris, and CRSA Inserm UMRS_938, Paris, France; [2]Sophie Tuffet, MsC, Alexandra Rousseau, PhD: Sorbonne University, Saint-Antoine Hospital, and Assistance Publique–Hôpitaux de Paris, Paris, France; [3]Michel D. Crema, MD: Institut d'Imagerie du Sport, Institut National du Sport, de l'Expertise et de la Performance, Paris, France; [4]Pascal Richette, MD, PhD: University of Paris, Lariboisière Hospital, and Assistance Publique–Hôpitaux de Paris, Paris, France; [5]Maxime Dougados, MD, PhD: University of Paris, Cochin Hospital, Assistance Publique–Hôpitaux de Paris, and INSERM PRES Sorbonne Paris-Cité, Paris, France.

Drs. Sellam and Courties contributed equally to this article.

Dr. Crema is shareholder of Boston Imaging Core Lab, LLC. Pascal Richette has received consulting fees and/or honoraria from Pierre Fabre and Labhra, (less than $10,000 each). Dr. Berenbaum has received consulting fees and/or honoraria from TRB Chemedica, 4P Pharma, AbbVie, Boehringer, Bone Therapeutics, Eli Lilly, Expanscience, Flexion, Galapagos, Gilead, GlaxoSmithKline, Janssen, Medivir, Merck, Merck-Serono, Nordic Pharma, Novartis, Peptinov, Pfizer, Regeneron, Regulaxis, Roche, Sandoz, Sanofi, Servier, TRB Chemedica, and UCB (less than $10,000 each). Dr. Sellam has received consulting fees and/or honoraria from MSD, Pfizer, AbbVie, Fresenius Kabi, Bristol Myers Squibb, Roche Chugai, Sandoz, Eli Lilly, Gilead, Novartis, and Janssen (less than $10,000 each) and research funding from Schwa Medico, MSD, and Roche. Dr. Courties has received consulting fees and/or honoraria from Meda Pharma and UCB (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Francis Berenbaum, MD, PhD, Hôpital Saint-Antoine, Service de Rhumatologie, 184 Rue du Faubourg Saint-Antoine, 75012 Paris, France. Email: francis.berenbaum@aphp.fr

Submitted for publication December 22, 2020; accepted in revised form April 29, 2021.

SIGNIFICANCE & INNOVATIONS

- Radiographic metacarpophalangeal (MCP) joint osteoarthritis (OA) was found in one-third of the present cases but was not clinically and radiographically severe.
- Radiographic MCP joint OA is associated with radiographic severity of proximal interphalangeal OA and with the presence of scaphotrapezial OA.
- Involvement of MCP joints in hand OA is associated more with mechanical factors than with systemic factors in this cohort of patients with symptomatic hand OA.
- The presence of MCP joint OA without proximal interphalangeal OA involvement should lead us to reconsider the diagnosis of hand OA.

well as public health (4). Among hand OA risk factors, some are well established (older age, female sex, family history), while others such as obesity or metabolic syndrome remain controversial (5,6).

The commonly used clinical definition of hand OA proposed by the American College of Rheumatology has a good radiographic correlation (94% sensitivity and 87% specificity, with 78% agreement) demonstrated in a secondary care population. The definition takes into account the distal interphalangeal (DIP) and proximal interphalangeal (PIP) joint involvement but excludes metacarpophalangeal (MCP) joint swellings (7). The consensus among rheumatologists is that MCP joints are not affected by OA but by inflammatory rheumatic diseases, such as rheumatoid arthritis, psoriatic arthritis, or crystal arthropathies.

However, previous evidence suggests that MCP joints can also be affected in OA. In the Framingham cohort (n = 2,301), radiographic MCP joint OA, defined by 1 or more MCP joints with a Kellgren/Lawrence (K/L) score of ≥2, was observed in 11.9% of men and 6.8% of women (2). In the Rotterdam study (n = 3,906), 8.2% of people age >55 years and 14.8% of people age >85 years were affected (8). In contrast, symptomatic MCP joint OA is less common and affects approximately 2% of people age >60 years (9).

Very few studies have addressed the prevalence of MCP joint OA in individuals with OA and its determinants. The role of mechanical stress may be critical because high manual strain at work and high grip strength are associated with radiographic and symptomatic MCP joint OA (i.e., farmers, dockers, pneumatic drillers, or jackhammer operators) (10–12). However, its association with systemic factors has never been studied. Finally, MCP joint OA might be associated with a more severe hand OA impairment, since it has been associated with hand disability (2,8,13) and accelerated OA (14). The aim of this study was to determine the prevalence, distribution, and features associated with radiographic MCP joint OA in a large hospital-based hand OA cohort.

PATIENTS AND METHODS

Study population. The Digital Cohort Osteoarthritis Design (DIGICOD, NCT01831570) is a French monocentric prospective cohort, including patients with symptomatic hand OA. A detailed description of the study design and data collection has been given elsewhere (15). For inclusion, people age >35 years with symptomatic hand OA defined as symptomatic hand OA (pain or nodes) on at least 2 joints among PIP or DIP joints or first IP joint with K/L score of ≥2, and/or symptomatic OA (pain or deformation) at the thumb base with K/L score of ≥2, were eligible to participate.

Patients with known polyarticular chondrocalcinosis or other inflammatory rheumatic diseases, such as psoriatic arthritis or rheumatoid arthritis, were excluded. At baseline, patients were clinically examined by a rheumatologist, completed self-reported questionnaires (the Australian Canadian Osteoarthritis Hand Index [AUSCAN], Functional Index for Hand Osteoarthritis [FIHOA], and the Cochin Hand Function Scale), provided a blood sample, and underwent a radiograph of both hands. As previously described, this radiograph was performed in a posteroanterior view, with both hands in the same film (18 × 24 cassette), with the same device, and with the requirement that wrists and MCP, PIP, and DIP joints had to be well visualized. Both hands were pronated and hands and wrists had to be flat, the fingers well-extended, very little apart, and the second metacarpal had to be in the extension of the radius. The focal-film distance was 1 meter and technical parameters were 40 to 65 kV, 5–10 mAs (15).

Each hand joint of both hands (DIP, PIP, MCP, trapeziometacarpal [TMC], and scaphotrapezial) was radiographically scored by a single trained musculoskeletal radiologist (MDC) blinded to the clinical data to obtain the K/L score (0–4 for each hand joint) and Verbruggen anatomical scores (16,17). The protocol for this cross-sectional analysis of the baseline visit assessments was approved by the local ethics board (Comité de Protection des Personnes Paris Ile de France IV). All participants provided written informed consent.

Clinical and biologic data collection. At baseline, the following demographic and biologic variables were collected: age, sex, occupation, body mass index (BMI), alcohol consumption, smoking, diabetes mellitus, hypertension, dyslipidemia, cardiovascular diseases, metabolic syndrome as defined by the National Cholesterol Education Program–Adult Treatment Panel III criteria (18), menopausal status, family history of OA, biologic inflammation (C-reactive protein [CRP] level ≥5 mg/liter), hyperferritinemia (defined by ferritinemia >300 μg/liter in men or >200 μg/liter in women), time of evolution of hand OA, and use of analgesics or nonsteroidal antiinflammatory drugs. All the definitions used for these characteristics are shown in Supplementary Table 1, available on the *Arthritis Care &*

Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24642. Current or previous occupations were classified using the Institut National de la Statistique et des Études Économiques classification and grouped into 3 categories: manual professions; intermediate professions and employees; and intellectual professions. In this study, we call "manual workers" farmers, craftsmen, traders, and workers.

Radiographic MCP joint OA and clinical severity. An MCP joint with OA was defined by a K/L score of ≥2, and radiographic MCP joint OA was defined by at least 2 MCP joints with a K/L score of ≥2 among the MCP joints of both hands. Radiographic severity excluded the MCP joints, since they contributed to the outcome, and was defined as the sum of the K/L score at the PIP, DIP, scaphotrapezial, and TMC joints (total score between 0 and 88); the number of PIP, DIP, scaphotrapezial, and TMC joints with K/L score of ≥2, each and together; and the number of erosive joints among the PIP and DIP joints defined by phase E or R of the Verbruggen radiographic score.

Clinical severity was defined by an AUSCAN pain score ≥40 of 100, FIHOA score >5 of 30, Cochin Hand Function Scale score between 0 and 90, the visual analog scale (VAS) score for hand pain ≥40 of 100, the number of painful joints on applying pressure, swelling joints among the PIP, DIP, and MCP joints and thumb base, and grip and pinch strengths on the dominant hand.

Statistical analysis. The prevalence of radiographic MCP joint OA defined by at least 2 MCP joints with a K/L score of ≥2 was evaluated as well as the prevalence of patients with at least 2 MCP joints with a K/L score of ≥3 or 4. The prevalence of painful MCP joints spontaneously or under pressure was assessed as well. Proportions of patients with MCP joint OA were compared between the dominant and nondominant hand using a McNemar test for paired data.

Baseline characteristics were reported according to the presence or absence of radiographic MCP joint OA using frequencies and percentages for categorical variables and mean ± SD or median and interquartile range (IQR) for continuous variables (according to the distribution). Differences in proportions between groups were reported with 95% confidence intervals (95% CIs), estimated by the exact method, and the difference of means between groups was reported with a 95% CI estimated by the pooled or Satterthwaite method, depending on whether variances were equal or unequal between groups, respectively.

Associations between the presence of radiographic MCP joint OA and demographic characteristics and radiographic and clinical severity were analyzed using logistic regression. A multivariable model was built using covariates selected from the univariable analysis (P value less than 0.20) and was systematically adjusted for age, sex, BMI, and family history of OA. The final model was determined using a backward elimination method. Results were expressed as odds ratios (ORs) with 95% CIs. All analyses were performed using SAS software, version 9.4. All tests were 2-sided, and a P value less than 0.05 indicated statistical significance.

RESULTS

In total, 436 patients were included in the DIGICOD cohort between April 2013 and June 2017 from the rheumatology department of Saint-Antoine Hospital (Assistance Publique-Hôpitaux de Paris). Eleven participants were excluded (1 withdrew consent, 3 had unavailable radiographs, 6 did not meet the selection criteria, 1 with missing K/L score). Finally, data from 425 patients were analyzed in the present study.

MCP joint OA is frequent and predominates on the dominant hand but is radiographically and clinically not severe. Among the 425 patients, 49.4% had at least 1 MCP joint K/L score of ≥2, 16.7% had at least 1 MCP joint K/L score of ≥3, and only 3.3% had K/L score of ≥4. The median MCP joint K/L total score (extremes 0–40) was 3.0 (IQR 1.0–6.0). MCP joint OA was more frequent in the dominant hand (40.6% versus 34.7%; $P < 0.05$) and was predominant in the thumb and index finger (20.2% and 20.7%, respectively) (Figure 1). Spontaneous painful MCP joint was rare (n = 29 of 424, 6.8%), while 139 of 424 patients (32.8%) had at least 1 painful MCP joint under pressure. Only 2 patients (0.5%) had 1 erosive MCP joint and none had more. In total, 138 patients (32.5%) had at least 2 MCP joints with a K/L score of ≥2 and were considered as patients with radiographic MCP joint OA for subsequent analyses.

Patients with MCP joint OA are older, have a more manual job, and are radiographically more severe. Patients with or without radiographic MCP joint OA were similar in their BMI (mean ± SD 25.0 ± 4.2 versus 25.2 ± 4.4) and OA family history (66.9% versus 70.5%). In addition, the percentages of patients with metabolic syndrome or its components were similar between the 2 groups. Patients with MCP joint OA were slightly older than those without MCP joint OA (mean ± SD 69.2 ± 6.9 versus 65.5 ± 7.3 years). Although there was a predominance of women in both groups, the proportion of women appeared to be slightly more frequent among individuals with MCP joint OA (86.2% versus 82.6%). There were twice as many patients with past or present work involving hands among the patients with MCP joint OA (10.4% versus 4.6%) compared to patients without MCP joint OA. In contrast, intellectual professions were less frequent

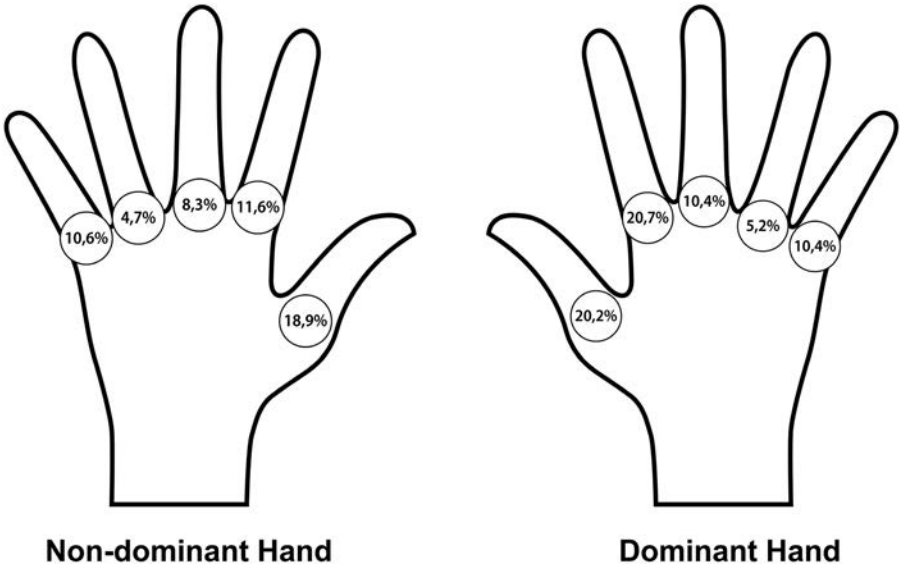

Figure 1. Hand diagram showing the percentages of patients of the cohort who have metacarpophalangeal joint osteoarthritis (Kellgren/Lawrence score ≥2) according to the fingers and the dominant hand.

among patients with MCP joint OA. There were no differences between the 2 groups in terms of iron parameters, CRP level, and calcemia (Table 1).

In the MCP joint OA group, the mean ± SD K/L score excluding MCP joints (0–88) was 50.3 ± 11.9 compared to 37.9 ± 15.3 in the group without MCP joint OA. The presence of

Table 1. Characteristics of patients according to the presence or absence of MCP joint OA*

Characteristic	No MCP joint OA (n = 287)		MCP joint OA (n = 138)		Difference (95% CI)
	No.†	Value	No.†	Value	
Age at inclusion, mean ± SD years	287	65.5 ± 7.3	138	69.2 ± 6.9	3.7 (2.2, 5.1)
Women	287	237 (82.6)	138	119 (86.2)	3.7 (–3.6, 10.9)
Body mass index, mean ± SD kg/m^2	285	25.2 ± 4.4	133	25.0 ± 4.2	–0.1 (–1.0, 0.8)
Family history of osteoarthritis	278	196 (70.5)	133	89 (66.9)	–3.6 (–13.2, 6.0)
Socioprofessional category	282		135		
Manual works	–	13 (4.6)	–	14 (10.4)	5.8 (0.06, 11.5)
Intermediate professions and employees	–	107 (37.9)	–	63 (46.7)	8.7 (–1.4, 18.9)
Intellectual professions	–	162 (57.4)	–	58 (43.0)	–14.5 (–24.6, –4.3)
Time of evolution of hand osteoarthritis, years	285		136		
<5	–	70 (24.6)	–	14 (10.3)	–14.3 (–21.4, –7.1)
5–15	–	143 (50.2)	–	68 (50.0)	–0.2 (–10.4, 10.0)
>15	–	72 (25.3)	–	54 (39.7)	14.4 (4.8, 24.1)
Current tobacco consumption	285	21 (7.4)	134	8 (6.0)	–1.4 (–6.4, 3.6)
Current alcohol consumption	285	215 (75.4)	134	109 (81.3)	5.9 (–2.4, 14.2)
Diabetes mellitus	287	24 (8.4)	138	9 (6.5)	–1.8 (–7.1, 3.4)
Hypertension	287	155 (54.0)	138	82 (59.4)	5.4 (–4.6, 15.4)
Cardiovascular disease except AHT	287	13 (4.5)	138	8 (5.8)	1.3 (–3.3, 5.8)
LDL, mean ± SD gm/liter	277	1.4 ± 0.4	134	1.4 ± 0.4	0.009 (–0.07, 0.09)
Metabolic syndrome	279	99 (35.5)	134	51 (38.1)	2.6 (–7.4, 12.5)
CRP, mg/liter	237	–	112	–	–
<5	–	213 (89.9)	–	101 (90.2)	0.3 (–6.4, 7.0)
≥5	–	24 (10.1)	–	11 (9.8)	–0.3 (–7.0, 6.4)
Hyperferritinemia	253	28 (11.1)	121	12 (9.9)	–1.1 (–7.7, 5.4)
TSC >45%	277	8 (2.9)	134	5 (3.7)	0.8 (–2.9, 4.6)
Calcemia, mean ± SD mmol/liter	279	2.4 ± 0.1	135	2.4 ± 0.1	0.003 (–0.02, 0.02)
Current analgesics or NSAIDs consumption	287	149 (51.9)	138	80 (58.0)	6.1 (–4.0, 16.1)

* Values are the number (%) unless indicated otherwise. 95% CI = 95% confidence interval; AHT = arterial hypertension; CRP = C-reactive protein; LDL = low-density lipoprotein; MCP = metacarpophalangeal; NSAIDs = nonsteroidal antiinflammatory drugs; OA = osteoarthritis; TSC = transferrin saturation coefficient.
† Number of patients with available data.

a high number of PIP joints with OA (9 or 10) was much more frequent in patients with radiographic MCP joint OA, while the presence of a low number of PIP joints with OA (0–4) was much more frequent in patients without MCP joint OA. In addition, 56.2% of the patients with MCP joint OA had erosive OA versus 39.2% of patients without MCP joint OA.

Concerning the clinical severity, 59.1% of patients with MCP joint OA had a VAS score of ≥40 of 100 compared to 57.2% of patients without MCP joint OA. There was no evident difference in the function scores on the FIHOA and Cochin scales (Table 2).

Age, socioprofessional category, and number of PIP and scaphotrapezial joints with OA are independently associated with MCP joint OA. The duration of hand OA, number of erosive joints, DIP or TMC joints with OA, painful joints on applying pressure, number of joints with synovitis, and prehension strength on the dominant hand were significantly associated with MCP joint OA in the univariable analysis ($P < 0.20$), but the association did not remain in multivariable analysis. No biologic factors were associated with MCP joint OA (calcemia, CRP level, hyperferritinemia, or high TSC). Neither BMI nor metabolic syndrome was associated with MCP joint OA (Table 3).

In the multivariable analysis, radiographic MCP joint OA was associated with higher age (OR 1.05 [95% CI 1.01, 1.10]), manual works (OR 3.74 [95% CI 1.21, 11.54] for manual workers versus intellectual professionals), and higher radiographic severity of hand OA (Table 3). Patients with 5–8 PIP joints with OA, representing 34.1% of the entire population, were 7.83 (95% CI 3.58, 17.16) times more at risk for MCP joint OA than patients with 0–4 PIP joints with OA. In addition, those with 9–10 PIP joints with OA (30% of the population) were 14.29 (95% CI 6.46, 31.64) times more at risk for MCP joint OA compared to patients with

Table 2. Clinical and radiographic severity of patients according to the presence or absence of MCP joint OA*

	No MCP joint OA (n = 287)		MCP joint OA (n = 138)		
	No.†	Value	No.†	Value	Difference (95% CI)
Radiographic severity					
Sum of K/L score except MCP joint (0–88), mean ± SD	276	37.9 ± 15.3	135	50.3 ± 11.9	12.4 (9.7, 15.1)
Number of joints K/L ≥2 excluding MCP joint (0–22)	276	13.0 (8.0–17.0)	135	18.0 (14.0–20.0)	–
Erosive joints: PIP and DIP (0–18), no. (%)	281		137		–
0	–	171 (60.9)	–	60 (43.8)	–
1–3	–	82 (29.2)	–	47 (34.3)	–
≥4	–	28 (10.0)	–	30 (21.9)	–
Number of DIP joint K/L ≥2 (0–8), no. (%)	282		138		–
0–3	–	48 (17.0)	–	4 (2.9)	–
4–6	–	89 (31.6)	–	25 (18.1)	–
7–8	–	145 (51.4)	–	109 (79.0)	–
Number of PIP joint K/L ≥2 (0–10), no. (%)	285		137		–
0–4	–	138 (48.4)	–	13 (9.5)	–
5–8	–	92 (32.3)	–	52 (38.0)	–
9–10	–	55 (19.3)	–	72 (52.6)	–
Number of ST K/L ≥2 (0–2), no. (%)	284		137		–
0	–	172 (60.6)	–	52 (38.0)	–
1	–	49 (17.3)	–	32 (23.4)	–
2	–	63 (22.2)	–	53 (38.7)	–
Number of TMC K/L ≥2 (0–2), no. (%)	284		135		–
0	–	93 (32.7)	–	36 (26.7)	–
1	–	56 (19.7)	–	17 (12.6)	–
2	–	135 (47.5)	–	82 (60.7)	–
Sum of K/L score at TMC + ST (0–16)	283	5.0 (2.0 –9.0)	135	8.0 (4.0 –10.0)	–
Clinical severity					
AUSCAN pain score ≥40/100, no. (%)	271	69 (25.5)	124	26 (21.0)	4.5 (–4.4, 13.3)
FIHOA score >5/30, no. (%)	272	129 (47.4)	132	73 (55.3)	–7.9 (–18.2, 2.5)
Cochin Function Hand Scale (0–90)	274	5.0 (1.0–13.0)	132	6.0 (2.0–15.0)	
Pain VAS (rest or activity) ≥40/100, no. (%)	285	163 (57.2)	137	81 (59.1)	–1.9 (–12.0, 8.1)
Painful joints on pressure (0–30)	287	3.0 (2.0–6.0)	137	3.0 (1.0–7.0)	–
Swelling joints (0–30)	287	0.0 (0.0–1.0)	136	1.0 (0.0–3.0)	–
Grip strength on dominant hand, kg	287	24.0 (20.0–31.0)	136	22.0 (18.0–28.0)	–
Pinch strength on dominant hand, kg	280	5.4 (4.3–6.7)	130	5.3 (4.4–6.5)	–

* Values are the median (interquartile range) unless indicated otherwise. 95% CI = 95% confidence interval; AUSCAN = Australian Canadian Osteoarthritis Hand Index; DIP = distal interphalangeal; FIHOA = Functional Index for Hand osteoarthritis; K/L = Kellgren/Lawrence; MCP = metacarpophalangeal; OA = osteoarthritis; PIP = proximal interphalangeal; ST = scaphotrapezial; TMC = trapeziometacarpal; VAS = visual analog scale.
† Number of patients with available data.

Table 3. Univariable and multivariable analysis of factors associated with MCP joint OA (n = 356, population with available values for all selected variables in univariable analysis)*

	Univariable analysis		Multivariable analysis†	
	OR (95% CI)	*P*	OR (95% CI)	*P*
Age at inclusion by year	1.09 (1.05, 1.13)	<0.0001	1.05 (1.01, 1.10)	0.0194
Sex	–	0.2947	–	0.2505
Men	1	–	1	–
Women	1.40 (0.74, 2.65)	–	1.59 (0.72, 3.48)	–
Body mass index, kg/m^2	1.00 (0.95, 1.05)	0.9912	0.99 (0.93, 1.05)	0.7547
Family history of osteoarthritis	–	0.7893	–	0.8558
No	1	–	1	–
Yes	0.94 (0.58, 1.51)	–	1.05 (0.59, 1.87)	–
Socioprofessional category	–	0.0114	–	0.0158
Intellectual professions	1	–	1	–
Manual works	3.16 (1.31, 7.62)	0.0104	3.74 (1.21, 11.54)	0.0218
Intermediate professions and employees	1.67 (1.05, 2.68)	0.0320	1.85 (1.07, 3.21)	0.0280
Time of evolution of hand osteoarthritis, years	–	<0.0001	–	–
<5	1	–	–	–
5–15	4.64 (1.32, 16.39)	0.0170	–	–
>15	9.67 (2.91, 32.14)	0.0002	–	–
Number of erosive joints among PIP and DIP (0–18)	–	0.0027	–	–
0	1	–	–	–
1–3	1.58 (0.94, 2.64)	0.0815	–	–
≥4	2.94 (1.57, 5.52)	0.0008	–	–
Number of DIP K/L ≥2 (0–8)	–	<0.0001	–	–
0–3	1	–	–	–
4–6	3.37 (1.11, 10.25)	0.0323	–	–
7–8	9.02 (3.16, 25.77)	<0.0001	–	–
Number of PIP K/L ≥2 (0–8)	–	<0.0001	–	<0.0001
0–4	1	–	1	–
5–8	6.78 (3.23, 14.23)	<0.0001	7.83 (3.58, 17.16)	<0.0001
9–10	15.34 (7.22, 32.59)	<0.0001	14.29 (6.46, 31.64)	<0.0001
Number of ST K/L ≥2 (0–2)	–	<0.0001	–	0.0045
0	1	–	1	–
1–2	2.60 (1.64, 4.12)	–	2.18 (1.27, 3.72)	–
Number of TMC K/L ≥2 (0–2)	–	0.0430	–	–
0–1	1	–	–	–
2	1.59 (1.01, 2.51)	–	–	–
Painful joints at pressure (0–30)	–	0.0232	–	–
0–1	1	–	–	–
2–3	0.45 (0.24, 0.82)	0.0097	–	–
4–6	0.48 (0.25, 0.93)	0.0289	–	–
≥7	0.86 (0.47, 1.59)	0.6319	–	–
Swelling joints (0–30)	–	0.0013	–	–
0–2	1	–	–	–
≥3	2.49 (1.43, 4.35)	–	–	–
Grip strength on dominant hand, kg	–	0.0143	–	–
>30	1	–	–	–
≤30	2.12 (1.16, 3.87)	–	–	–

* 95% CI = 95% confidence interval; DIP = distal interphalangeal; K/L = Kellgren/Lawrence; MCP = metacarpophalangeal; OA = osteoarthritis; OR = odds ratio; PIP = proximal interphalangeal; ST = scaphotrapezial; TMC = trapeziometacarpal.
† Systematic adjustment for age, sex, body mass index, and family history of OA.

0–4 PIP joints with OA. Lastly, scaphotrapezial impairment was also associated with radiographic MCP joint OA, with an OR of 2.18 (95% CI 1.27, 3.72).

DISCUSSION

Using a large hospital-based cohort of patients with symptomatic interphalangeal and/or thumb base hand OA, we found that radiographic MCP joint OA is frequent (32.5%) and predominates in the thumb and the index finger of the dominant hand. Radiographic MCP joint OA was independently associated with older age, manual work, PIP joint OA radiographic severity, and the presence of scaphotrapezial OA, but not with systemic or metabolic factors.

Here, we found that the prevalence of radiographic MCP joint OA is higher than that reported in previous studies, such as the Framingham study (11.9% of men and 6.8% of women in

Framingham) (2). One explanation is that we studied MCP joint OA in a symptomatic hand OA population (defined by PIP/DIP or thumb joint OA) (i.e., hospital-based cohort), while the Framingham cohort is a general population sample (i.e., a population-based cohort). Indeed, in a recent study, MCP joint impairment was studied using magnetic resonance imaging in hand OA patients compared to controls showing that 21 of 81 MCP joints had a loss of cartilage, including 5 with areas of full-thickness loss (19). This prevalence is close to that observed in our study. MCP joint OA is more frequent than expected and probably underestimated, especially when it is studied with radiographs only. Although the prevalence of MCP joint OA was high, the severity of the radiographic impairment was low (low prevalence of MCP joints at K/L score of 3 or 4 and low median MCP joint K/L sum) and so was the clinical burden (few spontaneously painful MCP joints).

We found that older age was associated with MCP joint OA (5% increase in risk per year) but not with sex. Conversely, in the Framingham cohort and the Zoetermeer study, MCP joint OA was slightly more frequent in men and at a younger age than other hand OA localizations. For example, in the age group 40–44 years, 9.6% of men had MCP joint OA versus 8.2% of women (2,20). These discrepancies can also be explained by the differences in the sample populations (individuals with hand OA versus general population). The presence of confounding factors in the Framingham study from a random general population would likely include more manual workers among men, explaining this difference. Radiographic MCP joint OA did not seem to be independently associated with disease duration, as may have been expected, probably because disease duration is strongly associated with age.

In this population of symptomatic hand OA patients, MCP joint impairment was more frequent and independently associated with manual jobs, which is consistent with the literature (10–12). Indeed, some observations among samples of up to 200 patients found an association between manual works (farmers, cotton workers, workers using vibrating-tools) and MCP joint OA in terms of frequency and of severity (10,21,22). Moreover, the distribution of MCP joint OA was consistent with previous publications, since MCP joint OA was more frequently described in the thumb, index, and middle fingers, probably because mechanical factors are stronger at these fingers (2,8). The higher frequency of MCP joint OA in the dominant hand and the association with manual occupations suggest that mechanical factors are strongly implicated in the development of MCP joint OA, which is in accordance with studies showing that higher grip strengths can be associated with the occurrence of MCP joint OA later (11,12). However, more specific mechanical factors such as quantitative data about the manual occupation (mechanical loading or repetitive tasks) and length of time in the employment or manual hobbies were not available, which limits the accuracy of manual activities assessment.

In contrast, systemic factors such as BMI, diabetes mellitus, metabolic syndrome, and elevated CRP level were not associated with MCP joint OA among these hand OA patients. The absence of such associations reinforces the importance of mechanical factors in MCP joint OA and suggests that MCP joints are more responsive to mechanical stress than PIP or DIP joints (5,23). There was no association with hyperferritinemia, elevated TSC, or hypercalcemia, but we expected this finding because the study population was primary OA patients. Some other systemic features may have influenced MCP joint OA such as genetic factors, but they were not investigated here.

Finally, MCP joint OA was associated with hand OA radiographic severity, especially with PIP joint OA severity and with the presence of scaphotrapezial OA. Although the scaphotrapezial joint is known to be related to mechanical forces, it is independently associated with MCP joint OA, even when adjusted on manual works. This finding is consistent with data showing that hand OA of other joints co-occurred in 86% of patients with MCP joint involvement (8). Thus, OA in the MCP joints could be part of a generalized and evolved form of hand OA. Since only 2.7% and 9.5% of patients with MCP joint OA have <3 DIP joints with OA and <4 PIP joints with OA, MCP joint OA without PIP or DIP joint OA should lead us to reconsider the diagnosis of primary hand OA.

While MCP joint OA was associated with symptoms in the univariable analysis, there was no association between MCP joint OA and hand OA clinical symptoms (pain, loss of function, or loss of grip strength) in the multivariable analysis. The reason may be because symptoms are closely related to radiographic severity. This result means that although radiographic MCP joint OA is common, its clinical impact is limited and is mediated by the severity of the PIP or thumb base joint OA.

The strengths of this study should be considered. First, DIGICOD is a large cohort of individuals with symptomatic hand OA, defined by worldwide validated criteria. The radiographs were analyzed by a trained radiologist blinded to the patient characteristics. Second, the definition of MCP joint OA was stringent compared to other studies because we used the presence of ≥2 MCP joints with K/L score of ≥2 (and not ≥1 MCP joint with K/L score ≥2) to be more specific in the association studied. Finally, patient characteristics were defined using composite criteria (patient interview, clinical examination, and biologic samples), thereby strengthening the results.

However, our study had some limitations. First, it was cross-sectional, which limits the possibility to demonstrate a causality link between MCP joint OA and manual activity. Second, it was conducted in a symptomatic hand OA population, which limits the generalizability of the findings to asymptomatic hand OA patients. Generalizability is particularly true for the high prevalence, which is not what is observed in the general population. Conducting sensitivity analysis using more stringent criteria would have been pertinent, but doing so was technically impossible because of the low number of patients with MCP joints with K/L score of 3 and 4. Moreover, because this is a symptomatic hand OA cohort, it was predominantly a female population, which prevented us from describing MCP joint OA among men. Finally, because of missing data, the analyzed population was smaller than the included population, with a slight overrepresentation of

individuals with diabetes mellitus and hypertension, as well as those with metabolic syndrome. However, since these factors were not associated with MCP joint OA, their influence over the results would be unlikely.

In conclusion, radiographic MCP joint OA is frequent in patients with symptomatic PIP/DIP or thumb base joint OA but is clinically rarely severe. Among these patients, the presence of radiographic MCP joint OA is cross-sectionally associated with mechanical factors rather than systemic factors such as metabolic factors and is especially associated with severe radiographic PIP joint OA. The presence of MCP joint OA without PIP joint OA involvement should lead us to reconsider the diagnosis of hand OA, which is consistent with the literature. Further prospective studies are necessary to confirm longitudinally these cross-sectional results.

ACKNOWLEDGMENTS

The authors thank all the patients from the DIGICOD cohort who participated, all the investigators who managed the patients' recruitment, and the monitoring of clinical visits (Dr. Camille Deprouw, Dr. Sandra Desouches, Dr. Ariane Do, Dr. Emeline Gaigneux, Dr. Camille Glanowski, Dr. Karine Louati, Dr. Stéphanie Malbos, and Dr. Sabine Trellu, all from the rheumatology department of AP-HP Saint-Antoine Hospital), the staff members of the Clinical Research Unit of Eastern Paris, and the Centre des Ressources Biologiques, AP-HP Saint-Antoine Hospital.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Berenbaum 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. Kouki, Richette, Dougados, Berenbaum, Sellam, Courties.

Acquisition of data. Kouki, Tuffet, Crema, Rousseau, Berenbaum, Sellam, Courties.

Analysis and interpretation of data. Kouki, Tuffet, Rousseau, Berenbaum, Sellam, Courties.

REFERENCES

1. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part II. Arthritis Rheum 2008;58:26–35.
2. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis 2011;70:1581–6.
3. Zhang Y, Niu J, Kelly-Hayes M, Chaisson CE, Aliabadi P, Felson DT. Prevalence of symptomatic hand osteoarthritis and its impact on functional status among the elderly: the Framingham Study. Am J Epidemiol 2002;156:1021–7.
4. Michon M, Maheu E, Berenbaum F. Assessing health-related quality of life in hand osteoarthritis: a literature review. Ann Rheum Dis 2011; 70:921–8.
5. Yusuf E, Nelissen RG, Ioan-Facsinay A, Stojanovic-Susulic V, DeGroot J, van Osch G, et al. Association between weight or body mass index and hand osteoarthritis: a systematic review. Ann Rheum Dis 2010;69:761–5.
6. Haugen IK, Magnusson K, Turkiewicz A, Englund M. The prevalence, incidence, and progression of hand osteoarthritis in relation to body mass index, smoking, and alcohol consumption. J Rheumatol 2017; 44:1402–9.
7. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. Arthritis Rheum 1990; 33:1601–10.
8. Dahaghin S, Bierma-Zeinstra SM, Ginai AZ, Pols HA, Hazes JM, Koes BW. Prevalence and pattern of radiographic hand osteoarthritis and association with pain and disability (the Rotterdam study). Ann Rheum Dis 2005;64:682–7.
9. Zhang Y, Xu L, Nevitt MC, Niu J, Goggins JP, Aliabadi P, et al. Lower prevalence of hand osteoarthritis among Chinese subjects in Beijing compared with white subjects in the United States: the Beijing Osteoarthritis Study. Arthritis Rheum 2003;48:1034–40.
10. Williams WV, Cope R, Gaunt WD, Adelstein EH, Hoyt TS, Singh A, et al. Metacarpophalangeal arthropathy associated with manual labor (Missouri metacarpal syndrome): clinical, radiographic, and pathologic characteristics of an unusual degeneration process. Arthritis Rheum 1987;30:1362–71.
11. Chaisson CE, Zhang Y, Sharma L, Kannel W, Felson DT. Grip strength and the risk of developing radiographic hand osteoarthritis: results from the Framingham Study. Arthritis Rheum 1999;42:33–8.
12. Chaisson CE, Zhang Y, Sharma L, Felson DT. Higher grip strength increases the risk of incident radiographic osteoarthritis in proximal hand joints. Osteoarthritis Cartilage 2000;8 Suppl A:S29–32.
13. Schaefer LF, McAlindon TE, Eaton CB, Roberts MB, Haugen IK, Smith SE, et al. The associations between radiographic hand osteoarthritis definitions and hand pain: data from the osteoarthritis initiative. Rheumatol Int 2018;38:403–13.
14. Davis JE, Schaefer LF, McAlindon TE, Eaton CB, Roberts MB, Haugen IK, et al. Characteristics of accelerated hand osteoarthritis: data from the Osteoarthritis Initiative. J Rheumatol 2019;46:422–8.
15. Sellam J, Maheu E, Crema MD, Touati A, Courties A, Tuffet S, et al. The DIGICOD cohort: a hospital-based observational prospective cohort of patients with hand osteoarthritis. Methodology and baseline characteristics of the population. Joint Bone Spine 2021;88:105171.
16. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
17. Verbruggen G, Veys EM. Numerical scoring systems for the anatomic evolution of osteoarthritis of the finger joints. Arthritis Rheum 1996;39: 308–20.
18. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III). JAMA 2001;285:2486–97.
19. Saltzherr MS, Muradin GS, Haugen IK, Selles RW, van Neck JW, Coert JH, et al. Cartilage evaluation in finger joints in healthy controls and early hand osteoarthritis patients using high-resolution MRI. Osteoarthritis Cartilage 2019;27:1148–51.
20. Van Saase JL, van Romunde LK, Cats A, Vandenbroucke JP, Valkenburg HA. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. Ann Rheum Dis 1989;48:271–80.
21. Lawrence JS. Rheumatism in cotton operatives. Br J Ind Med 1961; 18:270–6.
22. Gemne G, Saraste H. Bone and joint pathology in workers using hand-held vibrating tools: an overview. Scand J Work Environ Health 1987;13:290–300.
23. Puenpatom RA, Victor TW. Increased prevalence of metabolic syndrome in individuals with osteoarthritis: an analysis of NHANES III data. Postgrad Med 2009;121:9–20.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1704–1712
DOI 10.1002/acr.24613

American College of Rheumatology
Empowering Rheumatology Professionals

Socioeconomic Inequalities in All-Cause and Cause-Specific Mortality Among Patients With Osteoarthritis in the Skåne Region of Sweden

Maria Lindéus,[1] Aleksandra Turkiewicz,[1] Martin Englund,[1] and Ali Kiadaliri[2]

Objective. To assess the association between education and all-cause and cause-specific mortality among patients with osteoarthritis (OA) in comparison to an OA-free reference cohort.

Methods. Using data from the Skåne Healthcare Register, we identified all residents age ≥45 years in the region of Skåne in southern Sweden with doctor-diagnosed OA of peripheral joints between 1998 and 2013 (n = 123,993). We created an age- and sex-matched reference cohort without OA diagnosis (n = 121,318). Subjects were followed until death, relocation outside Skåne, or the end of 2014. The relative index of inequality (RII) and the slope index of inequality (SII) were estimated by the Cox model and Aalen's additive hazard model, respectively.

Results. We found an inverse association between education and mortality. The magnitude of relative inequalities in all-cause mortality were comparable in the OA, with an RII of 1.53 (95% confidence interval [95% CI] 1.46, 1.61), and reference cohorts (RII 1.54 [95% CI 1.47, 1.62]). The absolute inequalities were smaller in the OA cohort (all-cause deaths per 100,000 person-years, SII 937 [95% CI 811, 1,063]) compared with the reference cohort (SII 1,265 [95% CI 1,109, 1,421]). Cardiovascular mortality contributed more to the absolute inequalities in the OA cohort than in the reference cohort (60.1% versus 48.1%) while the opposite was observed for cancer mortality (8.5% versus 22.3%).

Conclusion. We found higher all-cause and cause-specific mortality in OA patients with lower education. The observed inequalities in the OA cohort reflect the inequalities in the population at large. The greater burden of cardiovascular diseases in OA patients suggests that proper management of cardiovascular risk factors in OA patients is important.

INTRODUCTION

Osteoarthritis (OA) commonly affects the knee and hip joints, is a major cause of disability (1,2), and the burden of OA in our aging population will likely keep rising (3). An increased all-cause mortality in OA patients compared to the general population has been reported by some studies (4–9), while others have found no such association (10–13). Certain studies have also examined cause-specific mortality and reported an association between OA and increased cardiovascular mortality (5,7,8,14). Physical inactivity and the use of nonsteroidal antiinflammatory drugs (NSAIDs) have been suggested to be the most plausible explanations for the observed excess all-cause mortality in OA patients (4–8,14).

There is a well-known association between socioeconomic status (SES) and health, wherein people with lower SES generally tend to have poorer health and increased mortality (15). In the context of OA, previous studies have shown an inverse association between both individual and area-level SES and OA prevalence (16–23). However, no previous study has, to our knowledge, examined the association between SES and all-cause and cause-specific mortality among patients with OA. In order to promote health equity, it is important to analyze and highlight health disparities. Further, it is important to study all-cause mortality as well as the burden of disease-specific mortality in order to improve public health and reduce preventable deaths. Considering that no widely accepted disease-modifying drug is on the market for OA treatment, understanding societal factors

Supported by grants from The Swedish Research Council, The Swedish Rheumatism Association, Governmental Funding of Clinical Research within National Health Service (ALF), Region Skåne, Greta and Kocks Foundation, and Crafoord Foundation. Funders had no role in study design, data collection and analysis, or preparation of the manuscript.

[1]Maria Lindéus, MD, Aleksandra Turkiewicz, PhD, Martin Englund, MD: Lund University, Faculty of Medicine, Lund, Sweden; [2]Ali Kiadaliri, PhD: Centre for Economic Demography, Lund University, Lund, Sweden.

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

Address correspondence to Maria Lindéus, MD, Skåne University Hospital, Clinical Epidemiology Unit, Remissgatan 4, SE-221 85 Lund, Sweden. Email: maria.lindeus@med.lu.se.

Submitted for publication April 9, 2020; accepted in revised form March 30, 2021.

SIGNIFICANCE & INNOVATIONS

- No previous study has assessed the magnitude of educational inequalities in a population with osteoarthritis (OA) compared to a reference population.
- People with lower education, with or without OA, have higher all-cause and cause-specific mortality compared to more highly educated people.
- The educational inequalities in mortality in OA patients reflect the health inequalities in the population at large.
- OA patients with lower education have a greater burden of cardiovascular diseases.

associated with adverse consequences of the disease is important. Education is an important factor to consider in OA management since several studies have reported an inverse association between educational attainment and prevalence/symptoms of OA (18,19,21–26). Thus, the aim of the present study was to assess the association between education and all-cause and cause-specific mortality among OA patients, and to compare these with a sex- and age-matched cohort free of OA.

MATERIALS AND METHODS

We conducted a register-based, open-cohort study. The region of Skåne in southern Sweden was the geographical area of interest. In 2013, the region had a population of ~1.3 million (~13% of Sweden's entire population) (http://www.scb.se). Ethical endorsement of the project was applied for and was approved by the Regional Ethics Review Board in Lund (Dnr 2014/276).

Data sources. The Skåne Healthcare Register (SHR) is a database that registers all health care consultations in the Skåne region. From the SHR, we identified subjects with doctor-diagnosed OA of peripheral joints (International Classification of Diseases, Tenth Revision [ICD-10] codes M15-M19). We retrieved data on vital status, sex, and place of residence from Statistics Sweden's Population Register. Data regarding each subject's highest educational attainment, country of birth, and marital status were collected from the Longitudinal integration database for health insurance and labor market studies (LISA). The LISA, which is maintained by Statistics Sweden, integrates data from educational and social sectors as well as from the labor market (https://www.scb.se/lisa-en). The National Board of Health and Welfare's Cause of Death Register was used to retrieve the underlying cause of death and the date of death reported on death certificates of the subjects who died during the observation period.

Outcome and follow-up period. The recruitment period was from January 1, 1998 to December 31, 2013. For inclusion in the OA cohort, a subject had to meet the following criteria: have at least 1 OA diagnosis (ICD-10 codes M15–M19) in the SHR, be age ≥45 years, and be a resident of the Skåne region at the time of OA diagnosis. To create the reference cohort, patients with no OA diagnosis (no OA diagnosis registered in the SHR from 1998 to 2013) were randomly sampled and frequency matched for age and sex to the OA cohort. Since 3 of the subjects in the reference cohort had unknown dates of death, they were each assigned a date of death at random. Each subject's follow-up started at the date of OA diagnosis (index date) or January 1 of the year they turned age 45 years, whichever occurred last. The subjects of the 2 cohorts were followed until relocation outside the Skåne region, death, or December 31, 2014, whichever occurred first.

Covariates. The level of education was categorized into the following 3 groups: low (0–9 years of education), medium (10–12 years of education), and high (>12 years of education). We considered age, sex, marital status, and nativity as potential confounders. Marital status was classified into 3 categories: never married, married, and previously married. Individuals who were in a registered partnership or had been in a registered partnership were also included in the 2 latter groups. Nativity was divided into 2 categories: native (i.e., born in Sweden with at least 1 Sweden-born parent) and non-native (i.e., born abroad, or born in Sweden with both parents born abroad).

Causes of death. The 5 following causes of death were used to examine cause-specific mortality: 1) diseases of the circulatory system (ICD-10 codes I00–I99), 2) malignant neoplasms (C00–C97), 3) mental and behavioral disorders (F00–F99) and diseases of the nervous system (G00–G99), 4) diseases of the respiratory system (J00–J99), and 5) other causes. These groups were chosen because they were the leading causes of death in the study population.

Statistical analysis. To examine educational inequalities in mortality, the relative index of inequality (RII) and slope index of inequality (SII) were estimated. The RII measures the relative inequality between the hypothetical best-placed and worst-placed person, ranked by educational attainment. Correspondingly, the SII measures the absolute inequalities between the 2 extremes of educational attainment. In order to estimate these indices, the populations in each educational group were assigned a fractional rank. The fractional rank was based on the midpoint of the educational group's range in the cumulative distribution of the study population (27–29). We used the Cox model to estimate the RII, and Aalen's additive hazard model to estimate the SII. Age was used as the time scale for both models. We used R codes from Moreno-Betancur et al (28) for these estimations. We also conducted subgroup analyses by sex, attained age throughout follow-up, and collected data on

prevalence of knee and hip OA, the 2 joint sites most commonly affected by OA. All subjects with registered knee OA were considered to have knee OA in the knee OA–specific analyses, regardless of whether they also had hip OA or any other type of OA. The same consideration was applied for those who had hip OA. All estimates were adjusted for sex, marital status, and nativity. Analyses were performed in R, version 3.5.1.

RESULTS

After excluding 4,195 people with unknown education status, 5 people with unknown nativity status, and 4 people with missing underlying cause of death, a total of 123,993 people were included in the OA cohort (60.9% women). A total of 123,993 subjects free of OA were initially included in the reference cohort. The final reference cohort consisted of 121,318 people (60.7% women) after 2,671 with unknown education status and 4 people with unknown nativity status were excluded. The mean ± SD age at study entry was 66.6 ± 11.6 years for the OA cohort and 66.5 ± 11.6 years for the reference cohort. In total, 978 subjects in the OA cohort had both knee and hip OA. The mean follow-up time was 7.1 years for the OA cohort and 6.6 years for the reference cohort. During follow-up, 24,493 and 28,843 deaths were identified in the OA and reference cohorts, respectively (Table 1).

The relative educational inequality, measured by the RII, showed that the all-cause mortality rate was ~1.5 times higher in the lowest-educated versus the highest-educated in both the OA (Table 2) and reference cohorts (Table 3), with an RII of 1.53 (95% confidence interval [95% CI] 1.46, 1.61) in the OA cohort and an RII of 1.54 (95% CI 1.47, 1.62) in the reference cohort. When examining cause-specific mortality, higher mortality was observed in patients with lower education in both cohorts for all specific causes (even though the CI included the null value of 1 for the cause of mental and behavioral disorders and diseases of the nervous system). Educational inequalities in favor of the more highly educated were generally seen in all subgroups in both cohorts. The magnitudes of relative inequalities were, in general, comparable between the OA and reference cohorts.

The absolute educational inequalities, measured by the SII, showed that there were 937 (95% CI 811, 1,063) more all-cause deaths per 100,000 person-years in the lowest versus highest educated OA patients (Table 2). The corresponding figure was greater in the reference cohort (1,265 [95% CI 1,109, 1,421] more deaths per 100,000 person-years in favor of the highest-educated) (Table 3). Greater absolute educational inequalities in all-cause mortality were also seen among women, and those age 45–74 years, in the reference cohort compared to the OA cohort. There were also greater absolute educational inequalities in the reference cohort for subjects with malignant

Table 1. Characteristics of the osteoarthritis (OA) and reference cohorts*

	Level of education, OA cohort				Level of education, reference cohort			
	Low	Medium	High	Missing	Low	Medium	High	Missing
Number of participants	48,157	49,946	25,890	4,195	47,108	46,298	27,912	2,671
Women	61.3	60.0	61.8	68.2	62.8	60.0	59.1	66.6
Age at entry, years								
45–64	31.5	53.1	56.5	13.8	29.4	52.6	62.4	26.4
65–84	57.8	42.4	40.5	45.5	59.8	43.0	35.1	61.4
85+	10.7	4.5	3.0	40.8	10.8	4.4	2.5	12.2
Marital status at entry								
Never married	7.9	9.2	10.0	5.5	10.4	11.5	12.7	10.9
Previously married	39.2	32.7	27.8	33.6	40.4	32.6	25.7	40.3
Married	53.0	58.1	62.2	26.8	49.2	55.9	61.5	47.7
Missing	0.0	0.0	0.0	34.1	0.0	0.0	0.0	1.1
Swedish nativity	88.1	85.9	85.4	30.2	87.8	85.3	84.7	24.1
Missing	0.0	0.0	0.0	34.1	0.0	0.0	0.0	1.2
OA type, number of participants								
Knee OA (ICD-10 code M17)	20,842	20,378	10,132	1,857	–	–	–	–
Hip OA (ICD-10 code M16)	10,892	9,567	5,076	999	–	–	–	–
Hand OA (ICD-10 codes M18, M15.1, M15.2)	2,961	4,638	2,719	147	–	–	–	–
Number of total deaths	14,154	7,687	2,652	2,748	16,795	8,855	3,193	874
Diseases of the circulatory system (I00–I99)	6,302	3,013	920	1,271	6,917	3,240	1,061	339
Malignant neoplasms (C00–C97)	3,046	2,176	845	432	3,998	2,608	1,091	165
Mental and behavioral disorders (F00–F99) and diseases of the nervous system (G00–G99)	1,053	572	217	178	1,493	796	296	78
Diseases of the respiratory system (J00–J99)	906	449	172	185	1,357	649	202	66
Other causes	2,847	1,477	498	682	3,030	1,562	543	226
Person-years of follow-up	352,367	351,369	173,710	16,203	315,601	306,450	180,713	15,505

* Except where indicated otherwise, values are the percentage of study participants. Low: 0–9 years of education, medium: 10–12 years of education, and high: >12 years of education. ICD-10 = International Classification of Diseases, Tenth Revision.

Table 2. Relative index of inequality (RII) and slope index of inequality (SII) in all-cause and cause-specific mortality in the osteoarthritis (OA) group*

Disease group (ICD-10 codes)	All (>45 years)		Men		Women		45–74 years		≥75 years		Knee OA (M17)		Hip OA (M16)	
	RII	SII†	RII	SII†	RII	SII†	RII	SII†	RII	SII†	RII	SII†	RII	SII†
All-cause	1.53 (1.46, 1.61)	937 (811, 1063)	1.54 (1.43, 1.66)	1,281 (1,059, 1,504)	1.52 (1.42, 1.62)	959 (799, 1,119)	1.86 (1.67, 2.07)	502 (409, 595)	1.44 (1.36, 1.52)	521 (426, 617)	1.51 (1.40, 1.63)	882 (684, 1,081)	1.51 (1.38, 1.65)	1,250 (941, 1,559)
Diseases of the circulatory system (I00-I99)	1.86 (1.71, 2.01)	563 (485, 640)	1.84 (1.63, 2.07)	739 (606, 873)	1.80 (1.62, 2.01)	547 (448, 647)	2.57 (2.08, 3.17)	210 (164, 255)	1.71 (1.57, 1.86)	351 (288, 414)	1.76 (1.56, 1.99)	523 (401, 645)	1.87 (1.62, 2.16)	806 (611, 1,000)
Malignant neoplasms (C00-C97)	1.17 (1.06, 1.29)	80 (14, 146)	1.27 (1.10, 1.46)	191 (81, 302)	1.16 (1.01, 1.33)	88 (8, 167)	1.46 (1.25, 1.72)	137 (77, 198)	1.06 (0.94, 1.20)	3 (−40, 46)	1.17 (1.01, 1.37)	74 (−25, 174)	1.17 (0.98, 1.40)	109 (−59, 276)
Mental and behavioral disorders (F00-F99) and diseases of the nervous system (G00-G99)	1.06 (0.88, 1.27)	6 (−31, 42)	0.95 (0.69, 1.31)	−7 (−59, 46)	1.08 (0.86, 1.36)	8 (−40, 56)	1.18 (0.65, 2.12)	5 (−12, 21)	1.02 (0.85, 1.24)	2 (−25, 29)	1.08 (0.80, 1.44)	3 (−51, 56)	0.89 (0.64, 1.23)	−36 (−122, 50)
Diseases of the respiratory system (J00-J99)	1.54 (1.25, 1.90)	55 (22, 88)	1.67 (1.23, 2.29)	99 (41, 156)	1.50 (1.13, 2.00)	56 (16, 95)	2.65 (1.60, 4.37)	38 (19, 57)	1.44 (1.15, 1.80)	32 (8, 55)	1.58 (1.13, 2.22)	50 (4, 96)	1.43 (0.97, 2.10)	59 (−22, 141)
Other causes	1.70 (1.51, 1.91)	234 (177, 291)	1.58 (1.33, 1.87)	259 (161, 357)	1.79 (1.53, 2.10)	261 (187, 335)	2.01 (1.57, 2.57)	112 (72, 152)	1.59 (1.40, 1.81)	133 (93, 173)	1.69 (1.42, 2.01)	233 (144, 321)	1.68 (1.35, 2.09)	312 (171, 454)

* Models were adjusted for sex, marital status, age, and nativity. 95% confidence intervals are shown in parentheses. ICD-10 = International Classification of Diseases, Tenth Revision.
† Per 100,000 person-years; SII values are rounded to integers.

Table 3. Relative index of inequality (RII) and slope index of inequality (SII) in all-cause and cause-specific mortality in the reference cohort, by sex and attained age*

Disease group (ICD-10 codes)	All ≥45 years		Men		Women		45–74 years		≥75 years	
	RII	SII†	RII	SII†	RII	SII†	RII	SII†	RII	SII†
All-cause	1.54 (1.47, 1.62)	1,265 (1,109, 1,421)	1.50 (1.40, 1.61)	1,498 (1,243, 1,752)	1.58 (1.48, 1.69)	1,355 (1,164, 1,546)	2.31 (2.15, 2.49)	954 (832, 1,077)	1.34 (1.27, 1.41)	543 (433, 653)
Diseases of the circulatory system (I00-I99)	1.76 (1.63, 1.90)	609 (516, 702)	1.66 (1.48, 1.86)	732 (575, 889)	1.79 (1.60, 1.99)	627 (512, 742)	2.68 (2.24, 3.20)	313 (254, 371)	1.53 (1.41, 1.66)	341 (268, 413)
Malignant neoplasms (C00-C97)	1.39 (1.27, 1.52)	282 (199, 366)	1.41 (1.24, 1.61)	360 (227, 494)	1.47 (1.30, 1.66)	325 (221, 429)	1.89 (1.65, 2.16)	354 (275, 433)	1.13 (1.01, 1.27)	44 (−6, 93)
Mental and behavioral disorders (F00-F99) and diseases of the nervous system (G00-G99)	1.02 (0.87, 1.20)	4 (−42, 50)	0.99 (0.77, 1.28)	−2 (−70, 67)	1.03 (0.84, 1.27)	6 (−56, 68)	1.51 (1.01, 2.26)	23 (0, 46)	0.96 (0.81, 1.13)	−10 (−45, 26)
Diseases of the respiratory system (J00-J99)	1.90 (1.59, 2.27)	142 (100, 184)	1.93 (1.48, 2.51)	180 (110, 250)	1.96 (1.54, 2.49)	147 (93, 200)	3.41 (2.29, 5.06)	86 (59, 113)	1.61 (1.33, 1.95)	74 (43, 105)
Other causes	1.58 (1.41, 1.77)	228 (163, 293)	1.44 (1.21, 1.70)	227 (129, 325)	1.64 (1.40, 1.92)	251 (169, 333)	2.43 (1.95, 3.02)	179 (135, 224)	1.33 (1.17, 1.52)	94 (46, 141)

* Models were adjusted for sex, marital status, age, and nativity. 95% confidence intervals are shown in parentheses. ICD-10 = International Classification of Diseases, Tenth Revision.
† Per 100,000 person-years; SII values are rounded to integers.

neoplasm (subjects age ≥45 years, women, and subjects age 45–74 years) and for subjects with diseases of the respiratory system (subjects age ≥45 years and subjects age 45–74 years).

Subgroup analyses based on OA type showed that the relative and absolute educational inequalities in mortality were similar in subjects with knee and hip OA (Table 2). In general, diseases of the circulatory system contributed the most, and mental and behavioral disorders and diseases of the nervous system contributed the least, to absolute educational inequalities in all-cause mortality in both the OA (Figure 1) and reference cohorts (Figure 2). The only exception was in subjects age 45–74 years in the reference cohort, where the cause of malignant neoplasm contributed the most (37.1%) to the absolute educational inequality. Diseases of the circulatory system tended to have a higher contribution to the absolute educational inequality in mortality in the OA cohort compared to the reference cohort, while deaths caused by malignant neoplasms tended to contribute more to the absolute educational inequality in mortality in the reference cohort. While the contribution of cardiovascular mortality rose with age, the opposite was seen for cancer-related mortality.

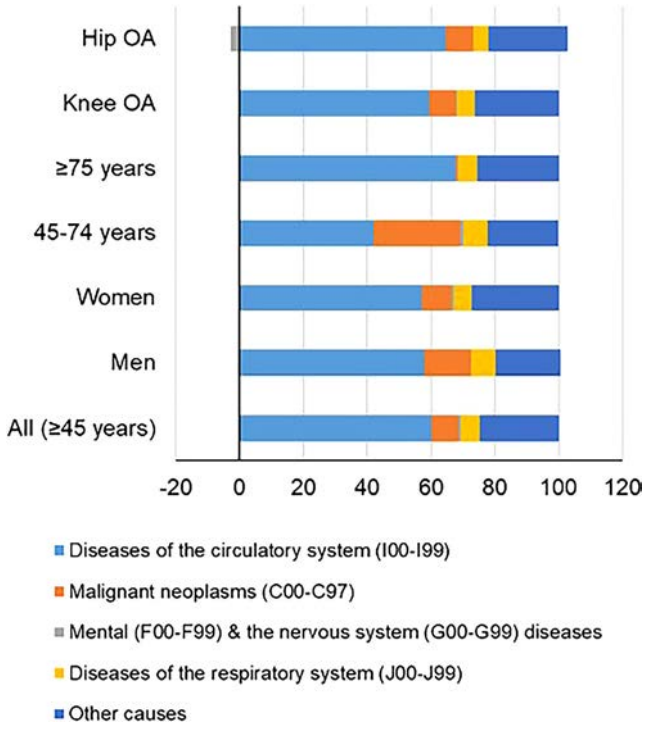

Figure 1. Contribution (%) of specific causes of death to the absolute educational inequality in all-cause mortality in the osteoarthritis (OA) cohort by sex, attained age, and OA type.

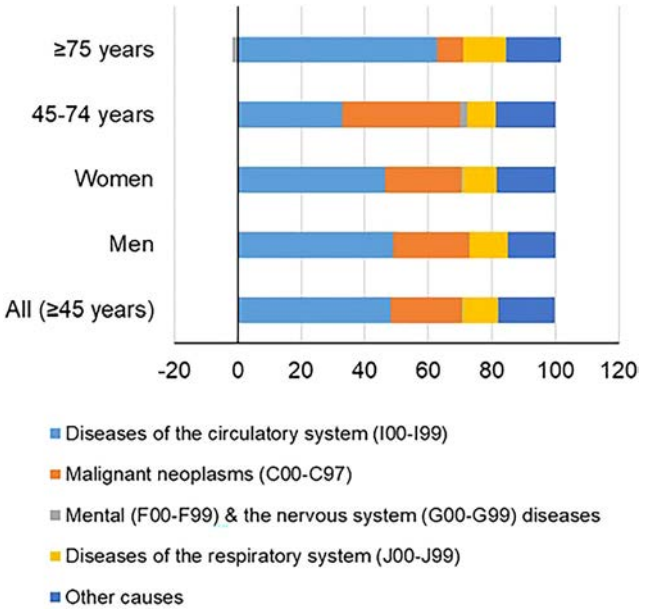

Figure 2. Contribution (%) of specific causes of death to the absolute educational inequality in all-cause mortality in the reference cohort by sex and attained age.

DISCUSSION

In this large register-based study, we found an inverse association between educational attainment and all-cause and cause-specific mortality among OA patients and people in the reference cohort. We found higher contribution of cardiovascular mortality to the absolute educational inequalities in mortality in the OA cohort compared to the reference cohort, which reflects the greater burden of cardiovascular diseases in the OA cohort, especially in persons with low education. While the magnitudes of the relative educational inequalities in mortality were comparable in people with and without OA, the absolute educational inequalities in mortality were greater in the reference cohort, which may be due to greater overall mortality in the reference cohort (23.8%) compared to the OA cohort (19.8%). Indeed, lower mortality in OA patients compared to the general population has been shown in previous studies based on populations in Sweden, the Netherlands, and the US (11–13). However, increased mortality in OA patients has also been suggested by other studies from Japan, China, the UK, and the US (4–9). A possible explanation of smaller absolute educational inequality in mortality in OA patients may be that OA patients who seek care for their OA may also get treatment for other underlying health conditions, leading to overall better health (11,12). However, patients with OA treated within health care facilities may not be entirely representative of all persons with OA (30). More broadly, we observe that the educational inequalities in mortality in the OA cohort also existed in the reference cohort, which suggests that the educational inequalities seen in the OA cohort reflect those in the population at large. Thus, these inequalities should mainly be addressed in the general population in order to even out the observed disparities between people with different educational backgrounds, both with and without OA.

From a global perspective, public health in Sweden is good. In 2016, the average life expectancy at birth in Sweden was 81 years in men and 84 years in women, compared to the average life expectancy of the global population of 72 years in the same year (https://www.who.int/). However, along with an increasing remaining average life expectancy at the age of 30 years for all educational groups in Sweden between 2006 to 2017, there has also been a trend of increasing disparity in the remaining average life expectancy between high- and low-educated people (31). In 2017, the remaining average life expectancy at age 30 years was ~6 years shorter in people with pre-secondary versus post-secondary education (31). The disparities in life expectancy are reflected in the present study, where higher mortality rates were observed among the less educated.

Differences in material, psychosocial, and behavioral factors between socioeconomic groups are commonly used to explain the creation of health inequality (32). It is more likely that higher education leads to better economic conditions than low education. A better economic situation enables better resources that influence health outcomes (32). In the psychosocial explanatory model, health inequality arises from differences among the social groups in psychological distress (the feeling of stress, discrimination, low social support, etc.). This psychological distress in turn has a negative effect on physical health (32). Another commonly used explanatory model, the behavioral model, explains health inequality as a product of behavioral differences (smoking prevalence, eating habits, cancer screening, etc.) among the groups (32). Lastly, differences in background health might also cause differences in academic performance, i.e., there is a reverse causality between education and health, which could worsen educational inequality in health. Studies have shown that there is an association between health in childhood and academic performance. For instance, prematurity/low birth weight is associated with lower educational qualifications in adulthood (33). In the present study, it is probable that both health benefits from education as well as differences in background health contribute to the observed educational inequalities in mortality.

The contribution of cardiovascular mortality to absolute educational inequalities in mortality was greater in the OA cohort in comparison to the reference cohort, reflecting the greater burden of cardiovascular diseases in the OA cohort. This result is in line with previous studies that have reported an association between OA and the metabolic syndrome (34), as well as an association between increased cardiovascular mortality and OA (5,7,8,14). Although the observed educational inequalities among OA patients reflect the inequalities in the population at large, the finding of greater burden of cardiovascular diseases together with inequalities in cardiovascular mortality in OA patients provides valuable information for clinicians. By highlighting the observed inequality, it is possible to take preventative measures in this defined group of patients. Reduced risk of cardiovascular mortality within this vulnerable group could potentially be achieved within, for example, standardized OA management programs such as Better management of patients with OsteoArthritis (BOA) and Good Life with osteoArthritis in Denmark (GLA:D) through personalized information, screening, and initiation of treatment of cardiovascular risk factors.

Interestingly, cancer-related mortality contributed more to the absolute educational inequality in mortality in the reference cohort in comparison to the OA cohort. There were minor differences between the cohorts in the distribution of cancer types causing death. The 5 leading causes of cancer-related deaths in the reference cohort were malignancies of bronchus and lung (19.2%), colon (9.1%), prostate (8.6%), pancreas (7.0%), and breast (6.7%), while the leading causes in the OA cohort were malignancies of bronchus and lung (15.9%), prostate (10.9%), colon (9.0%), pancreas (8.2%), and breast (7.5%). This minor difference in distributions might contribute to the higher cancer-related inequality in mortality in the reference cohort since there are variations in the strengths of association between different cancer types and SES (35). Another possible explanation is that the estimates in the reference cohort reflect the ongoing transition in the predominant cause of death from cardiovascular death to cancer-related death, which has been reported in high-income countries, including Sweden (36). This transition has been suggested to be caused by a reduction in cardiovascular mortality due to better prevention and treatment of cardiovascular diseases in high-income countries (36). Results from the Prospective Urban Rural Epidemiology (PURE) study showed this transition in the predominant cause of death in middle-age (36), which is consistent with our results wherein cancer-related mortality contributed more to the absolute educational inequality in mortality in patients age <75 years compared to older patients. If the shift in predominant cause of death explains the differences seen in cancer-related mortality between the 2 cohorts, this further emphasizes the importance of treatment of cardiovascular diseases in OA patients. A slightly higher respiratory-related mortality was also found for some subgroups in the reference cohort. However, it is difficult to speculate what may have caused this difference.

In general, we found smaller educational inequalities for deaths due to mental and behavioral disorders and diseases of the nervous system. Among deaths in mental and behavioral disorders and diseases of the nervous system, different forms of dementia contributed the most to mortality in both the OA cohort (77.6%) and the reference cohort (75.2%). Studies examining the association between SES and deaths due to dementia are sparse (37,38), but an inverse association between educational attainment and dementia mortality has been reported (37,38). The lack of inequality in our study may be due to the fact that mental and behavioral disorders and diseases of the nervous system contains heterogeneous diseases, which in turn may have different associations with SES.

We acknowledge several limitations of the present study. The mortality data used is based on information retrieved from death certificates, which in turn are coded by the responsible physicians. Although the Swedish cause of death register is a high-quality register with high completeness (39), problems with inaccurate mortality data in the register have been reported (40). Furthermore, if there are educational differences in reporting causes of death, our results may be biased. As with data from the cause of death register, data from the SHR are also prone to bias, and if these biases are associated with education (e.g., differences in OA diagnosis based on the patient's level of education due to health care avoidance or cognitive bias in physicians), then our estimates of educational inequality may be biased. Another limitation of the present study, which also applies to mortality studies in general, is the use of the single-cause of death method. By only taking the underlying cause of death into consideration, useful information is most likely lost since deaths are often caused by more than 1 disease, particularly among older people. Finally, our results may be generalizable to similar populations living with similar health care systems as the one in Sweden. In Sweden, the health care is mainly government-funded, thus the system facilitates relatively good access to health care services for all inhabitants.

The strengths of our study are the large study sample (virtually an entire geographically well-defined population) and the application of the most recent advances in methodology of estimating SII. We used the SII and RII to estimate educational inequalities and both indices accounted for the population size within the different socioeconomic groups, which is beneficial since this may change over time (27–29).

In conclusion, we found an inverse association between educational attainment and all-cause and cause-specific mortality among OA patients. The educational inequalities seen in the OA cohort were also seen in the reference cohort, indicating that the inequalities in the OA cohort reflect inequalities in the population at large. Hence, these inequalities should primarily be addressed in the general population in order to even out the observed disparities among both people with and without OA. However, our estimates also indicate a greater burden of cardiovascular diseases in OA patients compared to the reference cohort. The greater burden of cardiovascular diseases applies particularly to OA patients with low education. This finding generates an opportunity to reduce educational inequalities in cardiovascular mortality in OA patients by focusing on prevention and treatment of cardiovascular risk factors in OA patients.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Lindéus 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. Lindéus, Turkiewicz, Kiadaliri.
Acquisition of data. Turkiewicz,
Analysis and interpretation of data. Lindéus, Turkiewicz, Englund, Kiadaliri.

REFERENCES

1. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017;390:1211–59.
2. Kiadaliri AA, Lohmander LS, Moradi-Lakeh M, Petersson IF, Englund M. High and rising burden of hip and knee osteoarthritis in the Nordic region, 1990-2015. Acta Ortho 2018;89:177–83.
3. Turkiewicz A, Petersson IF, Bjork J, Hawker G, Dahlberg LE, Lohmander LS, et al. Current and future impact of osteoarthritis on health care: a population-based study with projections to year 2032. Osteoarthritis Cartilage 2014;22:1826–32.
4. Tsuboi M, Hasegawa Y, Matsuyama Y, Suzuki S, Suzuki K, Imagama S. Do musculoskeletal degenerative diseases affect mortality and cause of death after 10 years in Japan? J Bone Min Metab 2011;29:217–23.
5. Nüesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Jüni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. BMJ 2011;342:d1165.
6. 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.
7. Kluzek S, Sanchez-Santos MT, Leyland KM, Judge A, Spector TD, 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.
8. Barbour KE, Lui LY, Nevitt MC, Murphy LB, Helmick CG, Theis KA, et al. Hip osteoarthritis and the risk of all-cause and disease-specific mortality in older women: a population-based cohort study. Arthritis Care Res (Hoboken) 2015;67:1798–805.
9. Cleveland RJ, Alvarez C, Schwartz TA, Losina E, Renner JB, Jordan JM, et al. The impact of painful knee osteoarthritis on mortality: a community-based cohort study with over 24 years of follow-up. Osteoarthritis Cartilage 2019;27:593–602.
10. Xing D, Xu Y, Liu Q, Ke Y, Wang B, Li Z, et al. Osteoarthritis and all-cause mortality in worldwide populations: grading the evidence from a meta-analysis. Sci Rep 2016;6:24393.
11. Turkiewicz A, Neogi T, Bjork J, Peat G, Englund M. All-cause mortality in knee and hip osteoarthritis and rheumatoid arthritis. Epidemiology 2016;27:479–85.
12. Liu R, Kwok WY, Vliet Vlieland TP, Kroon HM, Meulenbelt I, Houwing-Duistermaat JJ, et al. Mortality in osteoarthritis patients. Scand J Rheumatol 2015;44:70–3.
13. Holbrook TL, Wingard DL, Barrett-Connor E. Self-reported arthritis among men and women in an adult community. J Commun Health 1990;15:195–208.
14. Turkiewicz A, Kiadaliri AA, Englund M. Cause-specific mortality in osteoarthritis of peripheral joints. Osteoarthritis Cartilage 2019;27:848–54.
15. Marmot M, Friel S, Bell R, Houweling TA, Taylor S. Closing the gap in a generation: health equity through action on the social determinants of health. Lancet 2008;372:1661–9.
16. Rossignol M, Leclerc A, Hilliquin P, Allaert FA, Rozenberg S, Valat JP, et al. Primary osteoarthritis and occupations: a national cross sectional survey of 10 412 symptomatic patients. Occup Environ Med 2003;60:882–6.

17. Rossignol M. Primary osteoarthritis and occupation in the Quebec national health and social survey. Occup Environ Med 2004;61: 729–35.
18. Callahan LF, Cleveland RJ, Shreffler J, Schwartz TA, Schoster B, Randolph R, et al. Associations of educational attainment, occupation and community poverty with knee osteoarthritis in the Johnston County (North Carolina) osteoarthritis project. Arthritis Res Ther 2011;13:R169.
19. Tang X, Wang S, Zhan S, Niu J, Tao K, Zhang Y, et al. The prevalence of symptomatic knee osteoarthritis in China: results rrom the China Health and Retirement Longitudinal Study. Arthritis Rheumatol 2016; 68:648–53.
20. Reyes C, Garcia-Gil M, Elorza JM, Mendez-Boo L, Hermosilla E, Javaid MK, et al. Socio-economic status and the risk of developing hand, hip or knee osteoarthritis: a region-wide ecological study. Osteoarthritis Cartilage 2015;23:1323–9.
21. Hannan MT, Anderson JJ, Pincus T, Felson DT. Educational attainment and osteoarthritis: differential associations with radiographic changes and symptom reporting. J Clin Epidemiol 1992;45:139–47.
22. Cleveland RJ, Schwartz TA, Prizer LP, Randolph R, Schoster B, Renner JB, et al. Associations of educational attainment, occupation and community poverty with hip osteoarthritis. Arthritis Care Res (Hoboken) 2013;65:954–61.
23. Kiadaliri AA, Gerhardsson de Verdier M, Turkiewicz A, Lohmander LS, Englund M. Socioeconomic inequalities in knee pain, knee osteoarthritis, and health-related quality of life: a population-based cohort study in southern Sweden. Scand J Rheumatol 2017;46:143–51.
24. Feldman CH, Dong Y, Katz JN, Donnell-Fink LA, Losina E. Association between socioeconomic status and pain, function and pain catastrophizing at presentation for total knee arthroplasty. BMC Musculoskelet Disord 2015;16:18.
25. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Prevalence and burden of osteoarthritis: results from a population survey in Norway. J Rheumatol 2008;35:677–84.
26. Juhakoski R, Tenhonen S, Anttonen T, Kauppinen T, Arokoski JP. Factors affecting self-reported pain and physical function in patients with hip osteoarthritis. Arch Phys Med Rehabil 2008;89: 1066–73.
27. Mackenbach JP, Kunst AE. Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe. Soc Sci Med 1997;44: 757–71.
28. Moreno-Betancur M, Latouche A, Menvielle G, Kunst AE, Rey G. Relative index of inequality and slope index of inequality: a structured regression framework for estimation. Epidemiology 2015;26: 518–27.
29. Oakes J, Kaufman J. Methods in social epidemiology. Hoboken, NJ: John Wiley & Sons, Inc; 2006.
30. Turkiewicz A, Nilsson PM, Kiadaliri A. Probabilistic quantification of bias to combine the strengths of population-based register data and clinical cohorts: studying mortality in osteoarthritis. Am J Epidemiol 2020;189:1590–9.
31. Public Health Agency of Sweden (Folkhälsomyndigheten). Folkhälsans utveckling. Årsrapport 2019. Yearly report on public health 2019 (In Swedish). 2019. URL: https://www.folkhalsomyndigheten.se/contentassets/d162673edec94e5f8d1da1f78e54dac4/folkhalsans-utveckling-arsrapport-2019.pdf.
32. Arcaya MC, Arcaya AL, Subramanian SV. Inequalities in health: definitions, concepts, and theories. Glob Health Action 2015;8:27106.
33. Bilgin A, Mendonca M, Wolke D. Preterm birth/low birth weight and markers reflective of wealth in adulthood: a meta-analysis. Pediatrics 2018;142:e20173625.
34. Dickson BM, Roelofs AJ, Rochford JJ, Wilson HM, De Bari C. The burden of metabolic syndrome on osteoarthritic joints. Arthritis Res Ther 2019;21:289.
35. Clegg LX, Reichman ME, Miller BA, Hankey BF, Singh GK, Lin YD, et al. Impact of socioeconomic status on cancer incidence and stage at diagnosis: selected findings from the surveillance, epidemiology, and end results. The National Longitudinal Mortality Study. Cancer Causes Control 2009;20:417–35.
36. Dagenais GR, Leong DP, Rangarajan S, Lanas F, Lopez-Jaramillo P, Gupta R, et al. Variations in common diseases, hospital admissions, and deaths in middle-aged adults in 21 countries from five continents (PURE): a prospective cohort study. Lancet 2020;395:785–94.
37. Strand BH, Skirbekk V, Rosness TA, Engedal K, Bjertness E. Income in midlife and dementia related mortality over three decades: a Norwegian prospective study. eNeurologicalSci 2015;1:24–9.
38. Russ TC, Stamatakis E, Hamer M, Starr JM, Kivimaki M, Batty GD. Socioeconomic status as a risk factor for dementia death: individual participant meta-analysis of 86,508 men and women from the UK. Brit J Psychiatr J Ment Sci 2013;203:10–7.
39. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. Euro J Epidemiol 2017;32:765–73.
40. Johansson LA, Bjorkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. J Clin Epidemiol 2009;62:1202–9.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1713–1722
DOI 10.1002/acr.24640

American College of Rheumatology
Empowering Rheumatology Professionals

Intermetatarsal Bursitis, a Novel Feature of Juxtaarticular Inflammation in Early Rheumatoid Arthritis Related to Clinical Signs: Results of a Longitudinal Magnetic Resonance Imaging Study

Bastiaan T. van Dijk,[1] Yousra J. Dakkak,[1] Xanthe M. E. Matthijssen,[1] Ellis Niemantsverdriet,[1] Monique Reijnierse,[1] and Annette H. M. van der Helm-van Mil[2]

Objective. Intermetatarsal bursae in the forefeet possess a synovial lining similar to joints and tendon sheaths. Inflammation of these bursae (intermetatarsal bursitis [IMB]) was recently identified as specific for early rheumatoid arthritis (RA). The present study was undertaken to determine if IMB is indeed an RA feature by assessing the following: 1) the association with other local inflammatory measures (synovitis, tenosynovitis, and osteitis), 2) the association with clinical signs, and 3) whether it responds to disease-modifying antirheumatic drug (DMARD) therapy similarly to other local inflammatory measures.

Methods. One hundred fifty-seven consecutive early RA patients underwent unilateral contrast-enhanced 1.5T forefoot magnetic resonance imaging (MRI) at diagnosis. MRIs were evaluated for IMB presence and for synovitis, tenosynovitis, and osteitis in line with the RA MRI Scoring (RAMRIS) system (summed as RAMRIS inflammation). MRIs at 4, 12, and 24 months were evaluated for IMB presence and size in patients who had IMB at baseline and received early DMARD therapy. Logistic regression and generalized estimating equations were used. Anti–citrullinated protein antibody (ACPA) stratification was performed.

Results. Sixty-nine percent of RA patients had ≥1 IMB. In multivariable analysis on bursa level, presence of IMB was independently associated with local presence of synovitis and tenosynovitis, with odds ratios (OR) of 1.69 (95% confidence interval [95% CI] 1.12, 2.57) and 2.83 (95% CI 1.80, 4.44), respectively, but not osteitis. On the patient level, IMB presence was most strongly associated with tenosynovitis (OR 2.92 [95% CI 1.62, 5.24]). IMB presence was associated with local joint swelling (OR 2.7 [95% CI 1.3, 5.3]) and tenderness (OR 1.7 [95% CI 1.04, 2.9]) independent of RAMRIS inflammation. During treatment, IMB size decreased between 0 and 12 months. This decrease associated with decrease in RAMRIS inflammation, which was driven by synovitis decrease. Within ACPA-positive and ACPA-negative RA, similar results were obtained.

Conclusion. IMB particularly accompanies inflammation of the synovial lining of joints and tendon sheaths, showed a similar treatment response after DMARD initiation, and associates with typical clinical signs. These findings suggest that IMB represents a frequently present novel RA feature of juxtaarticular synovial inflammation.

INTRODUCTION

Rheumatoid arthritis (RA)–related local inflammation in the hands and forefeet can be reliably and sensitively assessed using magnetic resonance imaging (MRI) (1,2), which is recommended by the European Alliance of Associations for Rheumatology for early detection of RA (3). Three features of local inflammation are assessed according to the conventionally used RA MRI Scoring (RAMRIS) system: synovitis, tenosynovitis, and osteitis (4). Although RA is traditionally known for targeting the synovial lining of (small) joints (synovitis), MRI studies have shown that juxtaarticular synovial inflammation in the form of tenosynovitis is typical

Supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation (starting grant agreement 634886) and by ReumaNederland.

[1]Bastiaan T. van Dijk, MSc, MD, Yousra J. Dakkak, PhD, MD, Xanthe M. E. Matthijssen, MSc, MD, Ellis Niemantsverdriet, PhD, Monique Reijnierse, PhD, MD: Leiden University Medical Centre, Leiden, The Netherlands; [2]Annette H. M. van der Helm-van Mil, PhD, MD: Erasmus Medical Center, Rotterdam, The Netherlands.

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

Address correspondence to Bastiaan T. van Dijk, MSc, MD, Department of Rheumatology, Leiden University Medical Centre, PO Box 9600, 2300 RC Leiden, The Netherlands. Email: B.T.van_Dijk@lumc.nl.

Submitted for publication February 15, 2021; accepted in revised form April 29, 2021.

SIGNIFICANCE & INNOVATIONS

- Inflammation of the synovium-lined intermetatarsal bursae (intermetatarsal bursitis [IMB]) is frequently present at diagnosis of rheumatoid arthritis (RA) (69%), both in anti–citrullinated protein antibody (ACPA)–positive (75%) and ACPA-negative (64%) patients, and associates with local joint tenderness and swelling.
- IMB also associates with known RA-related magnetic resonance imaging (MRI) inflammation (synovitis and tenosynovitis).
- After initiation of disease-modifying antirheumatic drugs, IMB decreases in a fashion similar to known RA-related MRI inflammation and disease activity (according to the Disease Activity Score), suggesting a treatment response.
- These findings imply that IMB is indeed a novel juxtaarticular inflammatory feature of RA.

for the disease as well; tenosynovitis at the small joints represents inflammation of the synovial lining of tendon sheaths, is specific for early RA, and contributes to RA-specific symptoms (5–8).

Forefoot involvement is frequent in RA and an important cause of symptoms and disability (9). Specifically in the forefeet, in addition to synovial joints and tendon sheaths, another distinct tissue with a synovial lining but without connection to the metatarsophalangeal (MTP) joints is present and may become inflamed: the intermetatarsal bursae (10–13). MRI-detected intermetatarsal bursitis (IMB) was recently identified as highly specific for early RA and to be less frequent in healthy controls and non-RA arthritides (14). Although IMB has been described in established RA and is associated with foot-related disability (15,16), its role in early disease has barely been explored. Two studies have thus far reported a prevalence of IMB in early RA of 63% and 69% (14,17). One of these studies also showed that IMB associates with RA independently from clinical factors (age, sex, and body mass index) (14). Bursae have a function in reducing mechanical strain and friction. Mechanical strain (e.g., due to deformities or altered mechanical loading) is suggested to be involved in bursitis development, but reports on its role in IMB development in early RA are contradictory (18–20). In short, there is some evidence suggesting that IMB is a feature of early RA, but scientific data are scarce.

Because RA is the most common inflammatory disease in the field of rheumatology and foot symptoms are common in individuals with RA, we believe it is essential to understand the pathophysiology of forefoot symptoms. The forefeet undergo mechanical loading during walking. In recent-onset RA, mechanical loading may possibly influence forefeet inflammation and/or aggravate forefeet symptoms. As such, IMB can be part of inflammation that relates to symptoms. We hypothesized that if IMB is indeed a feature of early RA, it should associate with known MRI inflammation measured by the RAMRIS (synovitis, tenosynovitis, and osteitis), as well as with typical signs related to RA (joint tenderness and swelling) at diagnosis. To determine this, we performed a large cross-sectional MRI study. Finally, we hypothesized that IMB should also respond to treatment with disease-modifying antirheumatic drugs (DMARDs), analogous to MRI inflammation measured by the RAMRIS (21,22). Follow-up MRIs were evaluated to study this.

PATIENTS AND METHODS

Patients. The Leiden Early Arthritis Clinic (EAC) is an inception cohort based in the Leiden University Medical Centre (LUMC) in The Netherlands and has been enrolling patients with clinically apparent arthritis of recent onset (symptom duration <2 years) since 1993. Its design has been described previously (23). At baseline, tender and swollen joint counts were conducted, Disease Activity Score (DAS) was assessed, and blood samples were taken to measure C-reactive protein level, erythrocyte sedimentation rate, IgG anti–citrullinated protein antibodies (ACPAs), and IgM rheumatoid factor (24). Follow-up visits were scheduled at 4 months, 12 months, and yearly thereafter. All patients provided written informed consent. This study was carried out in compliance with the Declaration of Helsinki, and all participating patients provided written informed consent. The Leiden EAC was approved by the Medical Ethics Committee of the LUMC (B19.008).

From June 2013 onwards, the EAC protocol included contrast-enhanced MRI of the forefoot. For the current study, we included 157 consecutive DMARD-naive, early RA patients from the EAC who were enrolled from June 2013 to March 2016. Fourteen patients with early RA were excluded because of missing baseline MRIs, and 5 others were excluded because of insufficient quality of MRIs. RA was defined as a clinical diagnosis plus fulfilment of the 2010 RA classification criteria within 1 year after inclusion (25).

Clinical signs typical for RA. Joint tenderness and swelling were assessed at physical examination by a trained research nurse. Joint swelling was also assessed independently by a rheumatologist and was considered positive if both assessors agreed on its presence in the same joint (26). Research nurses participate regularly in consensus exercises for joint examination led by a rheumatologist to maintain interobserver agreement.

MRI scanning and baseline evaluation. Shortly after the first visit (when clinical assessment was done) and before any DMARDs were started (the period between the first visit and MRI was 9 days on average), all patients underwent unilateral contrast-enhanced 1.5T ONI MRI (GE) of the first through the fifth MTP joints on the most painful side. In the case of symmetrically severe symptoms, the dominant side was scanned. The MRI protocol is described in more detail in Supplementary Appendix A,

available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640. All MRIs were scored blinded for clinical data.

The intermetatarsal bursae lie in the superior intermetatarsal spaces, which are bordered medially and laterally by the metatarsal heads, dorsally by the deep dorsal aponeurosis, and plantarly by the deep transverse metatarsal ligament (10,12). IMB was therefore defined as contrast enhancement of the bursa in the superior intermetatarsal space, with or without rim enhancement, visible on ≥2 consecutive slices in both planes (axial and coronal). For each superior intermetatarsal space (4 per foot), presence of IMB was recorded by 2 independent readers (YJD and MR), who then determined the final scores by consensus; an IMB lesion was considered present if both agreed on this. This IMB scoring method was described previously; the specificity for RA of IMB presence assessed in this manner was 84% compared to healthy controls (14). Next to IMB presence, the size of the lesions in dorsoplantar direction (in mm) was recorded (14) to enable assessment of changes in size at follow-up. Size measurements are described in more detail in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640.

To assess the relation between IMB and other MRI measures of local inflammation, MRIs were also evaluated for synovitis, tenosynovitis, and osteitis in line with the RAMRIS system by 2 independent readers, as described previously (2,27,28). To obtain the total RAMRIS inflammation score for each patient, the scores for synovitis, tenosynovitis, and osteitis were summed; the average of the scores of both readers was used (29). At joint level, presence of RAMRIS inflammation was stringently defined based on consensus: synovitis, tenosynovitis, or osteitis were considered present per location if that feature was scored as ≥1 by both readers independently, concordant to the literature (26). Detailed information on RAMRIS inflammation scoring is presented in Supplementary Appendix A, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640.

Follow-up MRIs. MRIs were repeated over time (scheduled at 4, 12, and 24 months from baseline) in patients included from June 2013 until February 2015; a flowchart illustrating patient selection for longitudinal analyses is presented in Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640. The course of IMB was evaluated longitudinally in patients who had ≥1 IMB lesion at baseline and received early DMARD therapy. "Early" therapy was defined as DMARD initiation (including glucocorticoids) within 100 days from first presentation at the outpatient clinic.

For IMB, both its presence and lesion size were evaluated. For the latter, a composite measure was used: the averaged IMB size (in mm), calculated by summing the dorsoplantar sizes of all IMB lesions in any intermetatarsal space and dividing the result by 4 (the maximum number of IMB lesions). The dorsoplantar size was used and not the transversal size because intermetatarsal bursae are confined transversally by the metatarsal heads and may, theoretically, distend dorsoplantarly more freely (30).

MRIs were scored in known time order. IMB presence and size were assessed without simultaneously performing RAMRIS scoring. In addition, the same set of MRIs was scored by another independent trained reader, who performed RAMRIS scoring. Interreader and intrareader intraclass correlation coefficients (ICCs) for IMB were 0.90 and 0.85, respectively (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640). For the RAMRIS system, interreader and intrareader ICCs were ≥0.90, as published previously (2).

Statistical analyses. First, the association between presence of IMB and RAMRIS inflammation at baseline was assessed at the patient and bursa levels. At the patient level, univariable logistic regression was used with continuous scores for RAMRIS inflammation (synovitis, tenosynovitis, osteitis, and total RAMRIS inflammation score) as independent variables. Bursa-level analyses were performed using univariable generalized estimating equations (GEEs), with presence of RAMRIS inflammation (synovitis, tenosynovitis, osteitis, and presence of any of these 3) in the 2 joints neighboring the bursa as independent variables. Both at the patient and bursa levels, multivariable models with synovitis, tenosynovitis, and osteitis as separate independent variables were performed because these features often co-occur.

Secondly, it was assessed at joint level whether presence of IMB at baseline contributes to 2 clinical signs typical for RA: joint tenderness and swelling. Univariable GEEs were used with tenderness or swelling of the MTP joint as outcome, and IMB presence in the adjacent intermetatarsal space as independent variable. Multivariable GEEs adjusted for concurrent presence of RAMRIS inflammation (synovitis, tenosynovitis, or osteitis). The first MTP joint was not included in these analyses because it is a predilection site for other diseases than RA (e.g., gout and osteoarthritis), which could introduce tenderness or swelling unrelated to RA (26).

Longitudinally, at patient level, the mean averaged IMB size and total RAMRIS inflammation were modeled over time using GEEs and visualized in 1 graph. The DAS score (calculated with a 4-component formula based on 44 joints) (31) was plotted as well. Associations between the time courses of IMB and RAMRIS inflammation were assessed at the patient level using univariable GEEs, with changes in averaged IMB size as dependent variables, and changes in RAMRIS inflammation (synovitis, tenosynovitis, osteitis, and total scores) as independent variables. Again, a multivariable GEE with the 3 inflammation features as separate independent variables was performed. GEE models were limited to the 0–4 and 4–12 month intervals to optimize the fit because thereafter, patient numbers were lower and MRI-detected inflammation was stable.

Table 1. Baseline characteristics of all rheumatoid arthritis patients studied according to presence of IMB*

Characteristic	All patients (n = 157)	IMB at baseline	
		Present (n = 109)	Absent (n = 48)
Age, mean ± SD years	59 ± 14	58 ± 14	61 ± 14
Female, no. (%)	109 (69)	74 (68)	35 (73)
BMI, mean ± SD	26 ± 5	26 ± 4	26 ± 5
Symptom duration, weeks	11 (5–28)	10 (5–27)	12 (5–31)
SJC†	7 (3–11)	8 (4–11)	4 (1–11)
TJC	5 (3–7)	5 (3–7)	5 (3–8)
Swollen MTP joint(s), no. (%)	57 (36)	43 (39)	14 (29)
ESR, mm/hour	28 (14–45)	28 (14–41)	27 (10–49)
DAS, mean ± SD	3.1 ± 0.8	3.1 ± 0.8	3.0 ± 0.8
ACPA positive, no. (%)	83 (53)	62 (57)	21 (44)
RF positive, no. (%)†	101 (64)	76 (70)	25 (52)
No. of IMB lesions	1 (0–3)	2 (1–3)	–

* Values are the median (interquartile range) unless indicated otherwise. ACPA = anti–citrullinated protein antibody; BMI = body mass index; DAS = Disease Activity Score; ESR = erythrocyte sedimentation rate; IMB = intermetatarsal bursitis; MTP = metatarsophalangeal; RF = rheumatoid factor; SJC = swollen joint count; TJC = tender joint count.
† Statistically significant at the 0.05 level.

Sensitivity analyses were performed by repeating the patient-level longitudinal analyses in the subgroup of patients who received methotrexate as initial DMARD therapy since methotrexate was most often used as a first-line DMARD therapy, as recommended by international guidelines for RA management (32).

Analyses were repeated with stratification for ACPA status because ACPA-positive and ACPA-negative RA are considered different entities (22,33,34). Effects of ACPA status on the time courses of IMB and total RAMRIS inflammation at the patient level were assessed by adding ACPA status and the interaction between ACPA status and MRI time point as independent variables to the longitudinal models. SPSS, version 25, was used. Two-sided *P* values less than 0.05 were considered statistically significant. Data are available from the corresponding author upon reasonable request.

RESULTS

Patient characteristics. Baseline characteristics are presented in Table 1. IMB was present in 109 patients (69%). IMB was more often present in patients with a higher swollen joint count and tended to be more often present in ACPA-positive RA (75% versus 64% in ACPA negative; $P = 0.13$).

IMB occurs together with tenosynovitis and synovitis. Patients with MRI-detected IMB were more likely to have higher total RAMRIS inflammation scores (Table 2). Also, evaluation of synovitis, tenosynovitis, and osteitis separately showed that patients with IMB were more likely to have higher scores for all these inflammatory features. Multivariable analyses showed that IMB presence was associated with high tenosynovitis scores. Thus, patients with a higher severity of tenosynovitis had IMB more frequently.

Analyses were then done at the bursa level. IMB was more often present at locations with synovitis, tenosynovitis, or osteitis in the adjacent MTP joints (Table 3). In multivariable analyses, the presence of IMB was associated with local presence of synovitis and tenosynovitis, with odds ratios (ORs) of 1.69 (95% confidence interval [95% CI] 1.12, 2.57) and 2.83 (95% CI 1.80, 4.44), respectively. In contrast, it was not associated with presence of inflammation in the adjacent bones (osteitis) in multivariable analysis. Two example MRI images of IMB co-occurring with tenosynovitis are presented in Figure 1. Additional example

Table 2. The association between the presence of IMB and RAMRIS inflammation scores in the forefoot at the patient level at first presentation (n = 157)*

	Univariable				Multivariable†	
	Range	Median (IQR)	OR (95% CI)	*P*	OR (95% CI)	*P*
Total RAMRIS inflammation	0–30	4 (1–9)	1.51 (1.28, 1.78)	<0.001	–	–
Synovitis	0–10	1 (0–3)	1.98 (1.45, 2.72)	<0.001	1.12 (0.77, 1.65)	0.55
Tenosynovitis	0–12	1 (0–3)	3.42 (1.97, 5.95)	<0.001	2.92 (1.62, 5.24)	<0.001
Osteitis	0–20	1 (0–3)	1.70 (1.27, 2.27)	<0.001	1.38 (0.97, 1.97)	0.074

* 95% CI = 95% confidence interval; IMB = intermetatarsal bursitis; IQR = interquartile range; OR = odds ratio; RAMRIS = Rheumatoid Arthritis Magnetic Resonance Imaging Scoring (system).
† Including synovitis, tenosynovitis, and osteitis scores as independent variables.

Table 3. The association between the presence of IMB and the presence of RAMRIS inflammation in neighboring joints at first presentation (n = 628 bursae)*

	Univariable		Multivariable†	
	OR (95% CI)	*P*	OR (95% CI)	*P*
Any feature	2.67 (1.91, 3.73)	<0.001	–	–
Synovitis	2.63 (1.84, 3.76)	<0.001	1.69 (1.12, 2.57)	0.013
Tenosynovitis	3.69 (2.40, 5.67)	<0.001	2.83 (1.80, 4.44)	<0.001
Osteitis	1.99 (1.33, 2.98)	0.001	1.30 (0.81, 2.08)	0.28

* 95% CI = 95% confidence interval; IMB = intermetatarsal bursitis; OR = odds ratio; RAMRIS = Rheumatoid Arthritis Magnetic Resonance Imaging Scoring (system).
† Including synovitis, tenosynovitis, and osteitis presence as independent variables.

MRI images are presented in Supplementary Figure 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640.

IMB contributes to joint tenderness and swelling independent of RAMRIS inflammation. One hundred fifty-seven RA patients contributed 628 MTP joints, of which 200 (32%) were tender and 81 (13%) were swollen. Joints with adjacent IMB were more likely to be tender (OR 2.1 [95% CI 1.3, 3.4]) and swollen (OR 3.1 [95% CI 1.6, 6.2]). Multivariable analyses showed that IMB presence was associated with both clinical signs independent of synovitis, tenosynovitis, and osteitis (adjusted OR [OR_{adj}] of 1.7 [95% CI 1.04, 2.9] for tenderness and OR_{adj} of 2.7 [95% CI 1.3, 5.3] for swelling).

IMB decreases after DMARD initiation in a fashion similar to synovitis and tenosynovitis. Of the 109 patients who were IMB positive at baseline, 101 received early DMARD therapy, of whom 73 patients were included before February 2015 (the period wherein follow-up MRIs were made; see Supplementary Figure 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640). Follow-up MRI was available for 55 (75%) of these patients. Longitudinal MRIs (at 4, 12, and 24 months) of these 55 patients were evaluated to assess the time course of IMB after DMARD initiation.

The time courses of IMB and total RAMRIS inflammation are depicted in Figure 2A. Both measures decreased statistically significantly between baseline and 12 months: the mean averaged IMB size was 6.7 mm at baseline and decreased by 3.1 mm (95% CI 2.2, 4.1; $P < 0.001$), while the mean total RAMRIS inflammation was 7.7 points at baseline and decreased by 4.1 points (95% CI 2.4, 5.7; $P < 0.001$) between 0 and 12 months.

Next, we assessed the relation between changes in IMB and simultaneous changes in RAMRIS inflammation over time. Between baseline and 12 months, greater decrease in averaged IMB size was statistically significantly associated with greater decrease in total RAMRIS inflammation (Figure 2A). The 3 RAMRIS inflammation features were also assessed separately for their relation to IMB decrease (see Supplementary Table 2, available

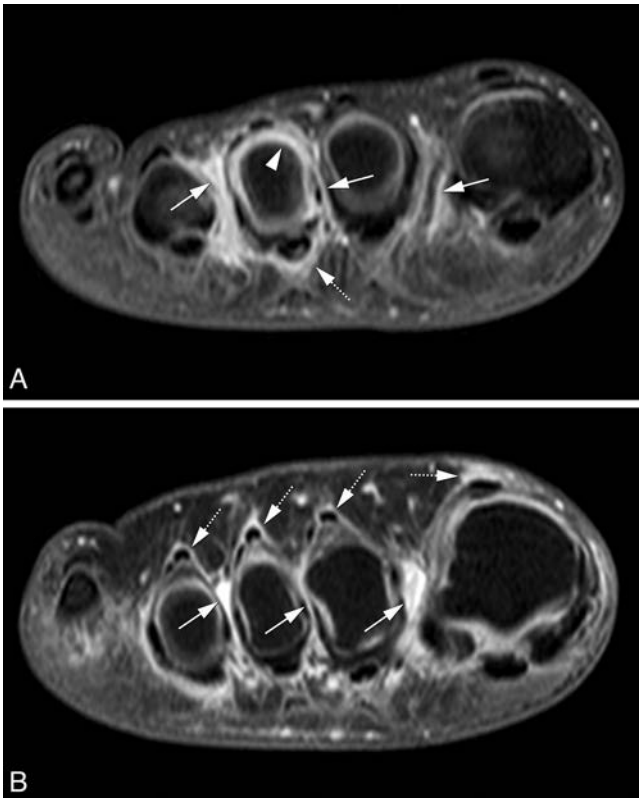

Figure 1. Magnetic resonance imaging (MRI) of 2 disease-modifying antirheumatic drug–naive, early rheumatoid arthritis patients showing MRI-detected intermetatarsal bursitis (IMB) co-occurring with synovitis and flexor tenosynovitis (**A**) and with extensor tenosynovitis (**B**) using coronal T1-weighted fat-suppressed images after gadolinium administration of the forefoot at the level of the metatarsal heads. Both patients show enhancement of thickened synovium in the intermetatarsal spaces 1–3, consistent with IMB (**arrows**). Patient A (female, 33 years old) (**A**) shows peripheral enhancement in the third intermetatarsal space with a central area of lower signal intensity consistent with fluid. At the third metatarsophalangeal joint, there is enhancement surrounding the flexor tendon consistent with tenosynovitis (46) (**dotted arrows**) as well as synovitis (**arrowhead**). Patient B (female, 41 years old) (**B**) shows peripheral enhancement in the first intermetatarsal space with a central area of lower signal intensity consistent with fluid. In addition, there is enhancement surrounding extensor tendons consistent with tenosynovitis (**dotted arrows**) (46).

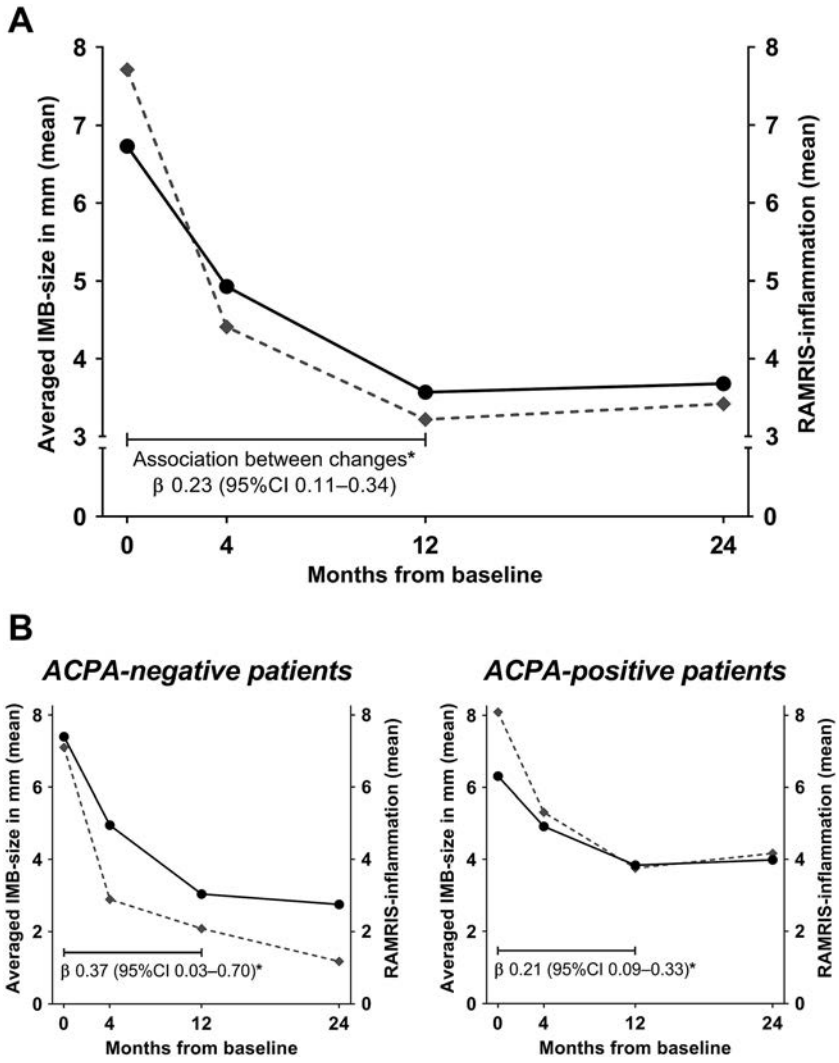

Figure 2. Intermetatarsal bursitis (IMB) size (circles) and total Rheumatoid Arthritis Magnetic Resonance Imaging Scoring (RAMRIS) system inflammation (diamonds) in the forefoot over time (n = 55) in all patients (**A**) and in anti–citrullinated protein antibody (ACPA)–positive and ACPA-negative patients separately (**B**). * = β signifies the association of change in averaged IMB size with change in total RAMRIS inflammation between 0 and 12 months, estimated using generalized estimating equations. IMB decrease was statistically significantly associated with total RAMRIS inflammation decrease at the 0.05 level in all patients (**A**) and in both ACPA subsets (**B**). 95% CI = 95% confidence interval.

on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640). In univariable analyses, patients with greater decreases in synovitis and tenosynovitis on average underwent a greater simultaneous decrease in IMB. Multivariably, IMB decrease was associated with synovitis decrease in the same time interval. Notably, IMB decrease was not related to osteitis decrease, both in univariable and multivariable models. Last, DAS score over time in relation to IMB was plotted (see Supplementary Figure 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640).

Longitudinal MRI images of a patient showing decreasing IMB are presented in Figure 3. An additional series is presented in the supplementary file (see Supplementary Figure 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640).

Sensitivity analysis. Forty-seven of 55 longitudinally studied patients (85%) received methotrexate as initial DMARD therapy. Longitudinal analyses were repeated in this subgroup. Results were similar to those of the main analyses (see Supplementary Tables 3–4 and Supplementary Figures 6–7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640).

Analyses stratified for ACPA positivity. Analyses of the relation between IMB and RAMRIS inflammation at baseline and in the first year of follow-up did not show meaningful differences between ACPA-positive and ACPA-negative RA (see Supplementary Tables 5–7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640), except that in univariable analyses at baseline the association between

IMB and osteitis was statistically significantly only in ACPA-positive patients. IMB presence at baseline was associated with local joint tenderness independent of RAMRIS inflammation only in ACPA-positive patients (OR_{adj} 3.0 [95% CI 1.6, 5.6]; see Supplementary Table 8, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24640). The association with joint swelling seemed present in both groups but only reached statistical significance in ACPA-positive patients (OR_{adj} 3.3 [95% CI 1.2, 9.1]).

Longitudinally, decreases in both averaged IMB size and total RAMRIS inflammation between 0 and 12 months appeared to be more pronounced in ACPA-negative RA (Figure 2B). This was statistically significant for IMB (2.4 mm [95% CI 0.4, 4.4] greater decrease in ACPA-negative versus ACPA-positive RA) but not for total RAMRIS inflammation (0.2 [95% CI –3.1, 3.6] points greater decrease in ACPA-negative versus ACPA-positive RA). Baseline values of averaged IMB size and total RAMRIS inflammation were not statistically significantly different.

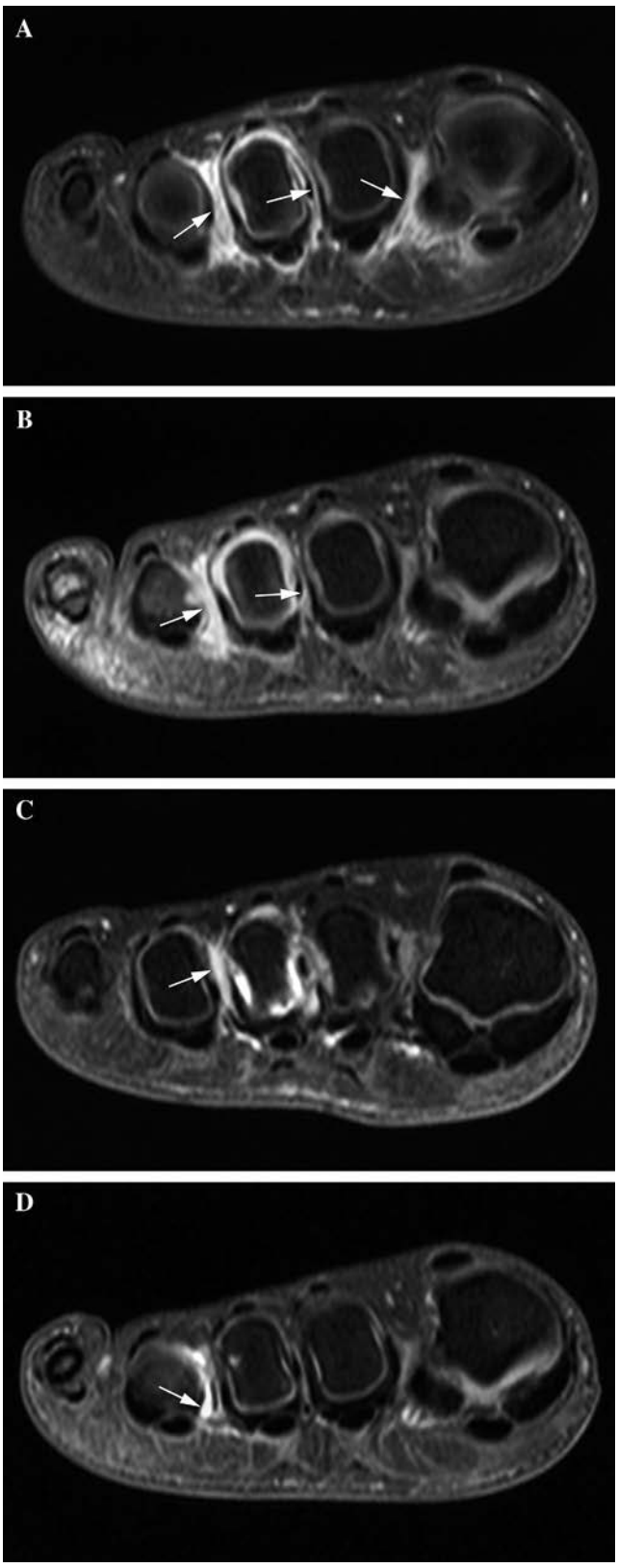

Figure 3. Longitudinal magnetic resonance imaging of decreasing intermetatarsal bursitis (IMB) in an early rheumatoid arthritis patient (female, 33 years old at baseline; corresponds to patient in Figure 1A) using coronal T1-weighted fat-suppressed images after gadolinium administration of the forefoot at the level of the metatarsal heads. The different time points are shown vertically: baseline (**A**), 4 months (**B**), 12 months (**C**), and 24 months (**D**). IMB (**arrows**) is visible in intermetatarsal spaces 1–3 with concomitant synovitis and flexor tenosynovitis at the third metatarsophalangeal joint. All inflammation decreased after initiation of disease-modifying antirheumatic drugs; minimal IMB in the third space remained visible after 2 years.

DISCUSSION

RA is traditionally known as an autoimmune disease targeting the synovial lining of (small) joints. There is mounting evidence indicating that juxtaarticular synovial inflammation is an important trait of the disease as well. Recently, tenosynovitis was the first feature of such juxtaarticular inflammation to be identified as a trait of RA (6). Our study shows that IMB frequently occurs in RA at the time of diagnosis, especially when synovitis and tenosynovitis were also present, and that it contributes to joint tenderness and swelling independent of known MRI features. These data enhance our understanding of forefoot inflammation in RA and support the notion that IMB might be another feature of juxtaarticular synovial inflammation in RA.

The current study is the first to investigate IMB in early RA during follow-up and in relation to known RA inflammation features (RAMRIS inflammation). We demonstrated that IMB decreased after DMARD initiation; a decrease that was most strongly related to a decrease in synovitis severity. This decrease of IMB was as one would expect from an RA treatment response. These findings may therefore further support the notion that IMB is truly a feature of RA.

Recognition of IMB is clinically relevant because it could add to the set of RA features and characteristics that physicians may

consider when evaluating patients with (suspected) RA. Our findings suggest that IMB contributes to the clinical appearance of metatarsalgia and arthritis. More specifically, it could aid in the interpretation of forefoot symptoms and walking disabilities in the absence of synovitis on imaging.

While IMB in RA has been described in small case reports and larger studies in long-standing disease (16,35–37), the current study is the first large MRI study in early RA. The prevalence of MRI-detected IMB in our baseline sample was published previously and amounted to 69% (14), which is in line with the 63% previously reported in a small MRI study in early RA (n = 30 patients) (17). The results of the present and previous imaging studies are concordant in their finding that IMB is prevalent in a majority of RA patients at diagnosis. In addition, data of our study suggest that IMB is especially present in patients presenting with extensive inflammation, measured by the swollen joint count or total RAMRIS inflammation.

The association between IMB and joint swelling was described previously in a study on early arthritis, which also included the RA patients studied here (26). However, the current finding that this association is present in RA patients specifically, independent of RAMRIS inflammation, is novel. Moreover, we now also show that IMB contributes to joint tenderness, which is a subject of utmost importance from the patient perspective (38). The association of IMB with joint swelling appeared somewhat stronger than its association with tenderness, generating the question whether the latter is partly caused by the former. When restricting analyses to nonswollen joints only (n = 540), the effect size changed only slightly (the OR_{adj} for RAMRIS inflammation went from 1.7 [95% CI 1.04, 2.9] to 1.5 [95% CI 0.8, 2.8]). In our view, this suggests that IMB contributes not only to joint swelling but also to tenderness without clinical swelling.

IMB was frequently present in both ACPA-positive and ACPA-negative RA at diagnosis. The prevalence was higher in the ACPA-positive group, but this finding was not statistically significant. The association of IMB presence with simultaneous presence of synovitis (at the joint level) and tenosynovitis (at the patient and joint level) was also positive in both groups. While ACPA-positive patients were more likely to have IMB in the presence of osteitis in univariable analyses, this association was not statistically significant in ACPA-negative patients. This difference between ACPA-positive and ACPA-negative RA is in line with previous findings showing that osteitis associates particularly with ACPA-positive RA (39,40). Despite similar associations of IMB with RAMRIS inflammation, associations with joint tenderness and swelling were more prominent in ACPA-positive than ACPA-negative RA. In both RA groups, however, IMB decreased significantly over time in a fashion similar to total RAMRIS inflammation. Moreover, patients in both groups showed greater IMB decrease when total RAMRIS inflammation decreased more strongly. Thus, although ACPA-positive and ACPA-negative RA have differences in risk factors, presumed etiology, and severity of disease course (33,34,41), IMB is prevalent and behaves similarly in relation to RAMRIS inflammation in both disease subsets.

Hypothetically, mechanical strain could promote development of bursitis. If so, one may assume that IMB decrease is secondary to decreasing mechanical pressure from reduction in neighboring synovitis (18–20). Exploratory analyses showed that IMB lesions with adjacent synovitis at baseline did not dissipate more often (44% after 12 months versus 57% for IMB lesions without adjacent synovitis; $P = 0.17$), arguing against a secondary treatment effect by decreasing mechanical pressure from neighboring synovitis.

We measured IMB size in dorsoplantar direction according to the literature and because this was expected to represent total lesion size more accurately than axial measurements. Intermetatarsal bursae are confined axially by the metatarsal heads and may therefore distend dorsoplantarly more freely (30). Influence of mechanical factors on dorsoplantar distension was most likely limited, as MRIs were made in supine, non–weight bearing position. Still, potentially relevant aspects of the time course of IMB might have gone unnoticed by focusing on the dorsoplantar dimension. Ideally, total IMB volume is used, but reliable measuring methods were not available and beyond the scope of the current investigation.

An important strength of the current study was the relatively large sample size at baseline compared to previous imaging studies of IMB in early RA (17,35). Second, results were robust across patient- and joint-level analyses. Last, owing to the design of the Leiden EAC, which is an inception cohort with extensive follow-up including MRI scans at multiple time points, we were able to perform novel longitudinal analyses of MRI-detected IMB in early RA.

There are also limitations. First, the method we used to score the presence and size of IMB lesions is novel and not yet systematically validated. On the other hand, it was developed in collaboration with a musculoskeletal radiologist (MR) with >20 years of experience, and interreader and intrareader reliability were good (ICCs ≥0.85). Second, MRIs were scored in chronological order to achieve better sensitivity to change, which is in line with the literature (2,27,42–44). Theoretically, this may have caused bias in the form of greater change scores than would have been the case with blinding for time order. However, impact on the main objective to assess associations between IMB and RAMRIS inflammation over time is assumed to be limited, as we have no reason to believe that the improvement in sensitivity to change is different between IMB and RAMRIS inflammation. Third, regression to the mean could have occurred in the longitudinal part of the study since only patients who were IMB positive at baseline were included. It might be interesting to assess in a subsequent study whether IMB-negative patients may develop IMB over time despite receiving DMARDs or during flares. Furthermore, as RA patients were treated and we did not perform a randomized clinical trial with a placebo arm, we interpreted the decrease in DAS

score, RAMRIS inflammation, and IMB as treatment response, but this was not formally proven. Although both IMB and RAMRIS inflammation decreased statistically and numerically significantly (by 46% and 53%, respectively), minimal reference values for determining a response in these measures are not available. We also had insufficient power to stratify analyses by individual DMARDs other than methotrexate. Finally, any association of IMB with deviations of forefoot bones (e.g., hallux valgus and hammer toes) that might hypothetically influence IMB could not be taken into account, as no weightbearing radiographs were made (18–20,45).

Recognition of IMB as an RA feature paves the way for further study of its properties in the disease. A case report suggested that IMB can be recognized by the feature "opening toes" related to enlargement of the space between adjacent toes (37). Although such a clinical sign to detect IMB has the advantage of being less costly and time-consuming than MRI, it has so far not been systematically studied. The contribution of IMB to walking disabilities, including the role therein of biomechanical factors such as pressure distribution, is another subject for further research. For RA patients with prominent foot symptoms and/or walking disabilities, it would be especially valuable to see if amelioration of IMB correlates with symptomatic and functional improvement and, if so, which individual DMARDs or additional therapeutic approaches influence IMB and forefoot symptoms. In addition, it could be studied whether IMB is of pathophysiologic relevance or just reflects extensive synovial inflammation pertaining to higher disease activity. For example, the causal relation between synovitis, tenosynovitis, and IMB could be investigated in a histopathologic study differentiating the types of synovitis. Last, although IMB has been reported to be detectable by ultrasound (15), which is more easily accessible in daily practice than MRI, its correlation with MRI-detected IMB in early RA has not yet been studied.

In conclusion, IMB behaves in line with known RA characteristics; it particularly accompanies inflammation of the synovial lining of joints and tendon sheaths, shows a similar treatment response after DMARD initiation, and contributes to typical clinical signs. These findings support the notion that IMB is a novel inflammatory feature of early RA.

ACKNOWLEDGMENTS

We thank G. Kracht for his assistance in preparing the MRIs. A. Boer, D. Boeters, W. Nieuwenhuis, and E. Newsum are acknowledged for RAMRIS scoring of baseline MRIs.

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 Dijk had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. van Dijk, van der Helm-van Mil.
Acquisition of data. van Dijk, Dakkak, Matthijssen, Reijnierse.
Analysis and interpretation of data. van Dijk, Dakkak, Matthijssen, Niemantsverdriet, Reijnierse, van der Helm-van Mil.

REFERENCES

1. Østergaard M, Bird P, Gandjbakhch F, Eshed I, Haugen IK, Haavardsholm EA, et al. The OMERACT MRI in Arthritis Working Group: update on status and future research priorities. J Rheumatol 2015;42:2470–2.
2. Dakkak YJ, Matthijssen XM, van der Heijde D, Reijnierse M, van der Helm-van Mil AH. Reliability of magnetic resonance imaging (MRI) scoring of the metatarsophalangeal joints of the foot according to the rheumatoid arthritis MRI score. J Rheumatol 2020;47:1165–73.
3. Colebatch AN, Edwards CJ, Østergaard M, van der Heijde D, Balint PV, D'Agostino MA, et al. EULAR recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis. Ann Rheum Dis 2013;72:804–14.
4. Østergaard M, Peterfy CG, Bird P, Gandjbakhch F, Glinatsi D, Eshed I, et al. The OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging (MRI) scoring system: updated recommendations by the OMERACT MRI in Arthritis Working Group. J Rheumatol 2017;44:1706–12.
5. Nieuwenhuis WP, van Steenbergen HW, Mangnus L, Newsum EC, Bloem JL, Huizinga TW, et al. Evaluation of the diagnostic accuracy of hand and foot MRI for early rheumatoid arthritis. Rheumatology (Oxford) 2017;56:1367–77.
6. Rogier C, Hayer S, van der Helm-van Mil A. Not only synovitis but also tenosynovitis needs to be considered: why it is time to update textbook images of rheumatoid arthritis. Ann Rheum Dis 2020;79:546.
7. Mankia K, Agostino MA, Rowbotham E, Hensor EM, Hunt L, Möller I, et al. MRI inflammation of the hand interosseous tendons occurs in anti-CCP-positive at-risk individuals and may precede the development of clinical synovitis. Ann Rheum Dis 2019;78:781.
8. Eshed I, Feist E, Althoff CE, Hamm B, Konen E, Burmester GR, et al. Tenosynovitis of the flexor tendons of the hand detected by MRI: an early indicator of rheumatoid arthritis. Rheumatology (Oxford) 2009;48:887–91.
9. Brooks F, Hariharan K. The rheumatoid forefoot. Curr Rev Musculoskelet Med 2013;6:320–7.
10. Theumann NH, Pfirrmann CW, Chung CB, Mohana-Borges AV, Haghighi P, Trudell DJ, et al. Intermetatarsal spaces: analysis with MR bursography, anatomic correlation, and histopathology in cadavers. Radiology 2001;221:478–84.
11. Awerbuch MS, Shephard E, Vernon-Roberts B. Morton's metatarsalgia due to intermetatarsophalangeal bursitis as an early manifestation of rheumatoid arthritis. Clin Orthop Relat Res 1982:214–21.
12. Chauveaux D, Le Huec JC, Midy D. The supra-transverse intermetatarsocapital bursa: a description and its relation to painful syndromes of the forefoot. Surg Radiol Anat 1987;9:13–8.
13. Jovanovic MS, Royer J, Roy PE, Caron P, Houde G, Houde JP, et al. A comparative study of the spaces between the metacarpal and metatarsal heads. Surg Radiol Anat 1990;12:31–6.
14. Dakkak YJ, Niemantsverdriet E, van der Helm-van Mil AH, Reijnierse M. Increased frequency of intermetatarsal and submetatarsal bursitis in early rheumatoid arthritis: a large case-controlled MRI study. Arthritis Res Ther 2020;22:277.
15. Bowen CJ, Culliford D, Dewbury K, Sampson M, Burridge J, Hooper L, et al. The clinical importance of ultrasound detectable forefoot bursae in rheumatoid arthritis. Rheumatology (Oxford) 2009;49:191–2.
16. Hammer HB, Kvien TK, Terslev L. Intermetatarsal bursitis is frequent in patients with established rheumatoid arthritis and is associated with

anti-cyclic citrullinated peptide and rheumatoid factor. RMD Open 2019;5:e001076.

17. Boutry N, Lardé A, Lapègue F, Solau-Gervais E, Flipo RM, Cotten A. Magnetic resonance imaging appearance of the hands and feet in patients with early rheumatoid arthritis. J Rheumatol 2003;30:671.
18. Bowen CJ, Culliford D, Allen R, Beacroft J, Gay A, Hooper L, et al. Forefoot pathology in rheumatoid arthritis identified with ultrasound may not localise to areas of highest pressure: cohort observations at baseline and twelve months. J Foot Ankle Res 2011;4:25.
19. Helliwell P, Siddle H, Redmond A, editors. The foot and ankle in rheumatology. Reports on the Rheumatic Diseases; 2011.
20. Nouh MR, Khalil AA. Forefoot: a basic integrated imaging perspective for radiologists. Clin Imaging 2014;38:397–409.
21. Sundin U, Aga AB, Skare Ø, Nordberg LB, Uhlig T, Hammer HB, et al. Conventional versus ultrasound treat to target: no difference in magnetic resonance imaging inflammation or joint damage over 2 years in early rheumatoid arthritis. Rheumatology (Oxford) 2020;59:2550–5.
22. Matthijssen XM, Niemantsverdriet E, Le Cessie S, van der Helm-van Mil AH. Differing time-orders of inflammation decrease between ACPA subsets in RA patients suggest differences in underlying inflammatory pathways. Rheumatology (Oxford) 2021;60:2969–75.
23. De Rooy DP, van der Linden MP, Knevel R, Huizinga TW, van der Helm-van Mil AH. Predicting arthritis outcomes: what can be learned from the Leiden Early Arthritis Clinic? Rheumatology (Oxford) 2011; 50:93–100.
24. Van der Linden MP, Batstra MR, Bakker-Jonges LE, on behalf of the Foundation for Quality Medical Laboratory Diagnostics, Detert J, Bastian H, et al. Toward a data-driven evaluation of the 2010 American College of Rheumatology/European League Against Rheumatism criteria for rheumatoid arthritis: is it sensible to look at levels of rheumatoid factor? Arthritis Rheum 2011;63:1190–9.
25. 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.
26. Dakkak YJ, Boer AC, Boeters DM, Niemantsverdriet E, Reijnierse M, van der Helm-van Mil AH. The relation between physical joint examination and MRI-depicted inflammation of metatarsophalangeal joints in early arthritis. Arthritis Res Ther 2020;22:67.
27. Haavardsholm EA, Østergaard M, Ejbjerg BJ, Kvan NP, Kvien TK. Introduction of a novel magnetic resonance imaging tenosynovitis score for rheumatoid arthritis: reliability in a multireader longitudinal study. Ann Rheum Dis 2007;66:1216–20.
28. Østergaard M, Peterfy C, Conaghan P, McQueen F, Bird P, Ejbjerg B, et al. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. J Rheumatol 2003;30: 1385–6.
29. Nieuwenhuis WP, Mangnus L, van Steenbergen HW, Newsum EC, Huizinga TW, Reijnierse M, et al. Older age is associated with more MRI-detected inflammation in hand and foot joints. Rheumatology (Oxford) 2016;55:2212–9.
30. Schwalbe G. The tendon sheaths and synovial bursae of the foot, 1896. Translated by Hartmann. Foot Ankle 1981;1:246–69.
31. Van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49:916–20.
32. Smolen JS, Landewé RB, Bijlsma JW, Burmester GR, Dougados M, Kerschbaumer A, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis 2020; 79:685.
33. Daha NA, Toes RE. Are ACPA-positive and ACPA-negative RA the same disease? Nat Rev Rheumatol 2011;7:202–3.
34. Matthijssen XM, Niemantsverdriet E, Huizinga TW, van der Helm-van Mil AH. Enhanced treatment strategies and distinct disease outcomes among autoantibody-positive and -negative rheumatoid arthritis patients over 25 years: a longitudinal cohort study in the Netherlands. PLOS Med 2020;17:e1003296.
35. Albtoush OM, Xenitidis T, Horger M. Intermetatarsal bursitis as first disease manifestation in different rheumatological disorders and related MR-imaging findings. Rheumatol Int 2019;39:2129–36.
36. Bowen CJ, Hooper L, Culliford D, Dewbury K, Sampson M, Burridge J, et al. Assessment of the natural history of forefoot bursae using ultrasonography in patients with rheumatoid arthritis: a twelve-month investigation. Arthritis Care Res (Hoboken) 2010;62:1756–62.
37. Endo Y, Koga T, Eguchi M, Okamoto M, Tsuji S, Takatani A, et al. Utility of power Doppler ultrasonography for detecting forefoot bursae in early rheumatoid arthritis: a case report. Medicine (Baltimore) 2018; 97:e13295.
38. Heiberg T, Finset A, Uhlig T, Kvien TK. Seven year changes in health status and priorities for improvement of health in patients with rheumatoid arthritis. Ann Rheum Dis 2005;64:191.
39. Stomp W, Krabben A, van der Heijde D, Huizinga TW, Bloem JL, van der Helm-van Mil AH, et al. Are rheumatoid arthritis patients discernible from other early arthritis patients using 1.5T extremity magnetic resonance imaging? A large cross-sectional study. J Rheumatol 2014;41:1630–7.
40. Tamai M, Kawakami A, Uetani M, Takao S, Tanaka F, Nakamura H, et al. The presence of anti-cyclic citrullinated peptide antibody is associated with magnetic resonance imaging detection of bone marrow oedema in early stage rheumatoid arthritis. Ann Rheum Dis 2006;65: 133–4.
41. Van der Woude D, van der Helm-van Mil AH. Update on the epidemiology, risk factors, and disease outcomes of rheumatoid arthritis. Best Pract Res Clin Rheumatol 2018;32:174–87.
42. Glinatsi D, Bird P, Gandjbakhch F, Haavardsholm EA, Peterfy CG, Vital EM, et al. Development and validation of the OMERACT Rheumatoid Arthritis Magnetic Resonance Tenosynovitis Scoring System in a multireader exercise. J Rheumatol 2017;44:1688–93.
43. Glinatsi D, Lillegraven S, Haavardsholm EA, Eshed I, Conaghan PG, Peterfy C, et al. Validation of the OMERACT Magnetic Resonance Imaging Joint Space Narrowing Score for the wrist in a multireader longitudinal trial. J Rheumatol 2015;42:2480–5.
44. Van Tuyl LH, van der Heijde D, Knol DL, Boers M. Chronological reading of radiographs in rheumatoid arthritis increases efficiency and does not lead to bias. Ann Rheum Dis 2014;73:391–5.
45. Fuhrmann RA, Layher F, Wetzel WD. Radiographic changes in forefoot geometry with weightbearing. Foot Ankle Int 2003;24:326–31.
46. Dakkak YJ, Jansen FP, DeRuiter MC, Reijnierse M, van der Helm-van Mil AH. Rheumatoid arthritis and tenosynovitis at the metatarsophalangeal joints: an anatomic and MRI study of the forefoot tendon sheaths. Radiology 2020;295:146–54.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1723–1735
DOI 10.1002/acr.24622

American College of Rheumatology
Empowering Rheumatology Professionals

Evidence-Based Research on Effectiveness of Periodontal Treatment in Rheumatoid Arthritis Patients: A Systematic Review and Meta-Analysis

Daniela S. Silva,[1] Flávio Costa,[2] Isabel P. Baptista,[3] Tânia Santiago,[2] Hans Lund,[4] Simon Tarp,[5] José A. P. daSilva,[6] and Robin Christensen[7]

Objective. To gauge the evidence of periodontal therapy's impact on measures of disease activity and systemic inflammatory burden in individuals with rheumatoid arthritis (RA).

Methods. A search for randomized trials and controlled cohort studies of RA patients with periodontitis was conducted on April 7, 2019, with an update on December 17, 2020 in PubMed, Cochrane Library (CENTRAL), Embase, ClinicalTrials.gov, and the World Health Organization International Clinical Trial Registry Platform portal. Two reviewers screened titles and abstracts and selected papers for full-text review. We used Outcome Measures in Rheumatology (OMERACT)–endorsed outcome domains for RA trials and summarized continuous outcomes using standardized mean differences (SMDs) with 95% confidence intervals (95% CIs). We evaluated inconsistency using the I^2 statistic and combined SMDs using random-effects models for the meta-analyses; fixed-effect meta-analyses were used for sensitivity analysis. To explore heterogeneity, we added stratified/meta-regression analyses, expressed in T^2.

Results. Of the 1,909 studies identified, 9 (including 10 comparisons) were eligible for quantitative synthesis (n = 388). Evidence suggested a favorable effect of periodontal treatment on disease activity (SMD –0.88 [95% CI –1.38, –0.38]; n = 311). The Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach was used to judge the estimates' certainty; evidence rated as having low or very low certainty indicated that any possible effect of periodontal treatment in RA is likely to change as more evidence is provided. Selection bias and RA medication stability were highlighted as sources of heterogeneity between studies.

Conclusion. There is an urgent need for a well-designed prospective cohort study (preferably a randomized controlled trial) of patients with RA and periodontitis using rigorous protocols, standardized diagnostic criteria, data collection, and adequate duration of follow-up.

INTRODUCTION

Periodontitis, a destructive inflammatory disease affecting the supporting tissues of teeth, is the most prevalent bacteria-driven chronic disease in humans (1). It has been considered a comorbidity for many other chronic conditions (2). Accumulating evidence indicates that effectively treating periodontitis may also result in significant beneficial effects on coexisting systemic diseases (3,4).

Rheumatoid arthritis (RA) is an autoimmune disease characterized by chronic, painful inflammation of the joints, which eventually leads to joint destruction, disability, and increased mortality (5). The current etiologic paradigm suggests that RA is triggered by a combination of genetic and (yet unknown) environmental factors that lead to the breakdown of immunotolerance at mucosal surfaces, specifically the lungs, gut, and periodontium (6). Researchers have hypothesized that bacteria from the inflamed gingivae operate as a source of citrullinated peptides, which can

Supported by the Oak Foundation (core grant OCAY-18-774-OFIL) and Associação de Reumatologia de Coimbra.

[1]Daniela S. Silva, DMD, MSc: University of Coimbra, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal, and the Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark; [2]Flávio Costa, MD, Tânia Santiago, MD: Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal; [3]Isabel P. Baptista, DDSc, PhD: University of Coimbra, Coimbra, Portugal; [4]Hans Lund, PT, PhD: Western Norway University of Applied Sciences, Bergen, Norway; [5]Simon Tarp, MSc, PhD: The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark; [6]José A. P. daSilva, MD, PhD: Centro Hospitalar e Universitário de Coimbra and University of Coimbra, Coimbra, Portugal; [7]Robin Christensen, MSc, PhD: The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark, and University of Southern Denmark, Odense University Hospital, Odense, Denmark.

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

Address correspondence to José A. P. daSilva, MD, PhD, Rheumatology Department, Centro Hospitalar e Universitário de Coimbra, Praceta Mota Pinto, 3000-075 Coimbra, Portugal. Email: jdasilva@ci.uc.pt.

Submitted for publication September 28, 2020; accepted in revised form April 13, 2021.

SIGNIFICANCE & INNOVATIONS

- Periodontal treatment studies suggest a clinical effect upon measures of rheumatoid arthritis disease activity and inflammation.
- Imprecision, inconsistency, and risk of bias render this study's evidence low to very low certainty.
- There is an urgent need for a well-designed randomized controlled trial.

initiate and fuel the biological processes leading to and maintaining RA (7). Such a progression would suggest that treating periodontitis might result in significant amelioration of the disease process in RA. Additionally, it is plausible that the increase of acute-phase reactants induced by periodontitis may increase RA evaluation scores and lead to overtreatment (8). Effective treatments for periodontitis, which are readily available, might improve disease activity or at least prevent undesirable overtreatment.

To reveal any previous evidence synthesis initiatives exploring the impact of periodontal interventions on patient-important outcomes in RA, we performed a pragmatic search for systematic reviews available from PubMed while preparing our review protocol (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24622). Only 2 systematic reviews were available that examined the nonsurgical periodontal treatment influence on clinical and biochemical measures for RA (9,10). As we were aware that new studies have been published since the last published systematic review search was closed in 2014, we applied an updated, unbiased, and transparent appraisal of the literature using current evidence synthesis techniques (11). So, following the principles of evidence-based research (12), we aimed to summarize the evidence regarding the effect of periodontal therapy on measures of disease activity and actual inflammatory burden in individuals with RA and to provide recommendations for future research addressing the knowledge gaps.

MATERIALS AND METHODS

Protocol and registration. We conducted our systematic review according to a prespecified and publicly registered protocol (PROSPERO CRD42018103359); the full Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol (PRISMA-P) is available in Supplementary Appendix A, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24622 (13). Results and the full manuscript are reported according to the recommendations given in the PRISMA statement (14).

Data sources and search. We conducted our database search on April 7, 2019, with an update on December 17, 2020, in Medline (PubMed), the Cochrane Library (CENTRAL), Excerpta Medica Database (Embase), ClinicalTrials.gov (http://www.clinicaltrials.gov), and the World Health Organization International Clinical Trial Registry Platform portal (WHO-ICTRP), as recommended by the Cochrane Musculoskeletal Group (15).

The following 3 search areas were combined in the study as medical subject headings/keywords, as subheadings, and as free text in the title or abstract: 1) RA; 2) periodontitis; and 3) controlled trials. The specific search strategies were created by a health sciences librarian/information specialist with expertise in systematic review searching. The Medline strategy was developed with input from the project team and then peer reviewed by a second librarian using Peer Review of Electronic Search Strategies (PRESS) (16). Once the Medline strategy was finalized, it was adapted to the syntax and subject headings of the other databases we used. The subsequent specific search strategies are detailed in Supplementary Appendix A, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24622.

The lists of references of retrieved studies and relevant reviews were manually checked to add any citations missed by the electronic searches (including citation search from retrieved reviews). Abstracts from the 2 major international rheumatology and periodontology scientific meetings in 2017 and 2018 of the American College of Rheumatology (ACR), the European Alliance of Associations for Rheumatology (EULAR), the American Academy of Periodontology, and the European Federation of Periodontology were also searched to identify unpublished studies.

Study selection. Eligible studies comprised randomized controlled trials (RCTs) or quasi-randomized clinical trials (qRCTs), i.e., trials or well-controlled cohort studies where participants were allocated to groups even if based on a clearly inadequate method of randomization (17). Eligible studies compared any type of periodontal treatment with usual care and with a sham or no-treatment comparator group in individuals with RA. We included studies reported as full-text, published as abstract only, and those presenting unpublished data. Reports had to have at least an abstract written in English, Portuguese, Spanish, French, Italian, Danish, Swedish, or Norwegian, and no restrictions were made regarding publication date.

Periodontal treatments included all surgical and mechanical nonsurgical periodontal treatment (NSPT) (i.e., scaling, root planing, and subgingival curettage). We also considered antimicrobial therapy (encompassing antiseptics and antibiotics), either locally applied (including mouth rinses, gels, or dentifrices) or systemically administered. Periodontal treatments included adjuncts as part of usual care, such as oral hygiene instruction or support sessions to improve self-help or self-awareness of plaque control. Interventions were compared with no treatment, usual care (e.g., supragingival prophylaxis and standalone oral hygiene instruction), or placebo.

Studies were selected if they included participants with RA as defined by the ACR/EULAR criteria (18) as well as concomitant periodontitis (19) who did not have any concomitant arthritic

conditions or any other rheumatic diseases. Two review authors (DSS and FC) independently screened the titles and abstracts yielded by the search against the inclusion criteria. These 2 reviewers screened the full-text reports and decided whether they met the inclusion criteria. Disagreement was resolved through discussion. Reasons for excluding trials in full text were recorded.

Data collection. Data were collected on a piloted form. Two review authors (DSS and FC) independently extracted study characteristics: methods (study design, total duration, details of any run-in period, date of study, number of centers and their locations, study setting, and withdrawals); participants (number, mean age, age range, sex, disease duration, diagnostic criteria, inclusion and exclusion criteria); interventions (experimental intervention, comparator, concomitant medications, and excluded medications); outcomes (primary and secondary outcomes specified and collected, and time points reported); characteristics of the design of the trial; and notes (funding and notable declarations of interest of trial authors).

The number of events and participants per treatment group for dichotomous outcomes, and means and SDs and number of participants per treatment group for continuous outcomes, were extracted into customized Excel (Microsoft) data tables and Review Manager (Cochrane). Disagreements were resolved by consensus or by involving a third person (JAPdS, RC, or both). The data that support the findings of this study are available from the corresponding author upon reasonable request.

Outcomes and prioritization. The outcome domains collected were developed for RA clinical trials according to the Outcome Measures in Rheumatology (OMERACT) guidance (20), similar to those endorsed by the ACR (21): disease activity, preferably change from baseline (Disease Activity Score in 28 joints [DAS28], Simplified Disease Activity Index [SDAI] score, Clinical Activity Disease Index [CDAI] score); systemic inflammation markers (acute-phase reactants: C-reactive protein [CRP], erythrocyte sedimentation rate [ESR]); response criteria (number of ACR criteria for 20% improvement in disease activity [ACR20], ACR50, ACR70 responders or EULAR response based on disease activity); life impact (patient-relevant outcomes, such as patient global assessment of disease activity [PtGA], visual analog scale or Numerical Rating Scale pain and disability subscales, and Health Assessment Questionnaire); attrition (number of completers versus number of starters); number of serious adverse events; and number of participants who withdraw due to adverse events.

Risk of bias in individual studies. Two review authors (DSS and RC) assessed risk of bias for each included study using the criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions tool, version 1.0 (22). The reviewing authors did not use the Risk of Bias in Non-Randomised Studies – of Interventions (ROBINS-I) tool (23), as all eligible studies were at least qRCTs. Disagreements were resolved by discussion or by involving another author (JAPdS.). Graphic representations of potential bias across studies were computed using Review Manager 5.1 (24). For ease of interpretation, each trial was also tentatively assigned an overall risk of bias: low risk (low for all key domains); high risk (high for ≥1 key domains); and unclear risk (unclear for ≥1 key domains).

Summary measures (effect sizes). Anticipating that different studies would report on major outcome domains using different outcome measures, we summarized continuous outcomes using standardized mean differences (SMDs) with 95% confidence intervals (95% CIs), with the differences in mean change between treatment groups divided by the pooled SD and applied with Hedges's g adjustment. When differences in mean change were unavailable, we used differences in mean values at the end of the treatment (25). When some of the required data were unavailable, we used various customized approximations (i.e., if SDs were missing, we obtained them from a study's confidence intervals, *P* [or t] values, interquartile range [IQR], or standard error of the mean). Binary outcomes were expressed as risk ratios (RRs) with 95% CIs.

Synthesis of results. A meta-analysis was performed for each outcome domain. Statistical heterogeneity was tested with Cochran's Q test, and inconsistency was evaluated as the I^2 (26). We used standard inverse variance random effects for the meta-analyses (27). When quantitative synthesis was not possible for a specific outcome domain, we provided a qualitative synthesis with information presented in the text and tables to summarize and explain the characteristics and findings of the included studies.

We added a number of stratified/meta-regression analyses, stratifying the available studies according to trial characteristics at the study level. Therefore, all the trial-level features collected were considered potential covariates. Among these covariates were various aspects of study design, intervention, comparator, RA definition, periodontitis definition, RA medication, and risk of bias. This stratifying was assessed fitting multiple mixed-effects, restricted maximum likelihood–based meta-regression models (28). A priori, we defined a relevant study level covariate as one that would decrease the between-study variance (tau-squared, estimated as T^2) as a consequence of inclusion in the (mixed-effects) statistical model (29). Evidence for the comparative effectiveness of periodontal treatment on RA-related outcomes was assessed using criteria suggested by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group (30).

RESULTS

Figure 1 summarizes the identification process of our database search, which initially identified 2,678 records and ultimately yielded 1,909 records after duplicates were removed. Some 41 studies were retrieved for full examination based on assessment of the title and abstract. The subsequent full-text assessment resulted in 8 studies (31–38) being eligible for systematic review. A search of CENTRAL, ClinicalTrials.gov, and the WHO-ICTRP contributed 5 additional unique trials: a PhD thesis (39) and 4 ongoing studies (40–43). Upon searching the 2 major international rheumatology and periodontology scientific meetings in 2017 and 2018, 1 unpublished study was identified (44). There was no disagreement concerning the selection of the final 14 studies (29–42) included in the qualitative synthesis. Of these 14 studies, 9 (31–39) were eligible for quantitative analysis (with 10 comparisons).

In the December 17, 2020 search update, we found 358 new references, excluding duplicates. Upon title and abstract assessment, 12 records were retrieved for full-text examination, which resulted in 3 records that fulfilled our inclusion criteria, 2 clinical trial registries (45,46), and a published study (47) (already included in our primary search as clinical trial registry [43]); however, the latter, surprisingly, did not report the core outcomes set for RA.

Study characteristics. The characteristics of the included studies are shown in Table 1. Of the 14 studies included, 5 (32–34,38,39) were RCTs, with a total of 190 participants; 5 (31,35–37,44) were qRCTs, with 244 participants; and 4 were ongoing studies (40–43), with 554 total expected participants. Among the 388 patients in the studies eligible for quantitative analysis (31–39), the median of the mean age was 50 years (IQR 41–62 years). Among

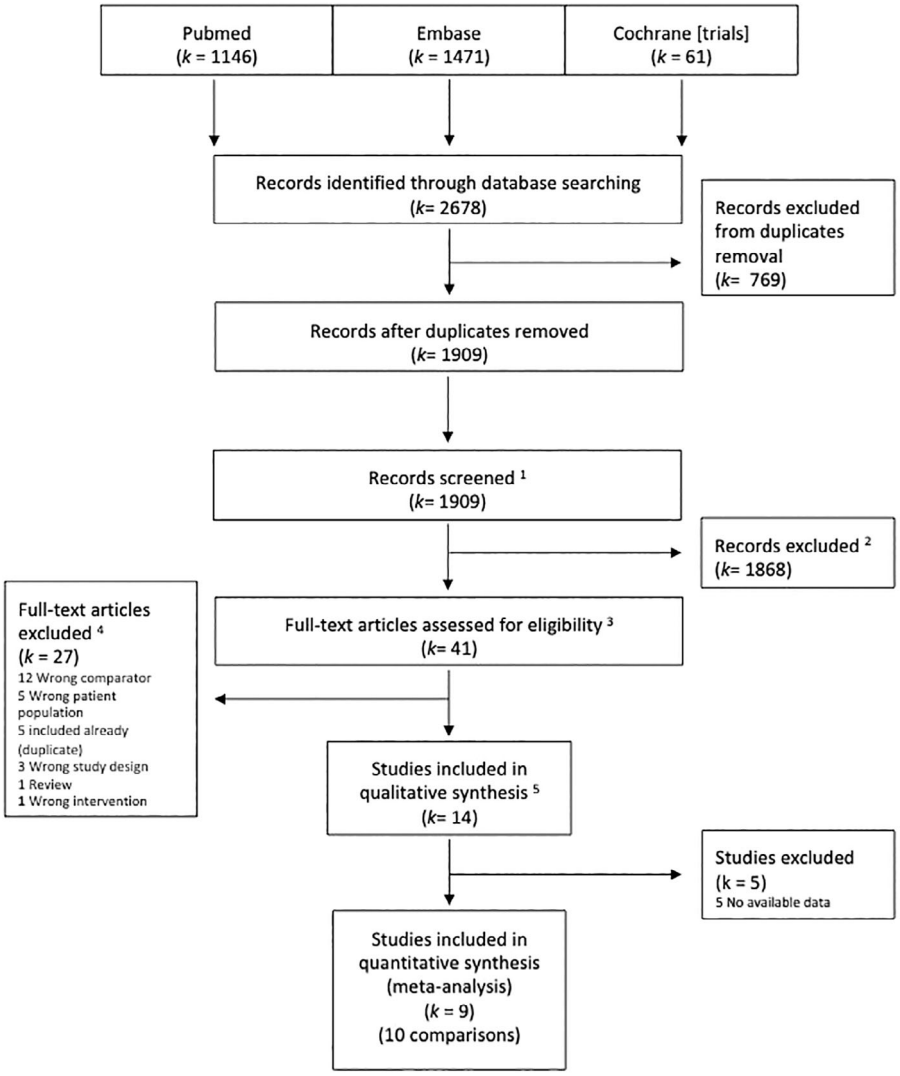

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram for studies retrieved through the search and selection process. December 17, 2020 search update: [1] = 358 new records screened; [2] = 346 records excluded; [3] = 12 full-text articles; [4] = 9 full-text excluded (3 wrong comparator, 2 editorial, 1 letter to the editor, 1 review, 1 wrong intervention, 1 clinical trial registry of excluded study); and [5] = 3 studies included in qualitative synthesis.

Table 1. Characteristics of the eligible studies (qualitative synthesis)*

Author, year (ref.)	Source	Baseline int./comp., no.	Study design	Duration of study and time points, weeks	Diagnosis (eligibility criteria and disease duration)		Mean age, years	Female, %	Interventions		Int./comp. completers, no.
					RA definition	CP definition			Int.	Comp.	
Ribeiro et al, 2005 (31)	Pub.	26/16	qRCT	12	ACR	At least 2 sites ≥5 mm PD; CAL ≥6 mm	48.8	91	OHI + STC + NSPT	OHI + STC	26/16
Al-Katma et al, 2007 (32)	Pub.	19/19	RCT	8	ACR; DAS28 score ≥2.5; RA medication stable during study duration	Generalized mild-to-moderate CP ≥3 years	53.6	86	OHI + NSPT	AT	17/12
Ortiz et al, 2009 (33)	Pub.	20/20	RCT	6	RA	Generalized severe CP	55.5	88	OHI + NSPT a) no anti-TNF, b) anti-TNF	AT a) no anti-TNF, b) anti-TNF	20/20
Pinho et al, 2009 (34)	Pub.	15/15	RCT	12, 24†	ACR	≥2 teeth with CAL ≥6 mm; ≥1 teeth with PD ≥5 mm	47.5	60	NSPT	AT	15/15
Khare et al, 2016 (35)	Pub.	30/30	qRCT	12	ACR	Generalized CP	50	85	OHI + NSPT	AT	30/30
Atarbashi-Moghadam et al, 2018 (36)	Pub.	28/28	qRCT	6, 12†	ACR; RA medication stable until next visit	Generalized moderate-to-severe CP	45.1	79	OHI + NSPT	OHI	28/28
Kaushal et al, 2019 (37)	Pub.	20/20	qRCT	8	≥6 joints involvement; patients receiving medication for RA since 1 month and stable during study	≥2 tooth sites with PD ≥4 mm or CAL ≥4 mm that bled on probing	41.2	80	OHI + NSPT	AT	20/20
Monsarrat et al, 2019 (38)	Pub.	11/11	RCT	12	RA ≥1 year; ACR moderately active RA (DAS28-ESR score 3.2–5.1); no change in RA medication ≤3 months	≥4 teeth with at least 1 site with a CAL of ≥3 mm and a PD of ≥4 mm	61.6	64	OHI + NSPT + systemic antibiotics	AT	11/11

(Continued)

Table 1. (*Cont'd*)

Author, year (ref.)	Source	Baseline int./ comp., no.	Study design	Duration of study and time points, weeks	Diagnosis (eligibility criteria and disease duration)		Mean age, years	Female, %	Interventions		Int./comp. completers, no.
					RA definition	CP definition			Int.	Comp.	
Buwembo et al, 2018 (43)	CTR	152/152	RCT	24	RA ≥2 years; DAS28 score ≥3.2 and ≤5.1; no change in RA medication ≤3 months	CP	18–90	NA	OHI + NSPT + systemic antibiotics	OHI	NA
Kaveri et al, 2016 (42)	CTR	30/30	RCT	12	RA patients taking DMARDs ≥6 months	≥1 site with PD >4 mm, CAL >3 mm	30–65	NA	NSPT	AT	NA
Lozano et al, 2017 (40)	CTR	60/60	RCT	4, 12, 24, 52	ACR/EULAR with high clinical activity	Moderate-to-severe CP	18–64 (n = 80); ≥65 (n = 40)	NA	NSPT + systemic antibiotics	2 STC sessions + CHX solution	NA
Serban, 2016 (39)	PhD thesis	30/30	RCT	12, 24†	ACR; DAS28 score = 3.2; DAS28 score >5.1 if patient taking biologics or patient unwilling to take biologics; treatment with DMARDs ≥3 months and stable dose ≥2 months or patient refusing to take DMARDs	Generalized moderate-to-severe CP; CAL ≥4 mm on at least 2 nonadjacent teeth and cumulative PD ≥40 mm	58	75	NSPT	OHI	23/26
Wolff et al, 2017 (41)	CTR	35/35	RCT	12	Active RA (DAS28 score >3.2)	CP and/or gingivitis: GBI >10%, PCR >30%	≥18	NA	OHI + STC + NSPT	AT	NA
Xiao et al, 2017 (44)	Poster (EULAR)	18/28	qRCT	6, 12†	No changes in DMARD therapy during the study period	Moderate-to-severe CP	NA	NA	OHI + NSPT	AT	NA

* The collected outcome domains were as follows: disease activity, systemic inflammation markers, response criteria, life impact, attrition, and adverse events and withdrawals due to adverse events. Also assessed were the following: radiographs, utility, self-reported painful joint count, psychological status, productivity losses, well-being, sleep disturbance, coping, and leisure (those outcomes were not included, as they were not mentioned among the studies). ACR = American College of Rheumatology; AT = absence of treatment; CAL = clinical attachment loss; CHX = chlorhexidine; Comp. = comparator; CP = chronic periodontitis; CTR = clinical trial registry; DAS28 = Disease Activity Score in 28 joints; DAS28-ESR = DAS28 using the erythrocyte sedimentation rate; DMARDs = disease-modifying antirheumatic drugs; EULAR = European Alliance of Associations for Rheumatology; GBI = gingival bleeding index; Int. = Intervention; NA = not available/applicable; NSPT = nonsurgical periodontal treatment (scaling and root planing); OHI = oral hygiene instructions; PCR = Plaque Control Record; PD = pocket depth; Pub. = published; qRCT = quasi-randomized clinical trial; RA = rheumatoid arthritis; RCT = randomized clinical trial; ref. = reference; STC = supragingival tooth cleaning; TNF = tumor necrosis factor.
† Time point selected for analysis.

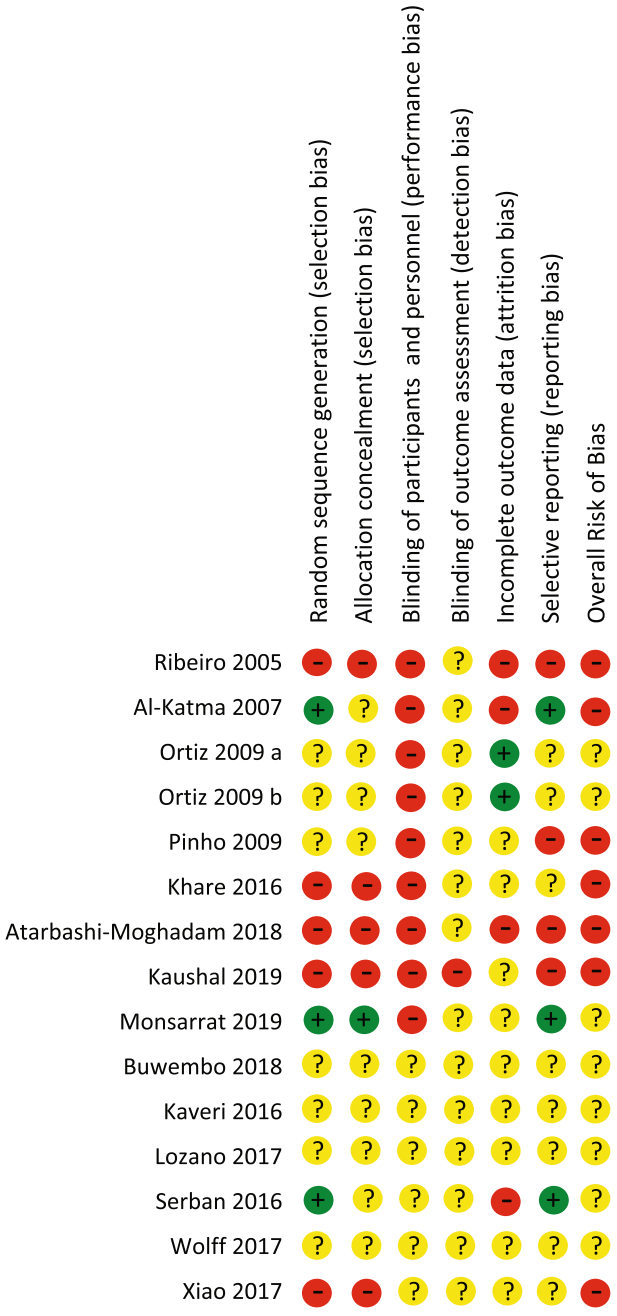

Figure 2. Review of authors' judgments about each risk of bias item for each included study. + indicates low risk of bias; – indicates high risk of bias; ? indicates unclear risk of bias.

enrolled patients, 80% were females (median of individual studies [IQR 60–91%]). The duration of the studies varied from 6 to 24 weeks. Most of the studies compared NSPT plus oral hygiene instructions with only oral hygiene instructions or absence of treatment, whereas some studies added antibiotics to the treatment group.

Risk of bias of individual studies. The measures of risk of bias were assessed as described above (24), and graphic representations of potential bias were computed (Figure 2). All 14 studies were at high risk for at least 1 of the above-mentioned domains: 3 (21%) (32,38,39) adequately generated their randomization sequence; 1 (7%) (38) adequately concealed allocation; and 0 studies (0%) blinded participants/personnel or outcome assessors. Two (14%) comparisons (33) had low risk of missing outcome data; 3 comparisons (21%) (32,38,39) presented low risk for reporting bias; 7 comparisons (47%) (31,32,34–37,44) presented an overall high risk of bias; and 8 comparisons (53%) (33,38–43) presented unclear overall risk.

Evidence synthesis. *Disease activity.* Of the 14 included studies, 1 (31) did not collect disease activity scores; 12 (32–36,38–44) collected or will collect disease activity scores as the DAS28; and 1 study (37) collected or will collect disease activity scores as the SDAI. One of the ongoing studies (40) will measure the DAS28 score using CRP level, and 8 of the studies (32–36,38,39,43) measured or will measure the DAS28 score using ESR. In 3 of the studies (41,42,44), it is not clear if they used or will use CRP level or ESR.

As illustrated in Figure 3, of the 13 studies that collected disease activity scores, only 7 studies could be included in the meta-analysis due to lack of data from the remaining trials: 4 (40–43) are ongoing studies with no data yet available; one (44) is a poster abstract; and one (34) does not show control group data. Of the 7 studies (8 comparisons) included in the meta-analysis, 6 studies (7 comparisons) (32,33,35–37,39) showed a reduction in disease activity scores, with statistically significant results in 5 of the comparisons.

When we pooled the results of all 7 studies that reported this outcome, we observed a statistically significant SMD of –0.88 (95% CI –1.38, –0.38) in disease activity reduction. The forest plot describes the effect of nonsurgical periodontal treatment on disease activity in terms of mean reduction from baseline; it also shows SD

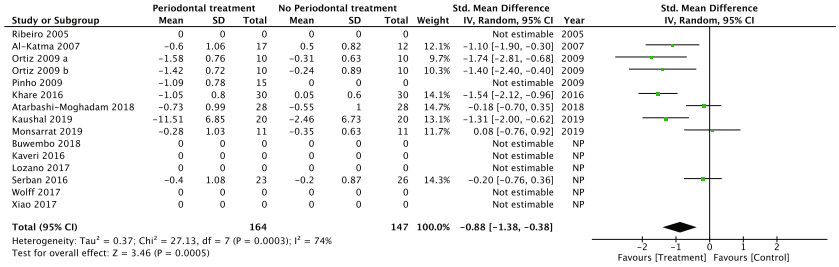

Figure 3. Forest plot of standardized mean difference (SMD) between periodontal treatment and no treatment (control) groups for disease activity. 95% CI = 95% confidence interval. Squares represent the individual SMD of each study. Horizontal lines indicate 95% CIs. Diamonds represent the pooled SMD and its 95% CIs.

Table 2. GRADE evidence profile of periodontal treatment versus controls for patients with RA in both RCT and qRCT studies*

No. of studies (n = 9)	Study design, no.		No. of patients (n = 388)			Serious risk of bias†	I^2, %‡	Serious indirectness or imprecision§	Relative effect (95% CI)	Absolute effect		Quality of evidence
	RCT	qRCT	Periodontal treatment	No periodontal treatment	Total					Control group estimate	Absolute effect (95% CI)	
Disease activity¶ 7	4	3	164	147	311	Yes	74	Yes	SMD 0.88 lower (1.38 lower to 0.38 lower)	4.1 ± 0.5#	3.6 (0.19, 0.69)	Very low
Systemic inflammation markers** 7	5	2	136	134	270	Yes	71	Yes	SMD 0.66 lower (1.14 lower to 0.18 lower)	29.8 ± 20.5#	19.7 (3.69, 23.37)	Very low
Response criteria†† 1	1	0	11	11	22	Yes	NA	Yes	RR 0.57 (0.23, 1.41)	7/11 (63.6)‡‡	362 fewer per 1,000 (from 896 fewer to 146 more)§§	Low
Life impact¶¶ 5	4	1	112	85	197	Yes	34	Yes	SMD 0.49 lower (0.79 lower to 0.18 lower)	43.2 ± 20.7#	21.2 (3.73, 15.35)	Very low
Attrition## 9	5	4	199 (77)	189 (77)	388 (154)	Yes	75	Yes	RR 1.56 (0.23, 10.38)	11/189 (5.8)‡‡	90 more per 1,000 (from 490 fewer to 1,000 more)§§	Very low

* 95% CI = 95% confidence interval; DAS28 = Disease Activity Score in 28 joints; DAS28-ESR = DAS28 using the erythrocyte sedimentation rate; GRADE = Grading of Recommendations Assessment, Development and Evaluation; NA = not applicable; qRCT = quasi-randomized clinical trial; RA = rheumatoid arthritis; RCT = randomized clinical trial; RR = risk ratio; SMD = standardized mean difference.
† Assessed using a modified Cochrane risk of bias instrument.
‡ An I^2 value between 50% and 90% may represent substantial heterogeneity.
§ Indirectness results if the intervention, patients, or outcomes are different from the research question under investigation. Imprecision refers to situations in which the 95% CI includes both benefit and harm.
¶ Assessed with the DAS28 and the Simplified Disease Activity Index.
# Values are the mean ± SD. In order to be able to interpret the overall SMD values for clinical efficacy, we present an algorithm (SMD into % improvement = SMD × [SD/mean]) based on the descriptive statistics published by Canhão et al (55). They assessed and evaluated data extracted from the METEOR database, a multinational collaboration on RA, from 2008 until 2013. The database provides data on patient and physician-reported outcome measures in RA. The mean ± SD DAS28-ESR score in the Portuguese population with moderate RA was 4.1 ± 0.5, whereas the mean ± SD ESR was 29.8 ± 20.5 mm/hour, and the mean ± SD patient global assessment of disease activity, assessed on a visual analog scale of 0–100 mm, was 43.2 ± 20.7. These data allow a generalized interpretation into a mean improvement expressed in percentage compared with placebo alone.
** Assessed with C-reactive protein level and ESR.
†† Assessed with the American College of Rheumatology criteria for 20% improvement in disease activity.
‡‡ Values are the no./total no. (%). Calculated absolute effect as the no. of subjects per 1,000.
§§ Control group risk from the studies included in this meta-analysis.
¶¶ Assessed with the patient global assessment of disease activity, pain and disability scales.
Assessed with the no. of starters (no. of completers).

as a comparison between treatment and control groups. The results were highly heterogeneous (I^2 = 74%), indicating that different treatment effects were observed across studies.

Systemic inflammation. Of the 14 studies in the analysis, 13 collected or will collect data related to this outcome domain. Of those 13 studies, 5 collected or will collect systemic inflammation data as ESR levels (31–33,36,43), 1 study will collect data as CRP levels (37), and 7 collected or will collect both ESR levels and CRP levels (34,35,38–41,44).

As can be seen in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24622, data were recorded and available for analysis in only 7 of the 13 data-collecting studies (32–35,37–39) (8 comparisons) for various reasons: 1 study (31)

Table 3. Results of the stratified meta-analyses*

Variable	Comparisons, no.	ES (95% CI)	Tau^2	Heterogeneity explained, %
All trials	8	−0.91 (−1.43, −0.39)	0.412	
Design			0.463	−12
RCT	5	−0.83 (−1.50, −0.16)		
qRCT	3	−1.07 (−1.99, −0.15)		
Intervention			0.270	34
NSPT	7	−1.08 (−1.57, −0.60)		
NSPT + AB	1	0.08 (−1.07, 1.23)		
Comparator			0.321	22
AT	6	−1.14 (−1.68, −0.59)		
OHI	2	−0.19 (−1.16, 0.78)		
RA definition			0.337	18
ACR	6	−0.72 (−1.27, −0.16)		
Unclear	2	−1.55 (−2.55, −0.55)		
PD definition			0.385	7
Advanced	6	−1.08 (−1.67, −0.49)		
Light	2	−0.45 (−1.43, 0.53)		
RA medication†			0.095	77
Stable	5	−0.43 (−0.86, 0.00)		
Unclear	3	−1.57 (−2.10, −1.04)		
Random sequence generation‡			0.172	58
Low	3	−0.34 (−0.93, 0.25)		
High	3	−1.13 (−1.81, −0.44)		
Unclear	2	−1.59 (−2.37, −0.80)		
Allocation concealment‡			0.344	16
Low	1	0.08 (−1.18, 1.34)		
High	3	−1.09 (−1.92, −0.25)		
Unclear	4	−1.08 (−1.76, −0.40)		
Blinding of participants and personnel§			0.394	4
Low	0	NA		
High	7	−1.03 (−1.58, −0.48)		
Unclear	1	−0.20 (−1.55, 1.15)		
Blinding of outcome assessment¶			0.462	−12
Low	0	NA		
High	1	−1.31 (−2.98, 0.36)		
Unclear	7	−0.87 (−1.44, −0.29)		
Incomplete outcome data#			0.362	12
Low	2	−1.58 (−2.57, −0.59)		
High	3	−0.50 (−1.32, 0.33)		
Unclear	3	−0.87 (−1.66, −0.08)		
Selective reporting**			0.080	81
Low	3	−0.31 (−0.79, 0.17)		
High	2	−0.77 (−1.60, 0.05)		
Unclear	3	−1.57 (−2.08, −1.06)		

* 95% CI = 95% confidence interval; AB = antibiotics; ACR = American College of Rheumatology; AT = absence of treatment; ES = effect size; NA = not applicable; NSPT = nonsurgical periodontal treatment; OHI = oral hygiene instructions; PD = pocket depth; qRCT = quasi-randomized clinical trial; RA = rheumatoid arthritis; RCT = randomized clinical trial.
† Test for interaction: $P = 0.0011$.
‡ Selection bias.
§ Performance bias.
¶ Detection bias.
# Attrition bias.
** Reporting bias.

reported ESR as the percentage of participants who have ≥28 mm/hour, which impeded meta-analysis; 1 study (36) did not report ESR-level data, although they were allegedly collected; 3 studies are ongoing (40,41,43); and 1 is a poster abstract (44). A decrease in systemic inflammation was observed in 6 of the 8 comparisons in RA patients undergoing periodontal treatment; 4 of these comparisons with the control group yielded statistically significant decreases (32,33,35). In contrast, the studies of Monsarrat et al (38) and Serban (39) showed an increase (although not statistically significant) in systemic inflammation after nonsurgical periodontal treatment.

The overall SMD between the treatment and nontreatment groups was 0.66 (95% CI –1.14, –0.18), indicating that systemic inflammation is significantly reduced in RA patients following periodontal treatment. The results were highly heterogeneous (I^2 = 71%).

Response criteria. Of the 14 studies included in our analysis, 1 study (38) reported response criteria data as ACR20 responders. According to their research plan and methodology, one study (39) collected ACR20 responders, but no data were available or reported. The proportion achieving the ACR20 response was higher for the control group compared to periodontal treatment group (RR 0.57 [95% CI 0.23, 1.41]). The effect of the intervention in this particular study is presented in Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24622.

Life impact. Of the 14 included studies, 11 collected or will collect life impact outcome measures. Of those 11 studies, 1 study (37) collected but did not report data; 2 ongoing studies (40,41) will collect data regarding the PtGA; and 1 study (44) reported an improvement in patients who received periodontal therapy compared with controls, although no data were available for the meta-analysis. Also, one study (36) collected pain data that were not available, and one study (34) did not show data regarding the comparator group.

As illustrated in Supplementary Figure 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24622, only 5 studies (6 comparisons) could be used in our meta-analysis. Of those 5 studies, 4 (32,33,38,39) reported PtGA, and 1 study (31) reported life impact as disability. When we pooled the results of all 5 studies (6 comparisons) that reported life impact measurements, we observed a statistically significant SMD of –0.49 (95% CI –0.79, –0.18) life impact reduction. The results were moderately heterogeneous (I^2 = 34%).

Attrition. As illustrated in Supplementary Figure 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24622, of the 9 published studies included, 3 studies reported dropouts. The attrition rate appeared higher for the treatment group (RR 1.56 [95% CI 0.23, 10.38]).

Summary of findings and exploring reasons for heterogeneity. Evidence was qualified using GRADE for both RCTs and qRCTs. In general, very low–quality evidence shows that periodontal treatment may be favorable when compared to no treatment, usual care, or placebo in reducing disease activity, systemic inflammation, and life impact (Table 2).

The prespecified stratified analyses performed with more advanced models are presented in Table 3. Only covariates reducing the variation across studies (significant decrease in T^2) were considered potentially relevant. These analyses supported the notion that stability in RA medication itself could be an important factor to the inconsistency seen across studies (effect size –0.43 versus –1.57, respectively, for RA medication stability versus unclear), with a corresponding 77% reduction in heterogeneity (i.e., T^2, changed from 0.412 to 0.095). For intervention, distinguishing between nonsurgical periodontal treatment alone versus nonsurgical periodontal treatment with antibiotics, a trend was observed indicating a potential interaction between strata (–1.08 versus 0.08, respectively), with a reduction in heterogeneity of 34%. Also, when analyzing comparator (absence of treatment versus oral hygiene instructions with an effect size of –1.14 versus –0.19, respectively), we can observe a 22% reduction in heterogeneity (T^2, changed from 0.412 to 0.321). Regarding RA definition (ACR definition versus unclear with an effect size of –0.72 versus –1.55, respectively), it explains 18% of heterogeneity across studies. In terms of periodontitis severity (advanced versus light cases with an effect size of –1.08 versus –0.45, respectively), a minor reduction of heterogeneity (7%) may be observed. The risk of bias assessment clearly showed that selection and reporting bias were associated with relevant reductions in heterogeneity (58% and 81%, respectively) due to poor study quality.

DISCUSSION

The studies included in this review and meta-analysis used validated composite indices of disease activity: the DAS28 (with ESR or CRP) or the SDAI (48,49). Unfortunately, they did not always clearly specify which type of DAS28 was used, which is crucial (50). This outcome has been the object of 2 previous reviews involving 3 studies, and 1 meta-analysis involving 2 studies (9,10). Both reviews identified a decrease in disease activity, which only reached statistical significance in the work by Calderaro et al (9). Of all 8 individual comparisons included in our meta-analysis, 5 revealed statistically significant differences in favor of periodontal treatment, and 3 showed nonsignificant differences. The study by Monsarrat et al (38) was the only one to suggest an increase in disease activity in association with treatment. In their study, mean probing depth was indicative of nonactive periodontitis, and gingival bleeding, a marker of inflammation, was lower than in other studies. Such findings may be an effect of concomitant abatacept, taken by 17 of the 22 patients (51). It is therefore plausible that the margin for improvement in periodontal inflammation was insufficient to identify an effect on RA disease-activity measures.

In order to further explore the impact of periodontal treatment on systemic inflammation, we analyzed data regarding inflammatory markers (ESR and CRP level). We concluded that a significant decrease of systemic inflammatory markers might be observed after periodontal treatment in RA patients. A detailed analysis of 2 studies (38,39) that reported a nonsignificant increase after periodontal treatment suggests that this increase may be linked to lower levels of periodontal inflammation, as previously described.

We observed a significant reduction of life impact measures in patients receiving periodontal treatment. Our analysis reveals promising results, albeit with a low-quality evidence. Another outcome of interest in this review was response criteria, but only 1 study, with several limitations, reported this outcome (38).

Finally, of the 14 studies included in this review and meta-analysis, only 3 studies reported dropouts, but they probably also occurred in the remaining ones. The higher attrition rates in the treatment groups, whatever their cause, questions the validity of the results and may cause treatment effectiveness to be overestimated.

Although data presented in this meta-analysis suggest an overall significant positive effect for treating periodontitis in RA patients, it should be emphasized that the quality of evidence is very low due to high risk of bias. The studies available for inclusion are mainly not truly randomized or do not clearly describe the randomization and/or allocation concealment method, which may lead to selection bias. Actually, our meta-regression analyses highlight selection bias as a considerable source of heterogeneity between studies. The sample size of the published studies is inferior to the sample size described in the corresponding available clinical trial registries. This lower sample size probably reflects difficulties in recruiting and maintaining patients in the studies, and it strongly affects the statistical power to identify true differences that may exist. Using the effect of treatment on disease activity obtained in this meta-analysis, we estimate that to detect an effect size of 0.88 with at least 80% power, a total sample size of at least 44 individuals would be needed, randomized 1:1. If we aim for a more robust power in the design phase, a sample size of >56 patients would allow for a statistical power of >90%.

The time frame used for follow-up among studies (6 weeks to 6 months) seems sufficient to improve periodontal condition and, probably, to demonstrate a systemic impact (52); it also avoids the ethical issues of prolonging the control condition for longer periods.

A lack of standardized case definition for periodontitis is a paramount observation in our review and probably makes a decisive contribution to the observed heterogeneity of results. In fact, and as expected, best effect size is observed among studies with advanced cases of periodontitis. Treating advanced periodontitis rather than light cases can be expected to result in a greater effect upon disease activity and systemic inflammation. We therefore advise that studies aiming at proving the concept should preferably include cases of moderate and severe periodontitis.

Periodontal intervention must be clearly described and designed. In our review, periodontal treatment encompassed not only mechanical debridement but also adjunctive therapies, such as antibiotics, which are not consensually recommended (53,54). Only 1 study used antibiotics in the intervention arm in a population with light cases of periodontitis. Our meta-regression analysis cannot support the use of antibiotics in this context. Further, the medication being taken by patients included in the study must be carefully described, kept stable throughout the study period, and be as similar as possible between treatment and control groups. Actually, changes in RA medication during the study explain a significant amount of inconsistency. We hypothesize that the large effect size in disease activity observed in studies where RA medication stability is unclear may be explained by unreported changes in medication in addition to the effect of periodontal treatment itself.

Future trials should strictly adopt current classification criteria for RA (18) and validated outcome measures for disease activity, response, and impact. In fact, the best effect sizes were observed in studies lacking a clear definition of RA, which actually puts the whole methodology under question. We recommend that the outcomes of interest should be standardized according to OMERACT guidance (20,21), which would facilitate comparisons of outcomes across studies and provide the best estimates of benefit and safety of the therapeutic intervention across differing patient populations. We believe that RA-related measurements ideally should be taken by experienced professionals blind to treatment allocation (blinding participants to treatment allocation is impractical for obvious reasons). Objective means of assessing inflammation, such as ultrasound, also should be considered.

In conclusion, our results highlight the effect of periodontal treatment on RA outcomes. Although available evidence describes a possible effect of periodontal treatment on RA, the evidence has low quality, with substantial heterogeneity. Critical analysis of published reports using an evidence-based research approach would enable specific recommendations for RA patients and help guide future research.

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. daSilva 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. Silva, Baptista, Santiago, Lund, Tarp, daSilva, Christensen.

Acquisition of data. Silva, Costa, Lund, Tarp.

Analysis and interpretation of data. Silva, daSilva, Christensen.

REFERENCES

1. Kassebaum N, Bernabé E, Dahiya M, Bhandari B, Murray C, Marcenes W. Global burden of severe periodontitis in 1990–2010: a systematic review and meta-regression. J Dent Res 2014;93: 1045–53.

2. Linden GJ, Lyons A, Scannapieco FA. Periodontal systemic associations: review of the evidence. J Clin Periodontol 2013;40 Suppl 14: S8–19.
3. Simpson TC, Needleman I, Wild SH, Moles DR, Mills EJ. Treatment of periodontal disease for glycaemic control in people with diabetes. Aust Dent J 2010;55:472–4.
4. Li C, Lv Z, Shi Z, Zhu Y, Wu Y, Li L, et al. Periodontal therapy for the management of cardiovascular disease in patients with chronic periodontitis. Cochrane Database Syst Rev 2017;11:CD009197.
5. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. Arthritis Res 2002;4 Suppl 3:S265–72.
6. Mikuls TR, Payne JB, Deane KD, Thiele GM. Autoimmunity of the lung and oral mucosa in a multisystem inflammatory disease: the spark that lights the fire in rheumatoid arthritis? J Allergy Clin Immunol 2016;137: 28–34.
7. Bartold PM, Lopez-Oliva I. Periodontitis and rheumatoid arthritis: an update 2012–2017. Periodontol 2000 2020;83:189–212.
8. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis 2016;75:3–15.
9. Calderaro DC, Corrêa JD, Ferreira GA, Barbosa IG, Martins CC, Silva TA, et al. Influence of periodontal treatment on rheumatoid arthritis: a systematic review and meta-analysis. Rev Bras Reumatol Engl Ed 2017;57:238–44.
10. Kaur S, Bright R, Proudman SM, Bartold PM. Does periodontal treatment influence clinical and biochemical measures for rheumatoid arthritis? A systematic review and meta-analysis. Semin Arthritis Rheum 2014;44:113–22.
11. Lund H, Juhl C, Christensen R. Systematic reviews and research waste. Lancet 2016;387:123–4.
12. Lund H, Brunnhuber K, Juhl C, Robinson K, Leenaars M, Dorch BF, et al. Towards evidence based research. BMJ 2016;355:i5440.
13. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;349:g7647.
14. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;6:e1000100.
15. Ghogomu EA, Maxwell LJ, Buchbinder R, Rader T, Pardo JP, Johnston RV, et al. Updated method guidelines for Cochrane musculoskeletal group systematic reviews and metaanalyses. J Rheumatol 2014;41:194–205.
16. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016;75:40–6.
17. Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898.
18. 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.
19. Armitage GC. Periodontal diagnoses and classification of periodontal diseases. Periodontol 2000 2004;34:9–21.
20. Tugwell P, Boers M. Developing consensus on preliminary core efficacy endpoints for rheumatoid arthritis clinical trials. OMERACT committee. J Rheumatol 1993;20:555–6.
21. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum 1993;36:729–40.
22. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions. John Wiley & Sons; 2011.
23. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.
24. Higgins JP, Green S, editors. Cochrane handbook for systematic reviews of interventions version 5.1.0. The Cochrane Collaboration; 2011. URL: www.cochrane-handbook.org.
25. Da Costa BR, Nüesch E, Rutjes AW, Johnston BC, Reichenbach S, Trelle S, et al. Combining follow-up and change data is valid in meta-analyses of continuous outcomes: a meta-epidemiological study. J Clin Epidemiol 2013;66:847–55.
26. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.
27. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.
28. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999;18:2693–708.
29. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1559–73.
30. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383–94.
31. Ribeiro J, Leão A, Novaes AB. Periodontal infection as a possible severity factor for rheumatoid arthritis. J Clin Periodontol 2005;32: 412–6.
32. Al-Katma MK, Bissada NF, Bordeaux JM, Sue J, Askari AD. Control of periodontal infection reduces the severity of active rheumatoid arthritis. J Clin Rheumatol 2007;13:134–7.
33. Ortiz P, Bissada NF, Palomo L, Han YW, Al-Zahrani MS, Panneerselvam A, et al. Periodontal therapy reduces the severity of active rheumatoid arthritis in patients treated with or without tumor necrosis factor inhibitors. J Periodontol 2009;80:535–40.
34. Pinho MN, Oliveira RD, Novaes AB Jr, Voltarelli JC. Relationship between periodontitis and rheumatoid arthritis and the effect of nonsurgical periodontal treatment. Braz Dent J 2009;20:355–64.
35. Khare N, Vanza B, Sagar D, Saurav K, Chauhan R, Mishra S. Nonsurgical periodontal therapy decreases the severity of rheumatoid arthritis: a case-control study. J Contemp Dent Pract 2016;17:484–8.
36. Atarbashi-Moghadam F, Maybodi FR, Dehghan A, Ardakani AH. Effect of non-surgical periodontal treatment on clinical signs of rheumatoid arthritis. J Periodontol Implant Dent 2018;10.
37. Kaushal S, Singh AK, Lal N, Das SK, Mahdi AA. Effect of periodontal therapy on disease activity in patients of rheumatoid arthritis with chronic periodontitis. J Oral Biol Craniofac Res 2019;9:128–32.
38. Monsarrat P, de Grado GF, Constantin A, Willmann C, Nabet C, Sixou M, et al. The effect of periodontal treatment on patients with rheumatoid arthritis: the ESPERA randomised controlled trial. Joint Bone Spine 2019;86:600–9.
39. Serban ST. Outcomes of Periodontal Treatment in Patients With Rheumatoid Arthritis (OPERA): quantitative and qualitative results of a pilot randomized controlled trial [dissertation]. University of Birmingham; 2016.
40. Dra Beatriz Lozano-Hospital Universitario de Canarias, sponsor. Periodontitis treatment in patients with reumatoid arthritis and high clinical activity in the randomized clinical trial. EudraCT Number: 2017-003259-40; 2017.
41. University Hospital Heidelberg, sponsor. Dental prophylaxis and rheumatoid arthritis (PREPARA II). ClinicalTrials.gov identifier: NCT03087240; 2017.

42. Institute of Rheumatology Madras Medical College, sponsor. Effect of non-surgical periodontal therapy on chronic periodontitis and rheumatoid arthritis and a correlation of anti-citrullinated protein antibody (ACPA) levels in South Indian population. Clinical Trials Registry India number: CTRI/2016/03/006751; 2016.
43. Makerere University, sponsor. RCT: evaluating the effect of oral care for periodontitis in rheumatoid arthritis patients in Uganda. ClinicalTrials.gov identifier: NCT03513263; 2017.
44. Xiao F, Zhang P, Li X, Mou Y, Chen H, Cai Y. AB0266 effects of periodontal basic treatment on periodontal condition, clinical response and serum inflammatory parameters in rheumatoid arthritis (RA) patients with moderate to severe periodontitis [abstract]. Ann Rheum Dis 2017;76:1141.
45. Birjand University of Medical Sciences, sponsor. Effect of non surgical periodontal treatment on the severity of rheumatoid arthritis (RA). Iranian Registry of Clinical Trials number: IRCT20180613040084N1; 2019.
46. Government Dental College and Research Institute Bangalore, sponsor. Evaluating the influence of non surgical periodontal treatment on the clinical and biochemical measures of rheumatoid arthritis: a randomized controlled study. Clinical Trials Registry India number: CTRI/2020/03/023692; 2020.
47. Buwembo W, Munabi IG, Kaddumukasa M, Kiryowa H, Mbabali M, Nankya E, et al. Non-surgical oral hygiene interventions on disease activity of Rheumatoid arthritis patients with periodontitis: a randomized controlled trial. J Dent Res Dent Clin Dent Prospects 2020;14:26.
48. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44–8.
49. Vander Cruyssen B, van Looy S, Wyns B, Westhovens R, Durez P, van den Bosch F, et al. DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: validation of the DAS28 score in patients under infliximab treatment. Arthritis Res Ther 2005;7:R1063–71.
50. Fleischmann RM, van der Heijde D, Gardiner PV, Szumski A, Marshall L, Bananis E. DAS28-CRP and DAS28-ESR cut-offs for high disease activity in rheumatoid arthritis are not interchangeable. RMD Open 2017;3:e000382.
51. Kobayashi T, Okada M, Ito S, Kobayashi D, Ishida K, Kojima A, et al. Assessment of interleukin-6 receptor inhibition therapy on periodontal condition in patients with rheumatoid arthritis and chronic periodontitis. J Periodontol 2014;85:57–67.
52. D'Aiuto F, Parkar M, Andreou G, Suvan J, Brett PM, Ready D, et al. Periodontitis and systemic inflammation: control of the local infection is associated with a reduction in serum inflammatory markers. J Dent Res 2004;83:156–60.
53. Feres M, Figueiredo LC, Soares GM, Faveri M. Systemic antibiotics in the treatment of periodontitis. Periodontology 2000 2015;67:131–86.
54. Santos RS, Macedo RF, Souza EA, Soares RS, Feitosa DS, Sarmento CF. The use of systemic antibiotics in the treatment of refractory periodontitis: a systematic review. J Am Dent Assoc 2016;147:577–85.
55. Canhão H, Rodrigues AM, Gregório MJ, Dias SS, Melo Gomes JA, Santos MJ, et al. Common evaluations of disease activity in rheumatoid arthritis reach discordant classifications across different populations. Front Med (Lausanne) 2018;5:40.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1736–1744
DOI 10.1002/acr.24625

Nurse-Led Consultation for Patients With Rheumatoid Arthritis at Low Disease Activity: A Randomized Noninferiority Trial

Suet-Kei Kwok,[1] Lai-Man Tang,[1] Hoi-Lun Tsang,[1] Ho-Yin Chung,[2] Man-Ho Chung,[1] Carmen T. K. Ho,[1] Chak-Sing Lau,[2] and Tommy T. Cheung[3]

Objective. To determine the effectiveness of nurse-led consultations in patients with stable rheumatoid arthritis (RA) in Hong Kong.

Methods. The present work was a single-center, randomized, open-label, noninferiority trial. Patients who had rheumatoid arthritis (RA) with low disease activity (LDA) were randomized at a 1:1 ratio to attend a nurse-led consultation or rheumatologist follow-up visit for 2 years. The primary end point was the proportion of patients whose RA remained at LDA. Secondary end points included the proportion of patients with RA in disease remission and the scores recorded on the Leeds Satisfaction Questionnaire at 2 years, changes from baseline on the Disease Activity Score in 28 joints using the C-reactive protein level (DAS28-CRP), modified Sharp/van der Heijde score (SHS), Health Assessment Questionnaire disability index (HAQ DI), Short Form 36 (SF-36) physical component score, and 19-item Compliance Questionnaire for Rheumatology (CQR-19) score.

Results. Among 280 patients who were randomized equally to either attend nurse-led consultations or rheumatologist follow-up visits, 267 patients completed the study. In the nurse-led consultation and rheumatologist follow-up groups, 92.1% and 91.4% patients, respectively, remained at LDA at 2 years. The 95% confidence intervals (95% CIs) of the adjusted treatment difference were within the predefined noninferiority margin in both the intention-to-treat analysis (95% CI 5.75, 7.15) and the per-protocol analysis (95% CI 1.67, 7.47). Although the changes in DAS28-CRP score over 2 years were significantly different between the 2 treatment groups ($P < 0.001$), there were no significant changes from baseline in SHS, HAQ DI, SF-36 physical component scores, and CQR-19 scores. At the end of the study, more patients expressed satisfaction with nurse-led consultations.

Conclusion. Nurse-led consultations were not inferior to rheumatologist follow-up visits in patients with stable RA.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic inflammatory disease affecting 0.35% of the population in Hong Kong (1,2). Suboptimal control of the disease may lead to permanent joint damage, progressive functional disability, and decreased quality of life (3). In order to minimize the adverse outcomes associated with RA, treatment should be aimed at achieving sustained remission or low disease activity (LDA) for every patient (4).

Conventionally, patients with RA are followed by rheumatologists in an outpatient setting. However, this type of care delivery model can no longer sufficiently accommodate the increasing number of patients in the public health care system. As a result, shared care by rheumatology nurses as an alternative to usual clinic assessment by rheumatologists has been implemented in many Western countries. The cost-effectiveness and satisfaction of this new care delivery model have been confirmed in many previous clinical studies (5–9). Therefore, the role of rheumatology nurses in the management of chronic inflammatory arthritis has been ascertained. According to recommendations by the European Alliance of Associations for Rheumatology (EULAR), nurses should participate in comprehensive disease management

[1]Suet-Kei Kwok, MNurs, Lai-Man Tang, MNurs, Hoi-Lun Tsang, MBBS, Man-Ho Chung, MBBS, Carmen T. K. Ho, MBBS: Queen Mary Hospital, Hong Kong, China; [2]Ho-Yin Chung, MD, Chak-Sing Lau, MD: The University of Hong Kong, Hong Kong, China; [3]Tommy T. Cheung, MBBS: The University of Hong Kong and Hong Kong Sanatorium & Hospital, Hong Kong, China;

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

Address correspondence to Tommy T. Cheung, MBBS, Department of Medicine, Hong Kong Sanatorium & Hospital, 2 Village Road, Hong Kong, China. Email: tcheungt@hku.hk.

Submitted for publication April 10, 2021; accepted in revised form April 15, 2021.

SIGNIFICANCE & INNOVATIONS

- Nurse-led consultation is not inferior to usual rheumatologist follow-up in patients with stable rheumatoid arthritis (RA) in Hong Kong.
- Patients with stable RA were satisfied with nurse-led consultations.
- Nurse-led consultation for patients with stable RA is feasible in the Chinese population.

to control disease activity, reduce symptoms, and improve patient-reported outcomes through nurse-led support, education, and consultations (10).

In recent years, there has been a significant development of rheumatology nursing service in Hong Kong. Rheumatology nurses are trained to evaluate disease activity, monitor adverse drug reactions, and provide education and psychosocial support to patients with RA (11). Although rheumatology nursing service has been established in Hong Kong for more than 15 years, nurse-led consultation for patients with RA has yet to be widely adopted as its feasibility and quality has not been confirmed in the Chinese population. Therefore, we have tailored this shared care model according to the health care system in Hong Kong. The objective of this study is to determine the effectiveness of nurse-led consultations in patients with stable RA in Hong Kong, with focus on the control of disease activity, patient-reported outcomes, and satisfaction.

PATIENTS AND METHODS

Patients ages 18 to 75 years who met the 2010 American College of Rheumatology/EULAR classification criteria for RA with LDA, which is defined as having a Disease Activity Score in 28 joints using the C-reactive protein (DAS28-CRP) level of ≤3.2 (12) for at least 6 months, were eligible to enroll in the present study. All patients received maintenance therapy of conventional synthetic disease modifying antirheumatic drugs (csDMARDs) at a stable dose for at least 3 months prior to randomization. Use of nonsteroidal antiinflammatory drugs or oral prednisolone of ≤7.5 mg per day was permitted.

Patients were excluded if they had any clinically significant laboratory abnormalities; systemic manifestations of RA, including but not limited to interstitial lung disease and systemic vasculitis; and/or a significant cognitive impairment or psychiatric condition that could interfere with the interpretation of study results. Full eligibility criteria are listed in the Supplementary Appendix, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24625.

Study design. The present study was a single center, randomized, open-label, noninferiority trial conducted at Queen Mary Hospital in Hong Kong. Patients were randomized at a 1:1 ratio to one of the treatment groups (i.e., nurse-led consultation versus usual rheumatologist follow-up [control group]). Randomization was performed using a computer-generated random number sequence, and the numbers were placed in sequentially numbered, opaque, and sealed envelopes. Study protocol was reviewed and approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW 15-378). All patients provided written informed consent and the study was conducted in accordance with the Declaration of Helsinki and all applicable regulatory requirements.

Procedures. Patients were required to attend scheduled visits every 4 months. Apart from joint assessment, the Stanford Health Assessment Questionnaire disability index (HAQ DI), Short Form 36 (SF-36) health survey, and Compliance Questionnaire of Rheumatology 19-item (CQR-19) were completed by patients at each scheduled visit. The Leeds Satisfaction Questionnaire (LSQ) was completed at 1 and 2 years. Radiographs of the bilateral hands and feet were taken at baseline, 1 year, and 2 years. The modified Sharp/van der Heijde score (SHS) was used to assess the radiographic progression of RA. All radiographs were reviewed by 2 independent readers in a time-blinded manner.

The treatment goal was to maintain LDA in patients with RA after randomization. Patients randomized to attend nurse-led consultations were scheduled to see a rheumatologist only once a year. All other study visits were conducted by the rheumatology nurses, who were trained to perform joint assessments, evaluate laboratory results, and monitor drug compliance and tolerability. Rheumatology nurses were allowed to titrate the dose of a csDMARD within the predefined maximum dosage according to disease activity and drug-related toxicity. Nurses were required to consult a senior rheumatologist if the patient developed any unexplained symptoms, laboratory abnormalities, or when initiating or switching to a new csDMARD was indicated. Apart from the scheduled study visits, rheumatology nurses were also allowed to arrange extra clinic visits, laboratory tests, and referrals to other allied health workers according to their judgment. Allied health service received by the patients during the study period are shown in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24625. During the yearly rheumatologist visits, rheumatologists performed the same assessments as rheumatology nurses did during the nurse-led consultations.

Patients randomized to receive rheumatologist follow-up were seen in the usual clinic setting. The rheumatology clinic in Queen Mary Hospital is run by rheumatologists, specialist trainees, junior doctors, and interns under the supervision of the in-charge rheumatologist. These providers performed tender and swollen joint counts and reviewed the results of blood tests arranged before the scheduled visits. Apart from titration of csDMARDs, extra clinic visits, laboratory tests, and referrals to other allied health workers could be arranged if necessary. Apart

from group education and counselling sessions, patients receiving usual care were not attended by the rheumatology nurses.

Additionally, independent joint assessors (M-HC and CTKH) who were blinded to treatment allocation performed joint assessments for patients in both treatment groups at 1 year and 2 years. Rheumatology nurses and doctors were blinded to the joint assessments performed by the independent assessors. Only the results obtained by the independent joint assessors were used for analysis at 1 year and 2 years.

Patients were withdrawn from the study if they were pregnant, developed any malignancy or systemic manifestation of RA during the study, or did not reach LDA 6 months after a disease flare despite the use of 3 csDMARDs at their maximum tolerated doses in combination or sequentially (concomitant use of methotrexate, sulphasalazine, and leflunomide was not permitted).

Outcome measures. The primary end point was the proportion of patients who remained at LDA at 2 years. Secondary end points included the proportion of patients in remission as determined by DAS score (DAS28-CRP score of <2.6) and the scores of LSQ at 2 years as well as changes from baseline in DAS28-CRP, SHS, HAQ DI, SF-36 physical component scores, and CQR-19 scores over 2 years. Other outcome measures included the proportion of patients with RA in LDA or remission at 1 year.

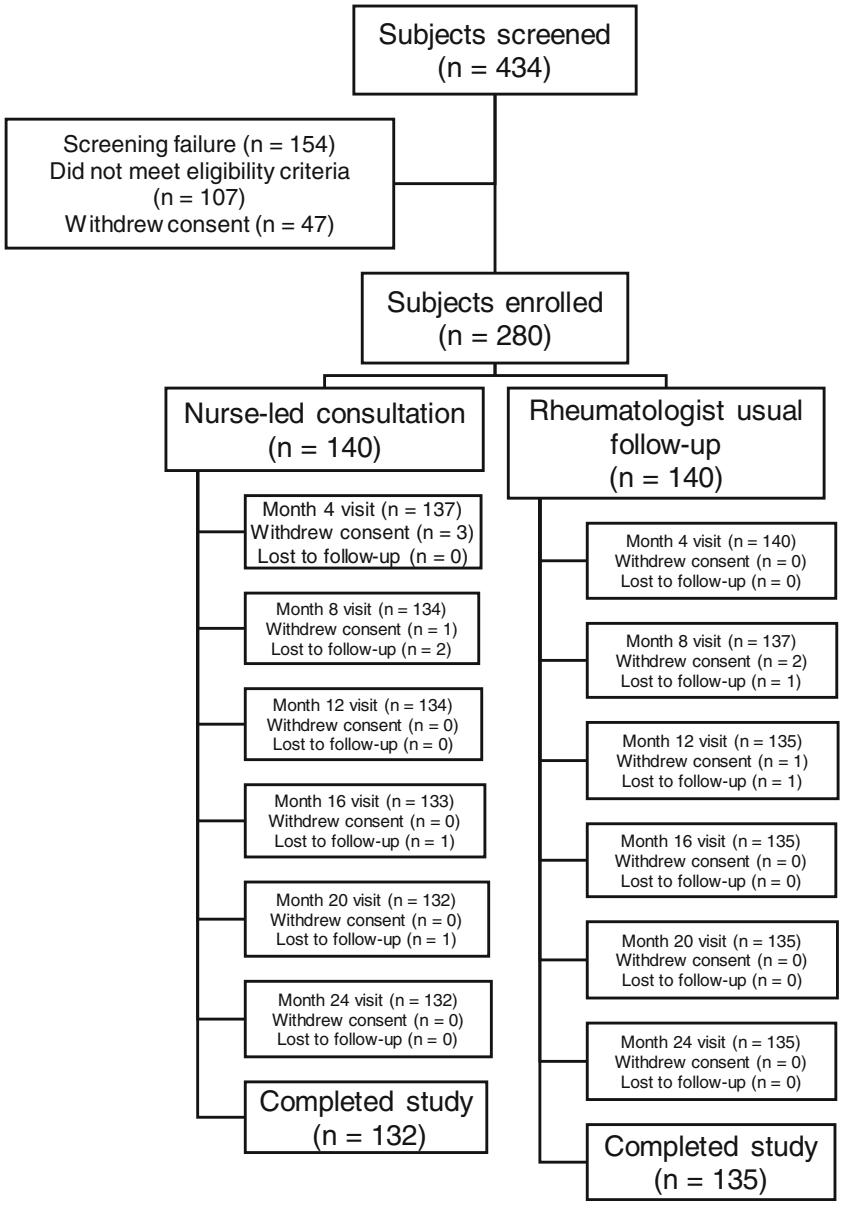

Figure 1. Patient flow diagram.

Sample size calculation. The success rate for nurse-led consultation was assumed to be 85%, and no difference in the proportions of patients with persistent DAS28-CRP scores of ≤3.2 could be observed at 2-year follow-up. This assumption was based on the results of an observational study that described the 5-year disease and patient-reported outcomes of continuous application of a treat-to-target strategy in patients with early RA in daily clinical practice (13). The study showed that the proportion of patients with RA in remission based on DAS28 score and low disease activity increased from 78.9% in the first year to 83.7% in the third year (13). Given a 2-sided alpha level of 0.05 and 90% power, the predefined 15% noninferiority margin required 120 subjects per group. Assuming a dropout rate of 10%, the study required a minimum of 264 subjects for randomization.

Statistical analysis. Analyses of primary and second end points were performed based on intention-to-treat analysis, which included all randomized patients in this study. Patients with missing data for assessment of DAS28-CRP at 2 years or who had withdrawn from the study were considered nonresponders. A sensitivity analysis for the primary end point was performed in patients who completed the study without missing data for assessment of DAS28-CRP at 2 years and major protocol deviations (the per-protocol set [PPS]). According to the study protocol, absence from more than 3 scheduled visits was considered a major protocol deviation, and these patients were excluded from the PPS. The proportions of patients in DAS-defined remission were analysed with the same approach.

In addition, a generalized linear mixed-effects model was used to determine the change in DAS28-CRP, SHS and its component scores, HAQ DI, SF-36 physical component score, and CQR-19 score between the 2 treatment groups over 2 years. All analyses were performed using SPSS, version 22, software based on a significance level of 0.05.

Table 1. Demographic and baseline characteristics of the study subjects*

	Nursing consultation (n = 140)	Rheumatologist follow-up (n = 140)	*P*
Age, years	56.01 ± 9.33	56.16 ± 10.44	0.909
Female sex, %	108 (77.1)	115 (82.1)	0.299
Education level, %			
None	6	11	–
Primary	37	30	0.636
Secondary	74	70	–
Tertiary	23	29	–
Disease duration, years	7.44 ± 4.94	7.32 ± 4.71	0.880
Rheumatoid factor positivity, %†	92 (67.6)	84 (64.1)	0.544
ACPA positivity, %†	85 (73.9)	80 (65.0)	0.138
28 tender joint count score	0.46 ± 1.02	0.57 ± 1.00	0.146
28 swollen joint count score	0.16 ± 0.47	0.22 ± 0.65	0.644
VAS pain (0–100)	18.61 ± 19.78	22.02 ± 20.39	0.120
VAS patient global assessment (0–100)	15.0 ± 18.11	18.0 ± 20.30	0.287
VAS physician global assessment (0–100)	4.0 ± 7.35	5.2 ± 7.91	0.119
ESR, mm/hour	32.63 ± 18.48	34.09 ± 18.98	0.513
CRP, mg/dl	0.60 ± 0.84	0.51 ± 0.33	0.414
DAS28-CRP SCORE	2.02 ± 0.53	2.12 ± 0.55	0.078
Number of csDMARDs	1.54 ± 0.65	1.52 ± 0.70	0.860
Use of methotrexate, %	92 (65.7)	92 (65.7)	1.000
Methotrexate dose, mg/week	12.50 ± 3.52	13.21 ± 3.48	0.173
Use of sulphasalazine	46 (32.9)	49 (35.0)	0.705
Sulphasalazine dose, mg/day	2,054 ± 508.0	1,969 ± 680.1	0.494
Use of leflunomide	9 (6.43)	16 (11.4)	0.142
Leflunomide, mg/day	17.22 ± 4.41	16.88 ± 4.43	0.852
Use of hydroxychloroquine, %	67 (47.9)	56 (40)	0.185
Hydroxychloroquine, mg/day	267.2 ± 89.42	258.9 ± 92.98	0.618
Use of prednisolone, %	4 (2.86)	4 (2.86)	1.000
Prednisolone, mg/day	4.25 ± 1.50	4.63 ± 2.69	0.816
CQR-19 score	66.94 ± 10.32	66.10 ± 9.44	0.657
HAQ DI score	0.41 ± 0.48	0.42 ± 0.61	0.168
Joint space narrowing score	16.82 ± 25.72	17.59 ± 26.74	0.807
Erosion score	10.29 ± 20.89	12.49 ± 26.70	0.450
Total modified Sharp/van der Heijde score	27.11 ± 45.28	30.08 ± 52.36	0.617

* Except where indicated otherwise, values are the mean ± SD. ACPA = anti–citrullinated protein antibody; CRP = C-reactive protein; csDMARDs = conventional synthetic disease-modifying antirheumatic drugs; CQR-19 = 19-item Compliance Questionnaire for Rheumatology; DAS28-CRP = Disease Activity Score in 28 joints using the CRP level; ESR = erythrocyte sedimentation rate; HAQ DI = Health Assessment Questionnaire disability index; VAS = visual analog scale.
† Data were not available for all study subjects at baseline.

RESULTS

From August 2015 to May 2017, 434 patients were screened for the present study. Of these patients, 280 were randomized to receive either nurse-led consultation or usual follow-up by rheumatologists at Queen Mary Hospital. An increase in the DAS28-CRP before randomization was the most common reason for screening failure. Some patients withdrew their consent because nurse-led consultations would be conducted at the Rheumatology Day Centre, which was located in a separate medical facility from Queen Mary Hospital. A total of 267 patients completed the study. No patient fulfilled the withdrawal criteria (Figure 1). Demographic and baseline disease characteristics are

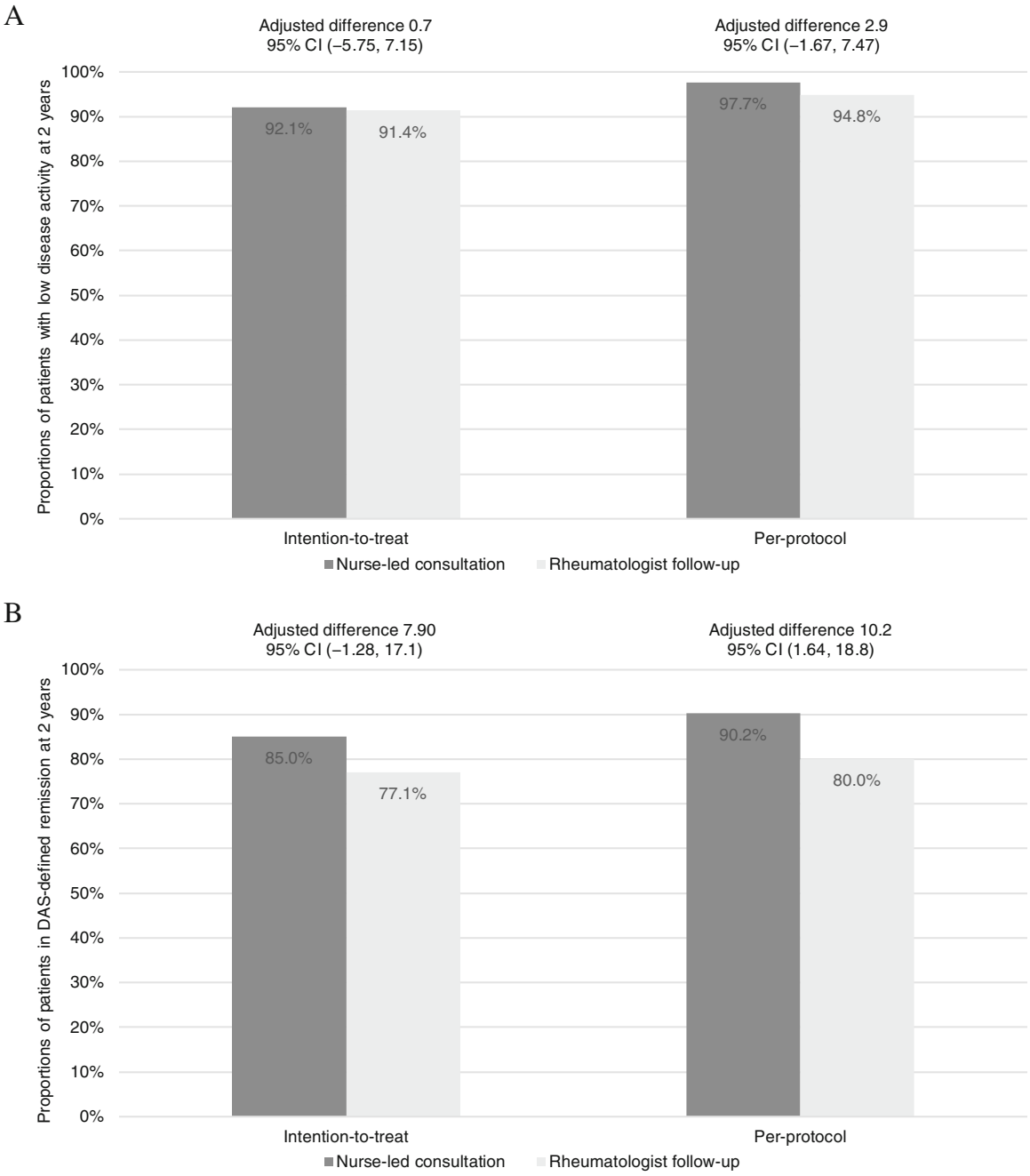

Figure 2. Percentage of study patients who maintained low disease activity or remission of disease activity in intention-to-treat and per-protocol analyses. **A**, Proportions of patients who remained at low disease activity (as defined by a Disease Activity Score in 28 joints using the C-reactive protein level [DAS28-CRP] score of <3.2) between treatment groups at 2 years in the intention-to-treat analysis and per-protocol analysis. **B**, Proportions of patients with disease that remained in DAS-defined remission (DAS28-CRP score of <2.6) between treatment groups at 2 years in the intention-to-treat analysis and per-protocol analysis. 95% CI = 95% confidence interval.

summarized in Table 1. All patients had LDA, and mean ± SD disease duration was 7.38 ± 4.82 years. Most of the study patients used more than 1 csDMARD, and 65% of patients were prescribed methotrexate.

As shown in Figure 2, the proportions of patients who remained in LDA at 2 years in the intention-to-treat analysis were 92.1% for nurse-led consultation and 91.4% for rheumatologist follow-up visits. The 95% confidence intervals (95% CIs) of the adjusted treatment difference were within the predefined noninferiority margin of 15% in both the intention-to-treat analysis (95% CI −5.75, 7.15) and the per-protocol analysis (95% CI −1.67, 7.47), showing that nurse-led consultation was not inferior to usual rheumatologist follow-up in terms of the primary end point. Although the proportion of patients in DAS28-defined remission in the PPS was significantly higher in the nurse-led consultation group than that in the rheumatologist follow-up group (95% CI 1.64, 18.8), the superiority of nurse-led consultation could not be demonstrated in the intention-to-treat analysis (95% CI −1.28, 17.1). The proportions of patients in remission according to other criteria are summarized in Supplementary Figures 2 and 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24625.

During the entire study period, 45 patients required titration of csDMARDs in the nurse-led consultation group compared to 55 patients in the rheumatologist follow-up group. Frequency of dose titration of csDMARDS among study patients over 2 years is shown in Supplementary Figure 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24625. Seventeen extra clinic visits were arranged for patients in the nurse-led consultation group, compared to 28 in the rheumatologist follow-up group. Repeated measures analysis showed a statistically significant interaction between the intervention and time on the changes in DAS28-CRP, $F = 3.97$ (4.24, 1,111.07); $P = 0.003$. Change in DAS28-CRP over 2 years between nurse-led consultation and rheumatologist follow-up were significantly different (mean ± SD 0.243 ± 0.065 [95% CI 0.116, 0.370]; $P < 0.001$); however, the difference remained clinically insignificant (Figure 3).

Additionally, there were no significant differences in the HAQ DI, SHS and its component scores, SF-36 physical component scores, and CQR-19 scores between nurse-led consultation and rheumatologist follow-up over 2 years (Figure 4). SF-36 results among study patients over 2 years are shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24625. As shown in Figure 5, more patients expressed satisfaction with the service provided by the rheumatology nurses compared to the satisfaction expressed with rheumatologist follow-up visits at the end of the study, with an overall satisfaction of 91.1% versus 79.3% ($P = 0.021$).

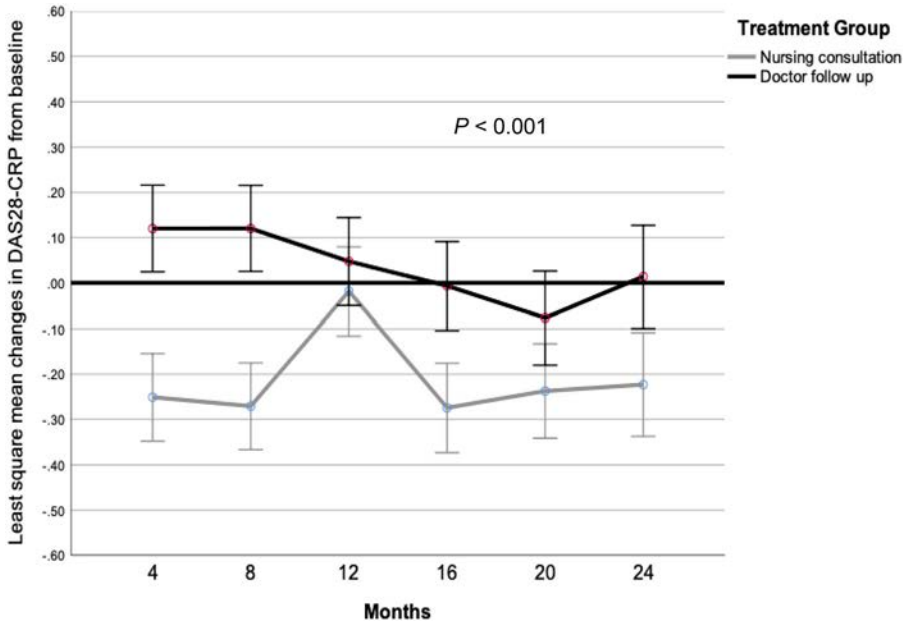

Figure 3. Change in Disease Activity Score in 28 joints using the C-reactive protein level (DAS28-CRP) score from baseline over 2 years among the study patients. A linear mixed-effects model was used to determine the least square means of the above variables and compare the effectiveness of nurse-led consultation versus rheumatologist follow-up. Baseline characteristics including age, sex, disease duration, rheumatoid factor positivity, anti–citrullinated protein antibody positivity, baseline DAS28-CRP, baseline degree of pain measured by visual analog scale, baseline 19-item Compliance Questionnaire for Rheumatology score, and baseline Health Assessment Questionnaire disability index score were the covariates used to construct the model. Circles show the mean change; error bars show the 95% confidence intervals. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24625/abstract.

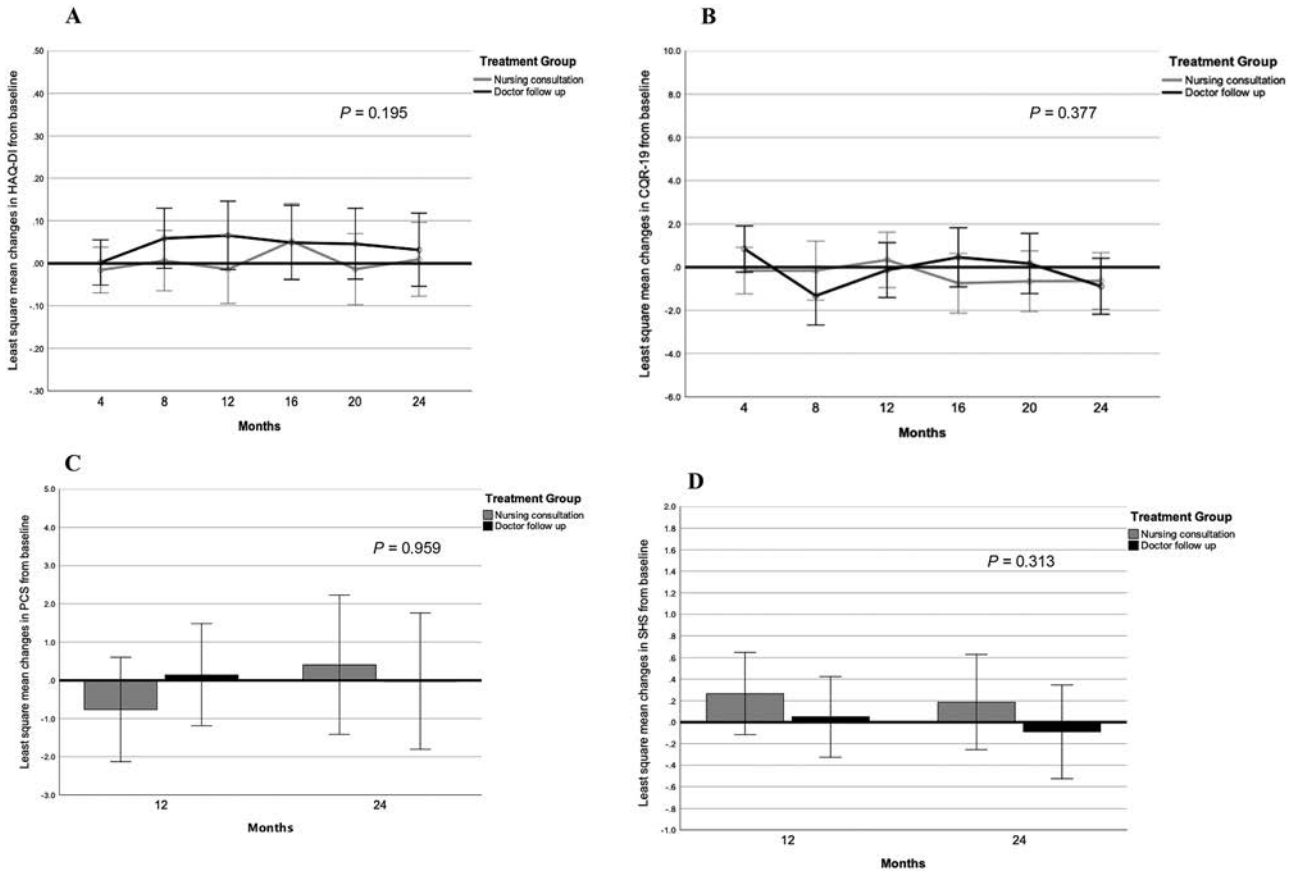

Figure 4. Changes in Health Assessment Questionnaire disability index (HAQ DI) score (**A**), 19-item Compliance Questionnaire for Rheumatology (CQR-19) score (**B**), Short Form 36 physical component scores (PCS) (**C**), and modified Sharp/van der Heijde score (SHS) (**D**) from baseline over 2 years between the nurse-led consultation and rheumatologist follow-up treatment groups. Circles show the mean change; error bars show the 95% confidence intervals.

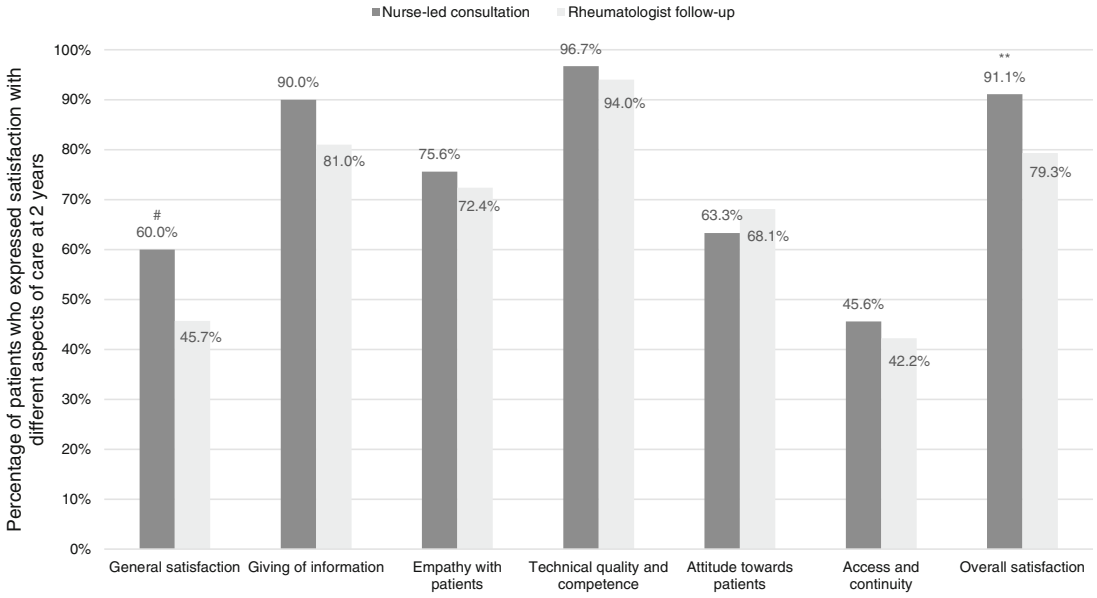

Figure 5. Patient satisfaction at 2 years between treatment groups.

DISCUSSION

The results of the present study showed that nurse-led consultation was not inferior to regular rheumatologist follow-up in RA patients with LDA, which is consistent with findings of other studies (6,7,14,15). The proportions of patients who remained at LDA at 2 years were comparable between the treatment groups. Tight disease control has been advocated for RA patients in order to achieve the lowest possible disease activity and to delay further joint damage and functional decline (16). This approach has been adopted in the protocol used in the present work for the rheumatology nurses with the aim of maintaining a DAS28 score of ≤3.2. The rheumatology nurses were able to measure the disease activity and adjust the dosage of csDMARDs according to the predefined protocol with the aim of achieving tight disease control. In total, 45 patients required dosage adjustments in the nurse-led consultation group within 2 years.

Although the repeated measures analysis showed that the mean changes in DAS from baseline were <0.6 in both treatment groups, patients in the nurse-led consultation group had significantly lower disease activity over 2 years compared to those in the rheumatologist follow-up group. Of note, patients randomized to receive nurse-led consultations had lower baseline disease activity compared to those in the rheumatologist follow-up group (mean ± SD 2.02 ± 0.53 versus 2.12 ± 0.55). In addition, more patients in the nurse-led consultation group were in remission at baseline according to different criteria (Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24625). Therefore, the risk of disease flare was different between the treatment groups (17).

There were no significant differences in radiographic progression, compliance to medications, and patient-reported outcomes as measured by the HAQ DI and SF-36. Most importantly, the present work showed that more patients were satisfied with nurse-led consultations at the end of the study. This is in line with studies conducted in both Western countries (18) as well as in the Chinese population (19).

In Hong Kong, patients with RA are accustomed to scheduled consultations with rheumatologists. Many of them know little about nurse-led consultation and its role in disease management. Due to the cultural difference, the acceptance of nurse-led consultations was also a major concern in implementing this new care model in our public health care system. Although some rheumatology centers in Hong Kong have been providing a similar care model to patients with stable RA, the present work is the first study that confirmed the effectiveness and acceptance of nurse-led consultations among patients with stable RA in Hong Kong. We believe that the results of the present study will encourage the formal implementation of nurse-led consultations and establish a platform for future validation studies. Meanwhile, accredited and structured rheumatology nursing training has been disseminated through network hospitals in Shanghai and Shenzhen to lay the foundation for this new care model in other regions of China.

However, the present study had some potential limitations. Patients treated with biologic or targeted synthetic DMARDs were excluded from the study. As the number of patients treated with these advanced agents is increasing, further studies are needed to evaluate the effectiveness of nursing consultations in this group of patients. In addition, nurse-led consultations were conducted in the Rheumatology Day Centre, which was located in a medical facility separate from Queen Mary Hospital. The change in environment can affect patient satisfaction in terms of technical quality and access. Last, it is important to note that other outcome measures, including work productivity and psychological evaluation, were not included in the present work.

Nurse-led consultation has been well established in many Western countries. The present study showed that nurse-led consultation with an evidence-based management was not inferior to usual rheumatologist follow-up and could be implemented in Chinese patients with stable RA. Further studies are needed to evaluate the effectiveness of nurse-led care in RA patients with higher disease activity and also the efficacy of nurse-led care in other types of inflammatory arthritis.

ACKNOWLEDGMENTS

The authors would like to thank the contributions of research staff members Ken C. Y. Yeung and Dorothy L. Y. Lam, who reviewed radiographs used in the present 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. Cheung had full access to all data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Kwok, Tang, Lau, Cheung.
Acquisition of data. Man-Ho Chung, Ho, Lau, Cheung.
Analysis and interpretation of data. Tsang, Ho-Yin Chung, Lau, Cheung.

REFERENCES

1. Lau E, Symmons D, Bankhead C, MacGregor A, Donnan S, Silman A. Low prevalence of rheumatoid arthritis in the urbanized Chinese of Hong Kong. J Rheumatol 1993;20:1133–7.
2. Xiang YJ, Dai SM. Prevalence of rheumatic diseases and disability in China. Rheumatol Int 2009;29:481–90.
3. Welsing PM, van Gestel AM, Swinkels HL, Kiemeny LA, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum 2001;44:2009–17.
4. Smolen JS, Landewé RB, Bijlsma JW, Burmester GR, Dougados M, Kerschbaumer A, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis 2020;79:685–99.

5. Ndosi M, Lewis M, Hale C, Quinn H, Ryan S, Emery P, et al. A randomised, controlled study of outcome and cost effectiveness for RA patients attending nurse-led rheumatology clinics: study protocol of an ongoing nationwide multi-centre study. Int J Nurs Stud 2011;48: 995–1001.
6. Ndosi M, Lewis M, Hale C, Quinn H, Ryan S, Emery P, et al. The outcome and cost-effectiveness of nurse-led care in people with rheumatoid arthritis: a multicentre randomised controlled trial. Ann Rheum Dis 2014;73:1975–82.
7. Primdahl J, Sorensen J, Horn HC, Petersen H, Hørslev-Petersen H. Shared care or nursing consultations as an alternative to rheumatologist follow-up for rheumatoid arthritis outpatients with low disease activity: patient outcomes from a 2-year, randomised controlled trial. Ann Rheum Dis 2014;73:357–64.
8. Sorensen J, Primdahl J, Horn HC, Hørslev-Petersen H. Shared care or nurse-led consultations as an alternative to rheumatologist follow-up for rheumatoid arthritis (RA) outpatients with stable low disease-activity RA: cost-effectiveness based on a 2-year randomized trial. Scand J Rheumatol 2015;44:13–21.
9. de Thurah A, Esbensen BA, Roelsgaard IK, Frandsen TF, Primdahl J. Efficacy of embedded nurse-led versus conventional physician-led follow-up in rheumatoid arthritis: a systematic review and meta-analysis. RMD Open 2017;3:e000481.
10. Bech B, Primdahl J, van Tubergen A, Voshaar M, Zangi HA, Barbosa L, et al. 2018 update of the EULAR recommendations for the role of the nurse in the management of chronic inflammatory arthritis. Ann Rheum Dis 2020;79:61–68.
11. Kwok WY, Kloppenburg M, Beaart-van de Voorde LJ, Huizinga TW, Vliet Vlieland TP. Role of rheumatology clinical nurse specialists in optimizing management of hand osteoarthritis during daily practice in secondary care: an observational study. J Multidiscip Healthc 2011;4:403–11.
12. Wells G, Becker JC, Teng J, Dougados M, Schiff M, Smolen J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009;68:954–60.
13. Versteeg GA, Steunebrink LM, Vonkeman HE, Ten Klooster PM, van der Bijl AE, van de Laar, MA. Long-term disease and patient-reported outcomes of a continuous treat-to-target approach in patients with early rheumatoid arthritis in daily clinical practice. Clin Rheumatol 2018;37:1189–97.
14. Larsson I, Fridlund B, Arvidsson B, Teleman A, Bergman S. Randomized controlled trial of a nurse-led rheumatology clinic for monitoring biological therapy. J Adv Nurs 2014;70:164–75.
15. Larsson I, Fridlund B, Arvidsson B, et al. A nurse-led rheumatology clinic versus rheumatologist-led clinic in monitoring of patients with chronic inflammatory arthritis undergoing biological therapy: a cost comparison study in a randomised controlled trial. BMC Musculoskeletal Disord 2015;16:354.
16. Bakker MF, Jacobs JW, Verstappen SM, Bijlsma JW. Tight control in the treatment of rheumatoid arthritis: efficacy and feasibility. Ann Rheum Dis 2007;66 Suppl 3:iii56–60.
17. Kuijper TM, Lamers-Karnebeek FB, Jacobs JW, et al. Flare rate in patients with rheumatoid arthritis in low disease activity or remission when tapering or stopping synthetic or biologic DMARD: a systematic review. J Rheumatol 2015;42:2012–22.
18. Koksvik HS, Hagen KB, Rødevand E, Mowinckel P, Kvien TK, Zangi HA. Patient satisfaction with nursing consultations in a rheumatology outpatient clinic: a 21-month randomised controlled trial in patients with inflammatory arthritides. Ann Rheum Dis 2013;72: 836–43.
19. Wang JZ, Zou X, Zhou L, Liu H. Patient satisfaction after nurse-led care in Chinese patients with rheumatoid arthritis: a China study. Biomed Res 2017;28:7.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1745–1750
DOI 10.1002/acr.24835

American College of Rheumatology
Empowering Rheumatology Professionals

REVIEW

Pregnancy Termination in Patients With Rheumatic Diseases

Sophy Mo,[1] Isabelle Malhamé,[2] Megan Schneiderman,[3] and Évelyne Vinet[2]

Rheumatic diseases affect women during their reproductive years. Many women with rheumatic diseases become pregnant; some undergo pregnancy termination. However, there are no official guidelines on pregnancy termination in patients with rheumatic diseases. This work provides an overview of factors that health care professionals must consider. We highlight areas that require further studies and the importance of pregnancy planning and contraception counseling. Patients with rheumatic diseases need to be informed of adverse maternal and fetal outcomes of pregnancy to make informed reproductive decisions and reduce the need for pregnancy terminations.

INTRODUCTION

Rheumatic diseases affect women during their reproductive years (1). Among 206 women who were younger than 45 years with systematic lupus erythematosus (SLE), 86 (42%) were at risk for unplanned pregnancy. Of women with SLE, 59% had not received contraception counseling in the last year, and 53% relied solely on barrier methods (2). Some women with unplanned pregnancy will seek a termination. However, there is currently little information regarding induced abortion in these patients.

We aim to review the literature to guide physicians caring for rheumatic disease patients considering pregnancy termination. We present a synthesis of considerations health care professionals must take into account. Ultimately, a patient-centered discussion about the acceptable risks in pregnancy is at the heart of shared decision-making about the best course of action for each patient.

Epidemiology of induced abortions

Every year, millions of women undergo an induced abortion, which is the medical or surgical termination of a pregnancy. Worldwide, half of pregnancies are unintended and half of those end in abortion (3). Induced abortions are most frequent in women ages 20–29 years old (4).

Patients with rheumatic diseases undergo induced abortion at a similar rate as the general population. For patients with rheumatic disease who had experienced a prior pregnancy, 21% and 25% had undergone at least one pregnancy termination based on data from the BVC (Barbara Volcker Center for Women and Rheumatic Disease) and PROMISSE (Predictors of Pregnancy Outcome: Biomarkers in Antiphospholipid Antibody Syndrome and Systemic Lupus Erythematosus) cohorts, respectively (5). These incidences parallel the average rates in North America, where 1 in 4 women will have an abortion over their lifetime (4,6). No complications, hospitalizations, or disease exacerbations associated with a termination were reported in the BVC cohort. Pregnancy termination was recommended for medical reasons in 4% and 1% of pregnancies in the BVC and PROMISSE cohorts, respectively (5).

In a Canadian population–based study, induced abortions among women with SLE were not higher than in the general population, with a standardized incidence ratio (SIR) of 1.10 (95% confidence interval [95% CI] 0.98, 1.24), when accounting for age and calendar time (7). In 2,508 women with SLE, Venne et al observed 293 induced abortions, with an incidence rate of 17.1 induced abortions per 1,000 person-years (95% CI 15.2, 19.2) (7). In the multivariable analysis, Venne et al did not find higher rates of induced abortions in women exposed to teratogenic immunosuppressive agents (rate ratio [RR] 0.37 [95% CI 0.13, 1.07]) or women who received glucocorticoids (RR 0.67 [95% CI 0.39, 1.16]) (7). These results suggest women with SLE do not have a lower rate

Dr. Vinet receives salary support from the Arthritis Society New/Mid Investigator: Stars Career Development Award STAR-19-0597 and Fonds de recherche du Québec Santé (FRSQ) Junior 2 Award 282178.

[1]Sophy Mo, MD: McGill University, Montreal, Quebec, Canada; [2]Isabelle Malhamé, MD, MSc, Évelyne Vinet, MD, PhD: McGill University, McGill University Health Centre, Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Health Centre (RI-MUHC), Montreal, Quebec, Canada; [3]Megan Schneiderman, MD: McGill University and St. Mary's Hospital, Montreal, Quebec, Canada.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24835&file=acr24835-sup-0001-Disclosureform.pdf.

Address correspondence to Évelyne Vinet, MD, PhD, Research Institute of the McGill University Health Centre, Centre for Outcome Research & Evaluation, 1650 Cedar Avenue, Building R, R1-107, Montreal, Quebec, Canada H3G 1A4. Email: evelyne.vinet@mcgill.ca.

Submitted for publication May 11, 2021; accepted in revised form December 9, 2021.

of unplanned pregnancies than the general population. These findings are concerning, as unplanned pregnancies in this patient population could lead to adverse maternal and fetal outcomes.

In a nested case–control study from Quebec, the same research group investigated whether methotrexate (MTX) exposure was a predictor of induced abortions in women with rheumatoid arthritis (RA) (8). Investigators identified 112 cases of induced abortions in women with RA and 5,855 RA controls. Exposure to MTX occurred in 10.7% of women with RA and 21.7% of control subjects. Women exposed to MTX had a lower rate of induced abortions than unexposed women (RR 0.47 [95% CI 0.25, 0.89]). The overall rate of induced abortions observed in women with RA (at the index date) was half the rate of the general population (incidence rate of 17.1 induced abortions per 1,000 person-years [95% CI 15.2, 19.2]). While women with RA have lower rates of induced abortions, given that they may be receiving teratogenic medicine, the rate of unplanned pregnancies still can be improved.

Potential indications in rheumatic diseases

Data on the safety and best practice of induced abortions in patients with rheumatic diseases is sparse. Yet, these patients and their treating physicians may face the eventuality of a pregnancy termination. In addition to personal or social factors, these patients may choose to terminate their pregnancy for unacceptable maternal health risks carried by a pregnancy, possible adverse fetal or maternal pregnancy outcomes, or exposure to teratogenic drugs. Each of these potential indications for induced abortion in rheumatic diseases are detailed in Tables 1, 2, and 3, respectively. However, pregnancy termination discussions are complex and necessitate an understanding of the patient's personal values as well as risks and benefits. In case of pregnancy continuation, a multidisciplinary team should work on optimizing maternal and fetal outcomes.

Induced abortion types

Rheumatologists and other health care professionals counseling patients with rheumatic diseases on reproductive issues should be informed about the different methods of induced abortion. Patients undergo an induced abortion in either the first or second trimester of pregnancy, either medically or surgically. In 2020, the term "medical abortion," was replaced with "medication abortion," to reflect that all abortions are considered medical

Table 1. High-risk cardiovascular and renal complications associated with pregnancy termination in persons with rheumatic diseases*

	Cardiovascular comorbidities	Renal comorbidities
Epidemiology	PH occurs in 5–27% and 0.5–14% of persons with SSc and SLE, respectively (24). Heart failure risk increased in persons with RA and SLE. Prevalence of LVDD is 9% in persons with RA compared to 6% in the general population (25). Valvulopathy (of variable severity) was noted in 18% of persons with SLE. Libman-Sacks endocarditis was noted in 11–74% of persons with SLE (26).	Of outpatient rheumatology patients, 18% reported to have a GFR of ≤60 ml/minute compared to 5% in the general population (27).
Relevant conditions	Modified WHO cardiovascular risk class. Class III includes mechanical valve. Class IV includes PH, severe LVDD with EF of <30–40%, and severe left-sided outflow tract obstruction (28). PH: Due to the vasculature's inability to relax, the cardiac output rise in pregnancy increases pulmonary arterial pressure and could precipitate right-sided heart failure (29).	SLE: Lupus nephritis (30). ANCA vasculitis: Crescentic (pauci-immune) GN (27). SSc: Scleroderma renal crisis (27).
Pregnancy-associated risk	Modified WHO cardiovascular risk class. Class III: Maternal cardiac event rate of 19–27% associated with pregnancy. Class IV: Maternal cardiac event rate of 40–100% associated with pregnancy (28). PH: Maternal mortality risk of 17–33% associated with pregnancy (28). The risk of severe right heart failure persists for at least 72 hours after delivery (31). Avoidance of pregnancy is recommended for patients with raised pulmonary arterial pressure, although patients with very mild PH may wish to get advice from a multidisciplinary team with experience in managing this condition in pregnancy (28).	Step-wise increase in maternal and fetal pregnancy-related risks from chronic kidney disease stage 1 to stages 4–5. Combined outcome of prematurity, need for neonatal intensive care unit, and small for gestational age newborns present in 34% of pregnancies for persons in stage 1 chronic kidney disease and 90% for persons with stage 4 or 5 chronic kidney disease (32). Adverse maternal outcomes: Worsening kidney function, proteinuria, and HTN (33). Adverse fetal outcomes: Prematurity, fetal loss, and intrauterine growth restriction (33).
Management	Modified WHO cardiovascular risk class. Class III: If pregnancy is decided upon, requires rigorous multidisciplinary management. Class IV: Pregnancy contraindicated and termination should be discussed. If decision is made to continue pregnancy, care as for Class III (28).	Patients with severe renal dysfunction may be advised not to conceive, and termination should be discussed if they become pregnant.

* ANCA = antineutrophil cytoplasmic antibodies; EF = ejection fraction; GFR = glomerular filtration rate; GN = glomerulonephritis; HTN = hypertension; LVDD = left ventricular systolic dysfunction; PH = pulmonary HTN; RA = rheumatoid arthritis; SLE = systemic lupus erythematosus; SSc = systemic scleroderma; WHO = World Health Organization.

Table 2. Possible adverse pregnancy outcomes associated with selected rheumatic diseases

Rheumatic disease	Possible adverse pregnancy outcomes
Systemic lupus erythematosus	Increased incidence of placenta-mediated complications (gestational hypertension, preeclampsia, fetal growth restriction, placental abruption, fetal death) and preterm births (34). Pregnancy is associated with a 60% higher rate of disease flare compared to the nonpregnancy period in systemic lupus erythematosus (34).
Systemic scleroderma	Associated with an increased risk of preterm birth, early pregnancy loss, and placenta-mediated complications (35). Renal crisis may complicate pregnancy. It can be life-threatening to both the mother and the fetus (35).
Vasculitic disorders	Associated with increased rates of preterm birth and pregnancy loss (36). Most common pregnancy complication in Takayasu arteritis is hypertension and/or preeclampsia, occurring in 43% of pregnancies (36). Large vessel disease may result in potentially fatal rupture of the vessel wall when exposed to pregnancy-related cardiovascular changes (28).
Antiphospholipid syndrome	High risk of venous and arterial thromboembolic events and placenta-mediated complications including fetal growth restriction, severe preeclampsia and HELLP syndrome (hemolysis, elevated liver enzyme levels, and low platelet levels) (37). High incidence of adverse pregnancy outcomes including fetal loss at or beyond the 10th week of gestation and placenta-mediated complications despite optimal treatment (38).

procedures. We will address the different methods and motivations for choosing a particular method. While one method may be recommended, it is ultimately the patient who decides which method to pursue, as patient choice and autonomy are strong predictors of patient satisfaction.

Medication abortion. First trimester medication abortions are performed by combination regimens of either mifepristone/misoprostol or MTX/misoprostol. Since the approval of the combination regimen of mifepristone/misoprostol for pregnancy termination in 2016 by the US Food and Drug Administration and Health Canada, it has become the most common method for first trimester medication abortions. It consists of the administration of 200 mg of mifepristone, followed by 800 mcg of misoprostol 24 to 48 hours later. The efficacy of medication abortion decreases with increasing gestational age. In pregnancies that are less than 49 days gestation, medication abortion is as effective as surgical abortion. While medication abortion is indicated for use in pregnancies up to until 63 days gestation in Canada, there is strong clinical evidence demonstrating safety and efficacy of resulting in expulsion of all products of conception between 90% and 96% before 70 days gestation (9). In the US, this method is indicated for pregnancies until 70 days gestation.

Reasons for choosing a medication abortion include that it is noninvasive with no need for surgery or anesthesia, can be done at home, and may be perceived as more "natural" (10). However, certain aspects of a medication abortion may deter patients from choosing this method. It involves increased cramping, lacks the immediacy of a surgical abortion, and requires frequent health care visits. In addition to patient preference, relative contraindications to medication abortion include severe anemia, coagulopathy, and/or anticoagulant drug use due to the increased risk of bleeding of this procedure (11). The risk of hemorrhage in medication abortion is 16% compared to 2% of that in surgical abortion (12,13). Second trimester medication abortions are performed with misoprostol regimens to induce labor and delivery and may also include mifepristone as it is thought to reduce the time to pregnancy expulsion.

Medications used in medication abortions each have specific contraindications. MTX is contraindicated in patients with important liver or renal diseases, severe immunodeficiency, and bone marrow suppression (14). There are no data on how to manage the care of patients taking MTX for other therapeutic purposes who wish to undergo medication abortion with a regimen that includes MTX. Mifepristone is contraindicated in patients with uncontrolled asthma as it may trigger bronchospasm and interfere with the efficacy of systemic glucocorticoid therapy used to treat severe asthma. This is also an important consideration for patients with eosinophilic granulomatosis with polyangiitis. Mifepristone is also contraindicated in chronic adrenal failure, which will be further detailed in a subsequent section (11).

Surgical abortion. First trimester surgical abortion is performed by suction dilation and curettage. To reduce the risk of infection, antibiotics are administered pre- and post-procedure. Analgesia usually consists of a combination of oral, intravenous, inhaled, and local (paracervical block) agents. The procedure lasts ~10 minutes, during which the cervix is gradually dilated and uterine contents are aspirated (15).

For second trimester pregnancies, surgical abortion is performed by dilatation and evacuation. Prior to the procedure, mechanical or medical cervical preparation is administered. The procedure is performed with the recipient placed under general anesthesia or deep intravenous sedation and takes 30 minutes to complete on average (15). First and second trimester surgical abortions have efficacy rates of 99% in the complete removal of all products of conception (16).

Motivations for choosing a surgical abortion include that this method is faster, requires fewer visits, may be perceived as simpler or easier by the patient, more effective, associated with less pain, and reduces the risk of requiring subsequent emergency

Table 3. Known teratogenic drugs used for the treatment of rheumatic diseases*

Drugs	Pregnancy contraindications†	Teratogenic effects
MTX	Contraindicated during pregnancy; stop 1–3 months prior to conception (39). Minimal dose resulting in congenital anomalies in studies was 7.5 mg/week. Most congenital anomalies occurred with much higher doses of MTX. Range of exposure was 2.5 mg/day to 100–200 mg biweekly to a single dose of 50 mg/m^2 (40). A prospective cohort study found exposure to MTX increased risk of major birth defects compared to women without autoimmune diseases (41). Based on a systematic review of 101 patients exposed to MTX at doses of 5–25 mg/week, birth defect rates were similar to those of a nonexposed population (42). Teratogenic period is 3–8 weeks post-conception (40).	Congenital anomalies (40): 1. Craniosynostosis 2. Microcephaly 3. Tetralogy of Fallot 4. Facial abnormalities 5. Limb reduction defects 6. Syndactyly Rate of major congenital anomalies in patients with post-conception exposure is 7% (41).
Cyclophosphamide	Contraindicated during pregnancy; stop 3 months prior to conception (39). Highest risk of teratogenicity when used in first trimester. Can sometimes be used for life-threatening conditions in the second or third trimester (39).	Congenital anomalies (43): 1. Growth restriction 2. Craniofacial defects 3. Limb defects Rate of congenital anomalies is 27% (44).
Mycophenolate mofetil	Contraindicated during pregnancy; stop at least 6 weeks prior to conception (39). Dose in mothers of infants with congenital anomalies is between 0.5–2 gm/day (45). Teratogenic period is exposure until week 8 in all pregnancies with malformations consistent with mycophenolate embryopathy (45).	Congenital anomalies (45): 1. Cleft lip and palate 2. Anomalies of distal limbs 3. Defects of eye, ear, brain, heart, esophagus, and kidneys Rate for congenital anomalies in first trimester exposure (0.5–2 gm/day) is 26% (45).
Leflunomide	Contraindicated during pregnancy. Cholestyramine washout recommended if serum concentration detectable. If concentration undetectable, pregnancy loss and birth defect risks are not elevated (39).	Teratogenicity in animals. No significantly increased risk of major malformations in humans, although data are very limited (46).
NSAIDs	Contraindicated in the third trimester (39).	Risk of premature closure of ductus arteriosus (39).

* MTX = methotrexate; NSAIDs = nonsteroidal antiinflammatory drugs.
† The recommendation for pregnancy termination must consider the dosage, duration of exposure, and timing of exposure to the drug. A discussion regarding the risks and benefits of pregnancy termination must be undertaken with the patient considering all the above factors. Shared decision-making on pregnancy termination can then be taken.

surgery in the event of a failed medication abortion (17). Although the complications rate is <1%, there remains a potential risk of infection, incomplete abortion, and damage to the cervix and uterus (18).

Although there are no absolute contraindications to surgical abortion, some medical comorbidities increase the risk of complications. For high-risk patients, physicians should consider the medical setting in which the procedure is performed and choice of anesthesia.

Special considerations regarding abortion in rheumatic diseases

Bleeding risks for patients receiving anticoagulants or with a bleeding diathesis. Some patients with rheumatic diseases present with a higher bleeding risk. For instance, thrombocytopenia commonly occurs in autoimmune diseases such as SLE. Chronic anemia (inflammatory or hemolytic) can also occur (19). Moreover, anticoagulants are used in the treatment of conditions such as antiphospholipid syndrome, thereby increasing bleeding risk (20).

Medication abortion generally leads to more prolonged and heavier bleeding than surgical abortion. An increased risk of bleeding puts the patient at a higher risk of hemorrhage. Moreover, chronic anemia increases patients' risk of needing blood transfusions. A surgical abortion could be recommended in this setting. Delayed bleeding is less likely as surgical abortions are completed prior to the patient leaving the health care facility. In addition, hemodynamic changes are noticed more promptly in a health care setting (21). If a surgical abortion is not chosen, consideration can be given to performing the medication abortion in-hospital.

Moreover, pain management should be tailored to the patient's comorbidities. Nonsteroidal antiinflammatory drugs are a first-line agent to manage pain during medication abortion; however, it should be avoided in patients at risk of bleeding.

Interaction with steroids and/or other medications. Mifepristone is a synthetic steroid. It competitively binds to the glucocorticoid receptor, thereby inactivating it. Following the intake of this drug, patients without pituitary or adrenal disorders can increase their secretion of adrenocorticotropic hormone and cortisol. While there are minimal adverse effects due to

mifepristone (22), chronic adrenal failure is an absolute contraindication to its use as this drug greatly decreases the efficacy of cortisol replacement therapy in these patients. Since adrenal insufficiency impairs compensation for this effect, mifepristone may lead to adrenal crisis (11). As such, long-term glucocorticoid therapy, a common treatment in rheumatic diseases, is a relative contraindication. Medication abortion without mifepristone or surgical abortion is recommended for these patients. Those who still opt for medication abortion with mifepristone would need a dose adjustment such as an increase of their glucocorticoid therapy for 1 week since mifepristone will diminish the effect of the steroids for 3–4 days (13).

Infectious risks for immunosuppressed patients. Infections are a possible complication for medication and surgical abortions. However, the risk of infection following abortion of any type remains <1%, with the risk of infection for medication abortion based on the combined data of 6 prospective studies estimated to be at 0.3% (23). Infections following medication abortions are rarely serious, and most do not require hospitalization. There are no data on the rates of infection in immunosuppressed patients following medication or surgical abortion, and specific recommendations on the best type of abortion for these patients are difficult to make.

Medication abortion for women with hypercoagulable state. Although hypercoagulable conditions such as antiphospholipid syndrome have been associated with an increased incidence of adverse pregnancy outcomes, there are no data to guide counseling on abortive methods for these patients. Medication and surgical abortions carry a similar risk of thromboembolic complications. Based on a study examining all induced abortions in Finland using national health registries (22,368 medication abortions and 20,251 surgical abortions), medication abortion and surgical abortion (at a gestational age of 63 days or less) both had a 0.08% risk of thromboembolic events at 42 days of gestation (12). Women receiving thromboprophylaxis for pregnancy and the postpartum state should continue anticoagulation therapy 6 weeks after abortion.

Conclusions

Summarily, many patients with rheumatic diseases will become pregnant. Some may undergo an induced abortion for various reasons including significant health risks and adverse pregnancy outcomes due to their medical condition, use of teratogenic drugs, or other personal factors. Multiple considerations are important to keep in mind when counseling patients on the method of pregnancy termination. Moreover, the high percentage of patients with SLE at risk of unplanned pregnancy (2) highlights the importance of better pregnancy planning and contraception counseling. These patients need to be informed of the maternal and fetal risks and possible adverse pregnancy outcomes in relation to their condition to make appropriate reproductive decisions and reduce the need for pregnancy terminations.

Although there is a substantial body of literature on pregnancy and rheumatic diseases, studies on induced abortion in patients with rheumatic diseases are scant. Given significant knowledge gaps in this area, it is difficult to provide evidence-based recommendations to these patients considering an induced abortion. We emphasize the need for research on induced abortion in these patients to better assess the safety, effectiveness, and complications of different options for pregnancy termination.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Oliver JE, Silman AJ. Why are women predisposed to autoimmune rheumatic diseases? Arthritis Res Ther 2009;11:252.
2. Yazdany J, Trupin L, Kaiser R, Schmajuk G, Gillis JZ, Chakravarty E, et al. Contraceptive counseling and use among women with systemic lupus erythematosus: a gap in health care quality? Arthritis Care Res (Hoboken) 2011;63:358–65.
3. Lohr PA, Fjerstad M, DeSilva U, Lyus R. Abortion. BMJ 2014;348:f7553.
4. Jones DC, Hayslett JP. Outcome of pregnancy in women with moderate or severe renal insufficiency. N Engl J Med 1996;335:226–32.
5. Lockshin MD, Guerra M, Salmon JE. Elective termination of pregnancy in autoimmune rheumatic diseases: experience from two databases. Arthritis Rheumatol 2020;72:1325–9.
6. Kortsmit K, Jatlaoui TC, Mandel MG, Reeves JA, Oduyebo T, Petersen E, et al. Abortion Surveillance – United States, 2018. MMWR Surveill Summ 2020;69:1–29.
7. Venne K, Scott S, Bernatsky S, Vinet É. Induced abortions in women with systemic lupus erythematosus. Lupus 2021;30:484–8.
8. Vinet É, Kuriya B, Pineau CA, Clarke AE, Bernatsky S. Induced abortions in women with rheumatoid arthritis receiving methotrexate. Arthritis Care Res (Hoboken) 2013;65:1365–9.
9. National Abortion Federation. 2020 clinical policy guidelines for abortion care. 2020. URL: https://prochoice.org/store/clinical-policy-guidelines/.
10. Henshaw RC, Naji SA, Russell IT, Templeton AA. Comparison of medical abortion with surgical vacuum aspiration: women's preferences and acceptability of treatment. BMJ 1993;307:714–7.
11. Bancsi A, Grindrod K. Medical abortion: a practice tool for pharmacists. Can Pharm J (Ott) 2019;152:160–3.
12. Niinimäki M, Pouta A, Bloigu A, Gissler M, Hemminki E, Suhonen S, et al. Immediate complications after medical compared with surgical termination of pregnancy. Obstet Gynecol 2009;114:795–804.
13. Costescu D, Guilbert E, Bernardin J, Black A, Dunn S, Fitzsimmons B, et al. Medical abortion. J Obstet Gynaecol Can 2016;38:366–89.
14. Bachman EA, Barnhart K. Medical management of ectopic pregnancy: a comparison of regimens. Clin Obstet Gynecol 2012;55:440–7.
15. World Health Organization. Institutional Repository for Information Sharing. Clinical practice handbook for safe abortion. 2014. URL: https://apps.who.int/iris/handle/10665/97415.

16. American College of Obstetricians and Gynecologists. Practice bulletin no. 143: medical management of first-trimester abortion. Obstet Gynecol 2014;123:676–92.
17. Murray ME, Casson M, Pudwell J, Waddington A. Patients' motivation for surgical versus medical abortion. J Obstet Gynaecol Can 2019;41: 1325–9.
18. Costescu D, Guilbert É. No. 360-induced abortion: surgical abortion and second trimester medical methods. J Obstet Gynaecol Can 2018;40:750–83.
19. Bashal F. Hematological disorders in patients with systemic lupus erythematosus. Open Rheumatol J 2013;7:87.
20. Derksen R, de Groot PG, Kater L, Nieuwenhuis HK. Patients with antiphospholipid antibodies and venous thrombosis should receive long term anticoagulant treatment. Ann Rheum Dis 1993;52:689–92.
21. Guiahi M, Davis A. First-trimester abortion in women with medical conditions. Contraception 2012;86:622–30.
22. Davey A. Mifepristone and prostaglandin for termination of pregnancy: contraindications for use, reasons and rationale. Contraception 2006; 74:16–20.
23. Achilles SL, Reeves MF. Prevention of infection after induced abortion. Contraception 2011;83:295–309.
24. Shahane A. Pulmonary hypertension in rheumatic diseases: epidemiology and pathogenesis. Rheumatol Int 2013;33:1655–67.
25. Renjith AS, Marwaha V, Aggarwal N, Koshy V, Singal VK, Kumar KVSH. Prevalence of left ventricular dysfunction in rheumatoid arthritis. J Family Med Prim Care 2017;6:622–6.
26. Miner JJ, Kim AH. Cardiac manifestations of systemic lupus erythematosus. Rheum Dis Clin North Am 2014;40:51–60.
27. Anders HJ, Vielhauer V. Renal co-morbidity in patients with rheumatic diseases. Arthritis Res Ther 2011;13:222.
28. Regitz-Zagrosek V, Roos-Hesselink JW, Bauersachs J, Blomström-Lundqvist C, Cífková R, De Bonis M, et al. 2018 ESC guidelines for the management of cardiovascular diseases during pregnancy: The Task Force for the Management of Cardiovascular Diseases During Pregnancy of The European Society of Cardiology (ESC). Eur Heart J 2018;39:3165–241.
29. Pieper PG, Lameijer H, Hoendermis ES. Pregnancy and pulmonary hypertension. Best Pract Res Clin Obstet Gynaecol 2014;28:579–91.
30. Smyth A, Oliveira GH, Lahr BD, Bailey KR, Norby SM, Garovic VD. A systematic review and meta-analysis of pregnancy outcomes in patients with systemic lupus erythematosus and lupus nephritis. Clin J Am Soc Nephrol 2010;5:2060–8.
31. McMillan E, Martin WL, Waugh J, Rushton I, Lewis M, Clutton-Brock T, et al. Management of pregnancy in women with pulmonary hypertension secondary to SLE and anti-phospholipid syndrome. Lupus 2002;11:392–8.
32. Piccoli GB, Cabiddu G, Attini R, Vigotti FN, Maxia S, Lepori N, et al. Risk of adverse pregnancy outcomes in women with CKD. J Am Soc Nephrol 2015;26:2011–22.
33. Blom K, Odutayo A, Bramham K, Hladunewich MA. Pregnancy and glomerular disease: a systematic review of the literature with management guidelines. Clin J Am Soc Nephrol 2017;12:1862–72.
34. Eudy AM, Siega-Riz AM, Engel SM, Franceschini N, Howard AG, Clowse ME, et al. Effect of pregnancy on disease flares in patients with systemic lupus erythematosus. Ann Rheum Dis 2018;77:855–60.
35. Lidar M, Langevitz P. Pregnancy issues in scleroderma. Autoimmun Rev 2012;11:A515–9.
36. Østensen M, Andreoli L, Brucato A, Cetin I, Chambers C, Clowse ME, et al. State of the art: reproduction and pregnancy in rheumatic diseases. Autoimmun Rev 2015;14:376–86.
37. Schreiber K, Hunt BJ. Managing antiphospholipid syndrome in pregnancy. Thromb Res 2019;181 Suppl 1:S41–6.
38. Bouvier S, Cochery-Nouvellon É, Lavigne-Lissalde G, Mercier É, Marchetti T, Balducchi J-P, et al. Comparative incidence of pregnancy outcomes in treated obstetric antiphospholipid syndrome: the NOH-APS observational study. Blood 2014;123:404–13.
39. Sammaritano LR, Bermas BL, Chakravarty EE, Chambers C, Clowse ME, Lockshin MD, et al. 2020 American College of Rheumatology Guideline for the Management of Reproductive Health in Rheumatic and Musculoskeletal Diseases. Arthritis Care Res (Hoboken) 2020;72:461–88.
40. Verberne EA, de Haan E, van Tintelen JP, Lindhout D, van Haelst MM. Fetal methotrexate syndrome: a systematic review of case reports. Reprod Toxicol 2019;87:125–39.
41. Weber-Schoendorfer C, Chambers C, Wacker E, Beghin D, Bernard N, Shechtman S, et al. Pregnancy outcome after methotrexate treatment for rheumatic disease prior to or during early pregnancy: a prospective multicenter cohort study. Arthritis Rheumatol 2014;66: 1101–10.
42. Martínez Lopez JA, Loza E, Carmona L. Systematic review on the safety of methotrexate in rheumatoid arthritis regarding the reproductive system (fertility, pregnancy, and breastfeeding). Clin Exp Rheumatol 2009;27:678–84.
43. Rengasamy P. Congenital malformations attributed to prenatal exposure to cyclophosphamide. Anticancer Agents Med Chem 2017;17: 1211–27.
44. Götestam Skorpen C, Hoeltzenbein M, Tincani A, Fischer-Betz R, Elefant E, Chambers C, et al. The EULAR points to consider for use of antirheumatic drugs before pregnancy, and during pregnancy and lactation. Ann Rheum Dis 2016;75:795–810.
45. Hoeltzenbein M, Elefant E, Vial T, Finkel-Pekarsky V, Stephens S, Clementi M, et al. Teratogenicity of mycophenolate confirmed in a prospective study of the European Network of Teratology Information Services. Am J Med Genet A 2012;158A:588–96.
46. Bérard A, Zhao JP, Shui I, Colilla S. Leflunomide use during pregnancy and the risk of adverse pregnancy outcomes. Ann Rheum Dis 2018;77:500–9.

Arthritis Care & Research
Vol. 74, No. 10, October 2022, pp 1751–1760
DOI 10.1002/acr.24620

American College of Rheumatology
Empowering Rheumatology Professionals

Rheumatic Disease Disclosure at the Early Career Phase and Its Impact on the Relationship Between Workplace Supports and Presenteeism

Arif Jetha,[1] Lori Tucker,[2] Catherine Backman,[3] Vicki L. Kristman,[4] Julie Bowring,[5] Elizabeth M. Hazel,[6] Louise Perlin,[7] Laurie Proulx,[8] Cynthia Chen,[5] and Monique A. M. Gignac[1]

Objective. Young adults with rheumatic disease face challenges communicating health needs, accessing workplace support, and sustaining productivity. Our objective was to examine whether disclosure modifies the relationship between workplace support and presenteeism.

Methods. An online survey was administered to Canadian young adults with rheumatic disease and asked about presenteeism (0 = health had no effect on work; 10 = health completely prevented working), workplace support need, availability, and use and whether health details were disclosed to an immediate supervisor. A multivariable robust linear regression model was conducted and stratified by those who did and did not disclose the details of their health to their supervisor.

Results. A total of 306 participants completed the survey with a mean ± SD presenteeism score of 4.89 ± 2.65. More than 70% disclosed health details to their supervisor; those who disclosed reported greater presenteeism (mean ± SD 5.2 ± 2.5) when compared to those who did not disclose (mean ± SD 4.2 ± 2.61). Greater disease severity was associated with disclosure. Half of the participants reported unmet workplace support needs (53%), 32% reported that their workplace support needs were met, and 15% reported exceeded workplace support needs. The relationship between presenteeism and workplace support needs was modified by disclosure. For participants who disclosed, workplace support needs that were unmet (β = 1.59 [95% confidence interval (95% CI) 0.75, 2.43]) and that were met (β = 1.25 [95% CI 0.39, 2.11]) were associated with greater presenteeism when compared to those with exceeded workplace support needs.

Conclusion. To address presenteeism, strategies should be developed for young adults with rheumatic disease to foster access to available workplace supports and to navigate disclosure decisions.

INTRODUCTION

Young adulthood, a period spanning 18–35 years, represents a critical transitional life phase where a person establishes themselves within the labor market, often making occupational changes toward achieving full-time work. Presenteeism (i.e., working while unwell) during young adulthood can impact early career success and contribute to difficulties with sustaining and advancing within the workforce (1). Increasingly, studies show that a rheumatic disease in young adulthood can be associated with barriers to employment that are attributed to the severity of symptoms and to work environments that lack appropriate supports (2). Moreover, the invisible and episodic nature of many rheumatic diseases may add to the complexity related to the disclosure of health needs and requesting workplace supports that are necessary to addressing employment barriers. We examined how disclosure of health details can modify the relationship between workplace support and presenteeism, using data from a Canadian survey of young adults with rheumatic disease.

Supported by a Young Investigator Operating Grant from the Arthritis Society (Canada) (#17-0095).

[1]Arif Jetha, PhD, Monique A. M. Gignac, PhD: Institute for Work and Health and Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; [2]Lori Tucker, MD: University of British Columbia and British Columbia Children's Hospital, Vancouver, British Columbia, Canada; [3]Catherine Backman, PhD, Reg. OT (BC): University of British Columbia and Arthritis Research Canada, Vancouver, British Columbia, Canada; [4]Vicki L. Kristman, PhD: Institute for Work and Health, Toronto, and Lakehead University, Thunder Bay, Ontario, Canada; [5]Julie Bowring, MA, Cynthia Chen, MSc: Institute for Work and Health, Toronto, Ontario, Canada; [6]Elizabeth M. Hazel, MD: McGill University Health Centre, Montreal, Quebec, Canada; [7]Louise Perlin, MD: University of Toronto and St. Michael's Hospital, Toronto, Ontario, Canada; [8]Laurie Proulx: Canadian Arthritis Patient Alliance, Ottawa, Ontario, Canada.

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

Address correspondence to Arif Jetha, PhD, Institute for Work and Health, 400 University Avenue, Suite 1800, Toronto, Ontario, Canada, M5G 2E9. Email: AJetha@iwh.on.ca.

Submitted for publication January 6, 2021; accepted in revised form April 8, 2021.

SIGNIFICANCE & INNOVATIONS

- Our study is one of the first to unpack the relationship between the disclosure of health details, workplace support needs, and presenteeism for young adults with rheumatic disease.
- More than half of young adults with rheumatic disease in our study reported that their workplace support needs were unmet.
- More than two-thirds of young adults with rheumatic disease described disclosing health details to their supervisor; those with a more severe disease were more likely to disclose.
- The relationship between unmet workplace support needs and presenteeism was significant for participants who disclosed health details.

Within industrialized countries, rheumatic disease is one of the most prevalent chronic health conditions affecting the working population and is a frequently reported cause of lost productivity (3–5). For young adults, a rheumatic disease may be associated with unique challenges in the school-to-work transition, including difficulties finding and sustaining stable employment, sustaining productivity, and achieving career advancement while balancing work, health, and personal responsibilities (2,6–9). Challenges in the school-to-work transition can be exacerbated by ongoing needs for health care that can change as a young person moves from pediatric to adult health care settings (10,11). A Canadian survey of young adults living with juvenile arthritis (JA) and systemic lupus erythematosus (SLE) indicated that >40% reported lost productivity, including health-related missed work days and job disruption. Productivity loss was more likely to be reported by young adults with more severe rheumatic disease symptoms (e.g., greater pain, fatigue, and activity limitations) and by those who reported more challenging work contexts (e.g., less job control or supervisor support) (12).

A supportive work environment can play an important role in addressing presenteeism and strengthening person–job fit for people with chronic disease (13–15). Studies with older age groups living with rheumatic disease show that diverse workplace support needs remaining unmet (e.g., job accommodation, work modification, and health benefits) are often associated with presenteeism and greater workplace activity limitations, while having workplace support needs met or even exceeded are related to less presenteeism (13,16,17). Importantly, existing research on workplace supports may not always be relevant to young adults with rheumatic disease, who have less established employment histories and who are more likely to work in nonstandard employment where formal accommodations are less likely to be available (18).

The disclosure of rheumatic disease by employees within the workplace may play an important role in determining access to workplace supports, especially for those requiring an accommodation (e.g., accessible workstation) or access to a work modification (e.g., scheduling flexibility) (19). When compared to older age groups, young adults with rheumatic disease report greater hesitancy in communicating details about their health to their supervisor (15,20). Qualitative research has found that intermittent and unpredictable disease symptoms coupled with less job tenure, inexperience with workplace self-advocacy, and poorly established relationships with a supervisor are commonly described barriers to communicating needs and requesting workplace supports (15,20). Life course research suggests that the timing of events can also impact work-related perceptions and behaviors. Specifically, a rheumatic disease is often seen by others as a condition of older adults. A rheumatic disease may be considered by others as occurring at a nonnormative time when experienced by a young adult (21–23). Consequently, there may be apprehension in requesting assistance out of concern of a negative reaction from supervisors (7,15,20). At the same time, privacy legislation within many industrialized countries means that workers with rheumatic disease are not legally obligated to reveal their health condition to an employer unless there is a safety concern (24). Other research indicates that the disclosure of a health condition to an immediate supervisor may modify the relationship between workplace supports and presenteeism; those who disclose may be more likely to have their workplace supports met, thereby attenuating the impact of health on work (19,24).

Little research has examined the association between rheumatic disease disclosure at the early career phases and how it modifies the relationship between workplace support needs and presenteeism. Our study aimed to address this knowledge gap in a cohort of Canadian young adults with rheumatic disease. We addressed 4 study objectives: 1) to describe the proportion of participants who reported that their workplace support needs were unmet, met, or exceeded, and who reported disclosing the details of their health to their immediate supervisor; 2) to examine the relationship between disclosure of health details to an immediate supervisor and whether workplace support needs were unmet, met, or exceeded; 3) to compare whether presenteeism differed according to whether workplace support needs were unmet, met, or exceeded and whether or not a participant disclosed the details of their health to an immediate supervisor; and 4) to examine whether the relationship between workplace support needs and presenteeism was modified by disclosure of health details when adjusting for sociodemographic, disease/health, and work context factors.

MATERIALS AND METHODS

We analyzed cross-sectional data from an ongoing longitudinal online survey of young adults with rheumatic disease. To be eligible, participants had to be between ages 18 and 35 years, report a doctor-diagnosed rheumatic condition (e.g., JA, SLE, rheumatoid arthritis) and have paid employment in the past

Table 1. Description of total sample of young adults with rheumatic disease and by whether a participant disclosed health details to their immediate supervisor or manager*

	Total sample (n = 306)	Disclosed (n = 216)	Not disclosed (n = 90)	*P*
Sociodemographic factors				
Age, years	28.5 ± 4.5	28.7 ± 4.5	28.0 ± 4.4	0.25
Sex, no. (%)				
Women	187 (63.1)	136 (63)	57 (63.3)	0.95
Men	113 (36.9)	80 (37.0)	33 (36.7)	–
Educational attainment, no. (%)				
<Postsecondary education	51 (16.7)	35 (16.2)	16 (17.8)	0.95
≥Postsecondary education†	255 (83.3)	181 (83.8)	74 (82.2)	–
Married/living as if married	138 (45.1)	105 (48.6)	33 (36.7)	0.06
Primary childcare responsibilities	69 (22.6)	55 (25.5)	14 (15.6)	0.06
Disease/health factors				
Pediatric disease onset (age <18 years), no. (%)	157 (51.3)	108 (50.0)	49 (54.4)	0.48
Pain (0–10)	5.5 ± 2.5	5.8 ± 2.4	5.1 ± 2.7	0.03
Fatigue (0–10)	6.1 ± 2.3	6.3 ± 2.2	5.6 ± 2.6	0.03
Disease activity (0–10)	4.9 ± 2.6	5.1 ± 2.5	4.4 ± 3.0	0.04
Self-rated health, no. (%)				
Poor	30 (9.8)	21 (9.7)	9 (10.0)	0.01
Fair	128 (41.8)	101 (46.8)	27 (30.0)	–
Good	101 (33.0)	65 (30.1)	36 (40.0)	–
Very good	38 (12.4)	24 (11.1)	14 (15.6)	–
Excellent	9 (2.9)	5 (2.3)	4 (4.4)	–
Depression (PHQ-2), no. (%)	109 (35.6)	80 (37.0)	29 (32.2)	0.42
Workplace activity limitations (WALS: 0–36)	11.6 ± 6.4	12.6 ± 6.2	9.2 ± 6.2	0.001
Work context factors				
Employed status, no. (%)				
Full-time (≥30 hours/week)	208 (68.0)	147 (68.1)	61 (68.0)	0.96
Part-time (<30 hours/week)	98 (32.0)	69 (31.9)	29 (32.2)	–
Employment contract, no. (%)				
Permanent	232 (75.8)	167 (77.3)	65 (72.2)	0.34
Temporary	74 (24.2)	49 (22.7)	25 (27.8)	–
Job tenure, years	2.9 ± 2.7	3.06 ± 2.8	2.53 ± 2.5	0.13
Job sector				
Trades, no. (%)	61 (19.9)	44 (20.4)	17 (18.9)	0.92
Sales and services	39 (12.8)	28 (13.0)	11 (12.2)	–
Professional services	67 (21.9)	45 (20.8)	22 (24.4)	–
Health care/social services	119 (38.9)	86 (39.8)	33 (36.7)	–
Technology	20 (6.5)	13 (6.0)	7 (7.8)	–
Job control (1–5)	2.8 ± 1.1	2.8 ± 1.0	2.8 ± 1.1	0.72
Workplace physical activity requirements (1–5)	3.0 ± 1.2	3.0 ± 1.2	3.0 ± 1.3	0.99
Mental job demands (1–5)	3.5 ± 1.1	3.5 ± 1.1	3.5 ± 1.1	0.86
Job stress (1–5)	3.1 ± 0.9	3.2 ± 0.9	3.0 ± 1.0	0.11
Organizational support (1–5)	3.2 ± 1.2	3.2 ± 1.11	3.1 ± 1.3	0.27
Workplace support needs, no. (%)				
Workplace support needs exceeded	48 (15.7)	32 (14.8)	16 (17.8)	0.80
Workplace support needs met	97 (31.7)	68 (31.5)	29 (32.2)	–
Workplace support needs unmet	161 (52.6)	116 (53.7)	45 (50.0)	–
Presenteeism (0–10)	4.89 ± 2.65	5.18 ± 2.47	4.19 ± 2.61	0.006

* Values are the mean ± SD unless indicated otherwise. PHQ-2 = 2-item Patient Health Questionnaire; WALS = Workplace Activity Limitations Scale.
† Postsecondary educational attainment includes training from a college or university.

year. Self-reported doctor diagnosis of rheumatic disease is considered a valid case-finding approach for public health research (25). Also, our decision to include different rheumatic diseases when constructing our cohort was informed by previous research showing that at-work experiences and workplace support needs are comparable even though clinical features may differ (26).

We used 3 recruitment approaches to maximize engagement. Participants were recruited directly from clinics in 3 Canadian provinces (British Columbia, Ontario, and Quebec). Eligible participants recruited through clinics were provided with a study invitation card by a clinic representative with a link to the online questionnaire. Second, young adult participants with rheumatic disease were recruited using an existing panel maintained by a

Table 2. Workplace supports that young adults with rheumatic disease reported as being needed, available, and used*

	Needed	Available	Used
Work schedule flexibility	277 (90.5)	239 (78.1)	260 (85.0)
Prescription drug coverage	264 (86.3)	224 (73.2)	236 (77.1)
Extended health benefits	261 (85.3)	208 (68.0)	216 (70.6)
Paid sick leave	246 (80.4)	196 (64.1)	197 (64.6)
Modified job duties	235 (76.8)	194 (63.4)	190 (62.1)
Informal work modification	216 (70.6)	201 (65.7)	192 (62.8)
Facilities or opportunities to manage health at work	209 (68.3)	195 (63.7)	168 (54.9)
Workstation adaptations	199 (65.0)	196 (64.1)	169 (55.2)
Work-from-home arrangements	192 (62.6)	140 (45.8)	153 (50.0)
Employee assistance program	184 (60.1)	179 (58.5)	141 (46.1)
Accessible workplace	184 (60.1)	214 (69.9)	166 (54.3)
Assistive devices or technology	166 (54.3)	163 (53.3)	146 (47.5)
Other workplace supports	112 (36.6)	45 (14.7)	65 (21.2)

* Values are the number (%).

research firm consisting of >1 million Canadians that is nationally representative according to region and income. Third, community-based recruitment was conducted through 3 nonprofit organizations that support the health and employment needs of young people with rheumatic conditions. Each community organization shared study advertisements through their listservs or social media accounts. All potential participants were provided with study information, and informed consent was obtained before they completed the questionnaire (27). Study procedures were approved by the University of Toronto Research Ethics Board (REB# 36588).

Survey. The online questionnaire was in English or French and took ~30 minutes to complete. Survey items were selected based on their psychometric properties and use in previous studies.

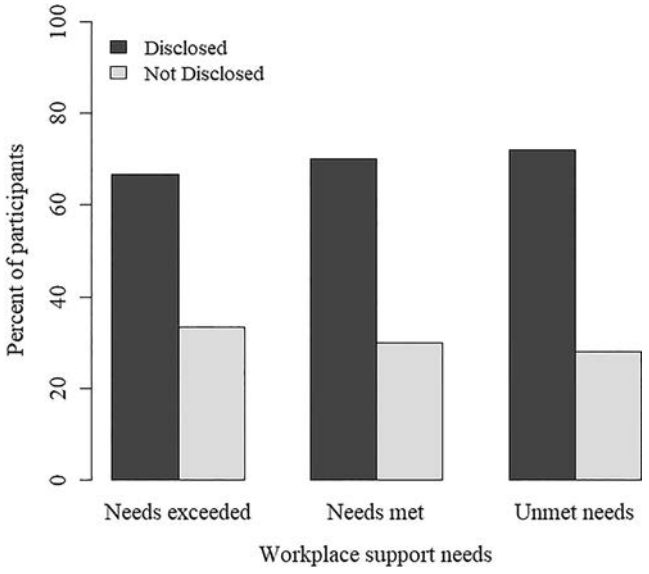

Figure 1. Frequency of disclosure of health details to an immediate supervisor or manager based on whether young adult participants with rheumatic disease reported that their workplace support needs were unmet, met, or exceeded.

Outcome measure: presenteeism. A global item from the Work Productivity and Activity Impairment Questionnaire asked participants to rate the extent to which their health affected presenteeism in the past week: "During the past seven days, how much did your health affect your productivity while you were working?" Response options were provided on an 11-point scale (0 = health had no effect on my work; 10 = health completely prevented me from working) (28).

Independent variable: workplace support needs unmet, met, or exceeded. A list of 13 job accommodations (e.g., workstation adaptation), modifications (e.g., work schedule flexibility), or benefits (e.g., prescription drug coverage) were presented to participants, based on previous studies of accommodation practices for people with rheumatic disease (15). Participants were asked whether a particular job accommodation, modification, or benefit was available (yes/no/don't know), needed (yes/no), and used (yes/no). Using responses, a 3-level variable was constructed: 1) unmet workplace support need (participant's need for workplace supports was greater than their use of available workplace support), 2) workplace support needs met (participant's need for workplace support was equal to their use of available workplace support), and 3) workplace support needs exceeded (participant's need for workplace support was less than the available workplace supports) (13).

Modifier variable: disclosure to supervisor. A single item asked whether a participant had disclosed their health details to their immediate supervisor: "Have you talked to your immediate supervisor/manager about any limitations you have that might affect your work and that are related to your rheumatic disease?" Respondents provided a dichotomous response (0 = no; 1 = yes) (19).

Covariates. Sociodemographic, disease/health, and work context factors were collected for descriptive purposes and were adjusted for in multivariable models. Sociodemographic factors included age (years), sex/gender, education (postsecondary

Table 3. Univariable linear regression models examining factors associated with presenteeism, including workplace support needs that were exceeded, met, or unmet, disclosure of health details to an immediate supervisor or manager, and study covariates*

	Values
Sociodemographic factors	
Age, years	0.02 (−0.0, 0.09)
Sex	
Men	0.35 (−0.27, 0.98)
Women	–
Educational attainment	
<Postsecondary education	–
≥Postsecondary education†	0.22 (−0.58, 1.02)
Married/living as if married	−0.07 (−0.67, 0.52)
Primary childcare responsibilities	0.52 (−0.19, 1.23)
Disease/health factors	
Pediatric disease onset (<18 years)	−0.78 (−1.37, −0.19)‡
Pain (0–10)	0.60 (−0.04, 0.09)
Fatigue (0–10)	0.54 (0.42, 0.65)‡
Disease activity (0–10)	0.56 (0.47, 0.66)
Self-rated health	
Poor/fair	–
Good/very good/excellent	−1.18 (−1.77, −0.60)‡
Depression (PHQ-2)	1.97 (1.39, 2.55)‡
Workplace activity limitations (WALS: 0–36)	0.19 (0.15, 0.24)‡
Work context factors	
Employment status	
Full-time (≥30 hours/week)	−0.30 (−0.94, 0.34)
Part-time (<30 hours/week)	–
Employment contract	
Temporary	–
Permanent	−0.09 (−0.79, 0.60)
Job sector	
Sales and services	–
Professional services	0.44 (−0.60, 1.49)
Health care/social services	0.18 (−0.78, 1.13)
Technology	0.76 (−0.66, 2.19)
Job control (1–5)	0.03 (−0.25, 0.31)
Workplace physical activity requirements (1–5)	0.16 (−0.08, 0.40)
Mental job demands (1–5)	−0.02 (−0.29, 0.25)
Job stress (1–5)	0.81 (0.50, 1.12)
Organizational support (1–5)	−0.24 (−0.49, 0.01)
Disclosed to supervisor or manager	0.99 (0.35, 1.64)‡
Workplace support needs	
Workplace support needs exceeded	–
Workplace support needs met	1.51 (0.61, 2.41)‡
Workplace support needs unmet	1.61 (0.77, 2.45)‡

* Values are the β estimate from univariable linear regression model (95% confidence interval). Presenteeism was measured on an 11-point scale (0 = health had no effect on my work; 10 = health completely prevented me from working). PHQ-2 = 2-item Patient Health Questionnaire; WALS = Workplace Activity Limitations Scale.
† Postsecondary educational attainment includes training from a college or university.
‡ Statistically significant.

educational attainment including training from college or university), and marital status (married/living as if married). Disease/health factors, including information on pediatric onset of a rheumatic disease (age <18 years) and self-rated health (1 = poor; 5 = excellent), were obtained. Self-reported pain, fatigue, and disease activity were measured using 11-point scales (0 = no pain/fatigue/disease activity; 10 = worst possible pain/fatigue/disease activity) (29). Participants were asked about the frequency of depressed mood in the last 2 weeks using the 2-item Patient Health Questionnaire (0 = not at all; 3 = nearly every day). A total sum score of >3 suggested the likelihood of depression (30). Participants also completed the Workplace Activity Limitation Scale (WALS) to measure difficulties with workplace activities and tasks. WALS is a 12-item scale that asks about problems with lower mobility, upper mobility, concentration, and the pace and schedule of work (0 = no difficulty/not applicable to job; 3 = unable to do). Items were summed to produce a score ranging from 0 to 36 (31).

Work context factors were obtained by asking participants about the details of their current or recent employment, including whether they worked part-time (<30 hours/week) or full-time hours (≥30 hours/week) and whether they had a permanent or temporary contract. Participants were also asked about their job tenure (years), job sector in which they were employed (trades/transportation, sales/services, professional services, health care/social services, technology), and the extent to which their employment had physical activity requirements (1 = not at all; 5 = a great deal) and mental job demands (1 = not at all; 5 = a great deal). Additionally, participants were asked about their perceptions of job control, job stress, and organizational support (1 = not at all; 5 = a great deal).

Statistical analysis. Descriptive statistics were used to examine variable distributions. Chi-square tests and *t*-tests were conducted to examine how study variables differed for those

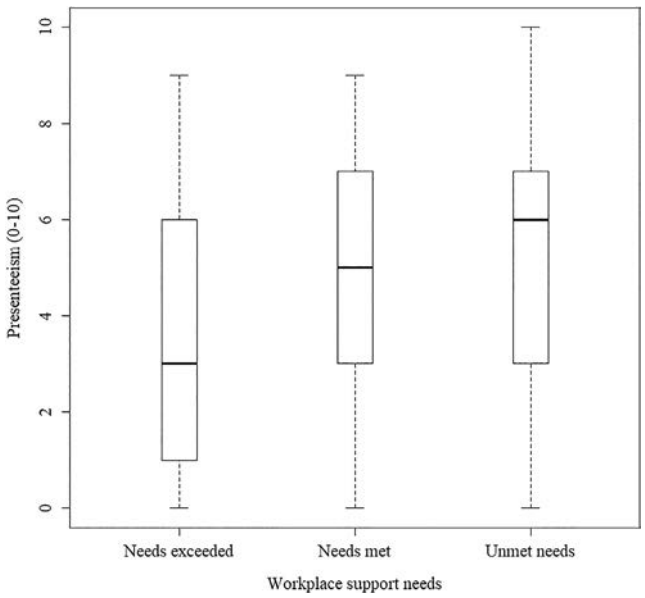

Figure 2. Box plot comparing presenteeism based on participants' workplace support needs being exceeded, met, or unmet. Solid line shows median presenteeism score; lower and upper whiskers represent lower and upper adjacent values.

Table 4. Stratified multivariable robust regression model examining the relationship between presenteeism and whether workplace support needs were exceeded, met, or unmet*

	Disclosed (n = 216)	Not disclosed (n = 90)
Sociodemographic factors		
Age, years	0.00 (−0.07, 0.07)	0.02 (−0.07, 0.11)
Sex		
Women	–	–
Men	0.52 (−0.12, 1.17)	0.09 (−0.93, 1.11)
Education		
<Postsecondary education	–	–
≥Postsecondary education†	0.81 (0.10, 1.54)‡	−0.32 (−1.46, 0.82)
Married/living as if married	−0.25 (−0.83, 0.33)	0.12 (−0.71, 0.95)
Primary childcare responsibilities	−0.43 (0.34, −1.10)	−0.46 (−1.61, 0.70)
Disease/health factors		
Pediatric disease onset (<18 years)	0.08 (−0.52, 0.69)	0.30 (−0.55, 1.15)
Pain (0–10)	0.40 (0.09, 0.22)‡	0.14 (−0.09, 0.38)
Fatigue (0–10)	0.17 (0.01, 0.33)‡	0.23 (−0.01, 0.47)‡
Disease activity (0–10)	0.06 (−0.11, 0.24)	0.47 (0.24, 0.71)‡
Self-rated health		
Poor/fair	–	–
Good/very good/excellent	0.04 (−0.56, 0.64)	0.27 (−0.69, 1.23)
Depression (PHQ-2)	0.62 (0.01, 1.23)‡	0.16 (−0.74, 1.06)
Workplace activity limitations (WALS: 0–36)	0.11 (0.06, 0.16)‡	0.10 (0.02, 0.17)‡
Work context factors		
Employment status		
Full-time (≥30 hours/week)	–	–
Part-time (<30 hours/week)	−0.20 (−0.84, 0.45)	0.17 (−0.70, 1.05)
Employment contract		
Temporary	–	–
Permanent	−0.98 (−1.64, −0.31)‡	−0.06 (−0.96, 0.83)
Job sector		
Sales and services	–	–
Trades and transportation	0.01 (−1.01, 1.03)	1.75 (0.32, 3.17)‡
Professional services	−0.20 (−1.25, 0.85)	0.90 (−0.36, 2.16)
Health care/social services	−0.39 (−1.33, 0.56)	0.87 (−0.44, 2.17)
Technology	0.65 (−0.72, 2.01)	1.27 (−0.36, 2.90)
Job control (1–5)	0.18 (−0.10, 0.46)	0.05 (−0.28, 0.37)
Workplace physical activity requirement (1–5)	−0.34 (−0.58, −0.10)‡	0.27 (−0.03, 0.57)
Mental job demands (1–5)	−0.02 (−0.28, 0.23)	−0.12 (−0.51, 0.28)
Workplace support needs		
Workplace support needs exceeded	–	–
Workplace support needs meet	1.25 (0.39, 2.11)‡	−0.58 (−1.66, 0.50)
Workplace support needs unmet	1.59 (0.75, 2.43)‡	−0.79 (−1.86, 0.27)

* Values are the β estimate from robust multivariable linear regression model (95% confidence interval). The model was stratified according to whether a young adult with rheumatic disease disclosed the details of their health to their immediate supervisor. Presenteeism was measured on an 11-point scale (0 = health had no effect on my work; 10 = health completely prevented me from working). PHQ-2 = 2-item Patient Health Questionnaire; WALS = Workplace Activity Limitations Scale.
† Postsecondary educational attainment includes training from a college or university.
‡ Statistically significant.

who had disclosed the details of their health to the supervisor and those who had not. Univariable linear regression analyses were conducted to examine the association between the study variables and presenteeism. Covariates of theoretical importance as well those significantly related to presenteeism at the univariable level were carried forward to the final multivariable model. To test study hypotheses, a multivariable linear regression model was developed that was stratified for those who did and did not disclose the details of their health to their supervisor. A robust regression was chosen because it is less sensitive to data with variables that may not exhibit normality or may possess atypical values when compared to linear regression models with ordinary least-squares estimators (32). Analyses were conducted using SAS software, version 9.3 (33).

RESULTS

A total of 306 young adults with a rheumatic disease completed the survey (mean age 28.5 ± 4.5 years), of whom less than two-thirds were recruited through the research firm (64%), 26% were recruited through community-based organizations, and 11% were recruited from rheumatology clinics (27).

The sample is described in Table 1. Approximately two-thirds of participants were women (63%), 45% were married/living

as married, 23% reported childcare responsibilities, and most reported obtaining a postsecondary education (83%). Half of the participants reported a pediatric onset of their rheumatic disease, over half of the sample indicated poor/fair self-rated health (52%), and 36% indicated depression. More than two-thirds of participants reported full-time employment (68%), and three-fourths held a permanent contract (76%). Mean ± SD job tenure was 2.9 ± 2.7 years. Notably, of a possible 10 points, participants reported a mean ± SD presenteeism score of 4.89 ± 2.65.

Close to 70% of participants reported disclosing the details of their health to their immediate supervisor or manager at work. When compared to those who did not disclose health details, participants who disclosed reported significantly greater mean pain (5.1 versus 5.8), fatigue (5.6 versus 6.3), disease activity (4.4 versus 5.1), and WALS scores (9.2 versus 12.6). Also, when compared to those who did not disclose, a significantly greater frequency of participants who reported disclosing the health details indicated fair/poor health (40% versus 57%). A greater frequency of participants who disclosed health details indicated depression when compared to those who did not disclose (37% versus 32%), but this relationship was not statistically significant.

An examination of the specific workplace supports needed, available, and used are reported in Table 2. The most needed workplace supports included work schedule flexibility (91%), prescription drug coverage (86%), extended health benefits (85%), paid sick leave (80%), and modified job duties (77%). With the exception of an accessible workplace, participants reported that their need for the different workplace supports exceeded the reported availability and use. More than half of participants reported that their workplace support needs were unmet (53%). In comparison, 32% reported that their workplace support needs were met, and 16% reported that their workplace support needs were exceeded. There were no significant differences in the frequency of participants disclosing the details of their health to their supervisor based on workplace support needs being reported as unmet, met, or exceeded (Table 1 and Figure 1).

Univariable analyses examined the relationships between workplace support needs, disclosure of health details, and presenteeism. Participants who reported disclosing health details to their supervisor reported greater presenteeism compared to those who did not disclose (β = 0.99 [95% confidence interval (95% CI) 0.35, 1.64]) (Table 3). Additionally, at the univariable level, those who reported that their workplace support needs were met (β = 1.51 [95% CI 0.61, 2.41]) or unmet (β = 1.61 [95% CI 0.77, 2.45]) reported greater presenteeism when compared to those who reported that their workplace support needs were exceeded (Figure 2).

The final multivariable model examined the relationship between workplace support needs and presenteeism and was stratified by disclosure (Table 4). The model was adjusted for sociodemographic, disease/health, and work context factors. The relationship between presenteeism and workplace support needs was statistically significant for participants who reported disclosure of health details. Workplace support needs reported as being unmet were associated with a 1.59-point increase in presenteeism when compared to those reporting workplace support needs as being exceeded (β = 1.59 [95% CI 0.75, 2.43]). Workplace support needs reported as being met were associated with a 1.25-point increase in presenteeism when compared to those who reported that their needs were exceeded (β = 1.25 [95% CI 0.39, 2.11]). Of note, for those who disclosed, greater pain (β = 0.40 [95% CI 0.09, 0.22]) and having depression (β = 0.62 [95% CI 0.01, 1.23]) were associated with greater presenteeism. For participants not disclosing their health details, the relationship between workplace support needs and presenteeism was not statistically significant.

DISCUSSION

Employment at an early career phase can shape longer-term experiences in the labor market. Our survey is one of the first to unpack the complex relationship between the disclosure of health details, supportiveness of the work environment, and productivity for Canadian young adults with rheumatic disease. Our survey highlighted the fact that workplace support needs at the early career phase can go unmet. What is more, we found that unmet workplace needs were associated with greater presenteeism for those who reported disclosing at least some health details to their supervisor. Encouraging productive employment at the early career phase can play an important role in enhancing work and health outcomes across the life course (1,34). Findings have implications for young adults with rheumatic disease and their clinical care teams to encourage the identification and acquisition of employment in supportive work environments and to help navigate disclosure decisions. Results also have implications for supervisors and other workplace stakeholders (e.g., human resource representatives and disability management professionals) to facilitate availability and access to diverse job accommodations, modifications, and benefits for young adults entering the workplace and to create work environments where employees are comfortable discussing their needs.

Our survey is one of the largest of its kind to ask young adults with rheumatic disease with employment experience about their workplace support needs. Aligning with previous research of older age groups with rheumatic disease, the most needed workplace supports among our young adult sample included employer-provided prescription drug coverage and extended health benefits, work schedule flexibility, and modified job duties (13,14). These accommodations play an important role in addressing the impact of rheumatic disease symptoms and activity limitations on employment participation (15). Of concern, more than half of the participants in our study reported that their workplace support needs were unmet. Our findings provide evidence that

young workers with rheumatic disease may start their career in work environments where supports are less accessible (18,35). Results can be explained by Canadian labor market analyses, which show that, when compared to older age groups, young adults are more likely to be employed precariously and in jobs where formal accommodations and extended health benefits are less likely to be provided (20,36). Additional research is required to expand on the barriers and facilitators within the work environment that may be unique to young adults with rheumatic disease and to determine access to workplace supports.

The relationship between unmet workplace support needs and presenteeism is complex and may depend on the extent to which a young person communicates the details of their health at work. More than two-thirds of young adults in our study indicated disclosing health details to their supervisor. Those who did disclose the details of their condition indicated greater disease severity, more workplace activity limitations, and greater presenteeism when compared to those who did not disclose the details of their condition. Moreover, the relationship between unmet workplace support needs and presenteeism was only significant for those who had disclosed the health details to their supervisor. Importantly, existing Canadian privacy legislation means that young workers with rheumatic disease are not obligated to disclose the details of their health condition (24). Those with a well-managed disease and less severe symptoms that interfere with work may not be required to disclose health details to obtain assistance.

Findings could also be explained by emerging research on disclosure decisions conducted in samples living with a broader range of invisible and episodic chronic health conditions. These studies find that individuals may choose to communicate health details when there is a crisis situation (e.g., severe flares of pain or depressive episode) and when a workplace support is necessary to address lost productivity (24). Additionally, having a rheumatic disease at a young age, which may be invisible to others, could be associated with unique challenges in communicating health needs. Past studies have found that young adults with rheumatic disease may choose to not disclose so as to protect themselves from the potential of a negative reaction from a supervisor or to ensure that they are not excluded from career advancement opportunities (e.g., job upskilling, business travel) (7,36). Our results draw greater attention toward the development of resources that are directed to the unique needs of young people with rheumatic disease to understand the pros and cons of communicating health needs at work (20).

Of interest, when compared to those with unmet or met workplace support needs, participants who indicated that their needs were exceeded by their employer were significantly less likely to report presenteeism. Importantly, these study findings are cross-sectional, and causation cannot be determined. Nonetheless, our study adds to growing evidence on the importance of the work context to fostering the productivity of people with rheumatic disease (37–39). In particular, our study shows that employers who offer diverse workplace supports can attenuate the relationship between rheumatic disease and presenteeism at the early career phase. Alternatively, for young adults with rheumatic disease, the absence of workplace supports may contribute to a lack of fit between health needs and characteristics of the work environment. Importantly, employers often report being unaware of the number of employees who are living with chronic disease and the types of accommodations and modifications that are most needed (40). Our study suggests that targeted knowledge translation efforts to employers may be needed to increase awareness of the benefits of a supportive work environment for young workers with rheumatic disease and to provide recommendations on ways the employers can support employment success. Findings may also inform vocational rehabilitation recommendations for young adult patients in transitional rheumatology settings, to encourage a consideration of the importance of supportive work environments and the specific accommodations, modifications, and benefits that can address health needs and boost productivity.

A strength of our study was our diverse sample of employed young adults with rheumatic disease from across Canada. In particular, our purposive recruitment approach enabled us to construct a cohort of young people with rheumatic disease from clinical and community settings who may differ in terms of access to health care and who ranged according to personal, disease/health, and work context factors. At the same time, a majority of participants in our study indicated having a postsecondary education. Also, while our multivariable model controlled for work context factors, we did not have the statistical power to examine differences in the availability and need of workplace support according to specific job sectors. Additional research of participants who may range in educational attainment and are working across a broader range of job sectors could be beneficial in further understanding the experiences of young adults with rheumatic disease within different work environments and occupations, and such research may enhance the generalizability of our findings. While our survey captured self-reported information on disclosure of health details, details on the content and amount of information shared with a supervisor are unclear. Future research on the disclosure processes (e.g., details communicated to an employer, timing of disclosure) could advance recommendations provided to young adults with rheumatic disease on the communication of their health information. Finally, our study was cross-sectional. Longitudinal research across the school-to-work transition is needed to expand on our results and determine causal pathways in the relationship between disclosure, workplace supports, and presenteeism.

Supporting productivity at the early career phase can have important implications for young people with rheumatic disease as they enter the labor market and across their working life. Our study highlights the complex interrelationship between disease

disclosure, unmet workplace support needs, and presenteeism. Indeed, employers who offer diverse job accommodations, modifications, and health benefits provide a more supportive work environment and play an important role in ensuring that young workers with rheumatic disease are able to sustain productivity. However, the benefits of workplace support may only be accessed by those who communicate their needs. Findings underscore the importance of equipping young people with resources that can be used to navigate disease disclosure and requests for support as they establish their careers with a rheumatic disease.

ACKNOWLEDGMENTS

We would like to acknowledge the Arthritis Society (Canada), the Cassie + Friends Society, and the Canadian Arthritis Patient Alliance for their support with recruitment. We would also like to acknowledge Jocelyn Dollack for project coordination support.

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. Jetha 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. Jetha, Tucker, Backman, Kristman, Bowring, Proulx, Gignac.

Acquisition of data. Tucker, Bowring, Hazel, Perlin, Gignac.

Analysis and interpretation of data. Jetha, Chen.

REFERENCES

1. Scarpetta S, Sonnet A, Manfredi T. Rising youth unemployment during the crisis: how to prevent negative long-term consequences on a generation? URL: https://ideas.repec.org/p/oec/elsaab/106-en.html.
2. Jetha A. The impact of arthritis on the early employment experiences of young adults: a literature review. Disabil Health J 2015;8:317–24.
3. Burton W, Morrison A, Maclean R, Ruderman E. Systematic review of studies of productivity loss due to rheumatoid arthritis. Occup Med 2006;56:18–27.
4. Perruccio AV, Power JD, Badley EM. Revisiting arthritis prevalence projections: it's more than just the aging of the population. J Rheumatol 2006;33:1856–62.
5. Zhang W, Anis AH. The economic burden of rheumatoid arthritis: beyond health care costs. Clin Rheumatol 2011;30:25–32.
6. Jetha A, Badley E, Beaton D, Fortin PR, Shiff NJ, Rosenberg AM, et al. Transitioning to employment with a rheumatic disease: the role of independence, overprotection, and social support. J Rheumatol 2014;41:2386–94.
7. Hanson H, Hart RI, Thompson B, McDonagh JE, Tattersall R, Jordan A, et al. Experiences of employment among young people with juvenile idiopathic arthritis: a qualitative study. Disabil Rehabil 2018; 40:1921–8.
8. Jetha A, Theis KA, Boring MA, Barbour KE. Education and employment participation in young adulthood: what role does arthritis play? Arthritis Care Res (Hoboken) 2017;69:1582–9.
9. Packham J, Hall M. Long-term follow-up of 246 adults with juvenile idiopathic arthritis: education and employment. Rheumatology (Oxford) 2002;41:1436–9.
10. Hazel E, Zhang X, Duffy CM, Campillo S. High rates of unsuccessful transfer to adult care among young adults with juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2010;8:1–6.
11. Tucker LB, Cabral DA. Transition of the adolescent patient with rheumatic disease: issues to consider. Rheum Dis Clin North Am 2007;33: 661–72.
12. Jetha A, Badley E, Beaton D, Fortin PR, Shiff NJ, Gignac MA. Unpacking early work experiences of young adults with rheumatic disease: an examination of absenteeism, job disruptions, and productivity loss. Arthritis Care Res (Hoboken) 2015;67:1246–54.
13. Jetha A, Johnson SR, Gignac MA. Unmet workplace support needs and lost productivity of workers with systemic sclerosis: a path analysis study. Arthritis Care Res (Hoboken) 2021;73:423-31.
14. Gignac MA, Cao X, McAlpine J. Availability, need for, and use of work accommodations and benefits: are they related to employment outcomes in people with arthritis? Arthritis Care Res (Hoboken) 2015; 67:855–64.
15. Jetha A, Gignac MA, Bowring J, Tucker S, Connelly CE, Proulx L, et al. Supporting arthritis and employment across the life course: a qualitative study. Arthritis Care Res (Hoboken) 2018;70: 1461–8.
16. Gignac MA, Kristman V, Smith PM, Beaton DE, Badley EM, Ibrahim S, et al. Are there differences in workplace accommodation needs, use and unmet needs among older workers with arthritis, diabetes and no chronic conditions? Examining the role of health and work context. Work Aging Retire 2018;4:381–98.
17. Gignac MA, Ibrahim S, Smith PM, Kristman V, Beaton DE, Mustard CA. The role of sex, gender, health factors, and job context in workplace accommodation use among men and women with arthritis. Ann Work Expo Health 2018;62:490–504.
18. Martin JC, Lewchuk W. The generation effect: Millennials, employment precarity and the 21st century workplace. Hamilton (ON): Poverty and Employment Precarity in Southern Ontario; 2018. URL: https://pepso.ca/documents/the-generation-effect-full-report.pdf.
19. Gignac MA, Cao X. "Should I tell my employer and co-workers I have arthritis?" A longitudinal examination of self-disclosure in the workplace. Arthritis Rheum 2009;61:1753–61.
20. Jetha A, Bowring J, Tucker S, Connelly CE, Martin Ginis KA, Proulx L, et al. Transitions that matter: life course differences in the employment of adults with arthritis. Disabil Rehabil 2018;40:3127–35.
21. Gignac MA, Backman CL, Davis AM, Lacaille D, Cao X, Badley EM. Social role participation and the life course in healthy adults and individuals with osteoarthritis: are we overlooking the impact on the middle-aged? Soc Sci Med 2013;81:87–93.
22. Carstensen LL. Social and emotional patterns in adulthood: support for socioemotional selectivity theory. Psychology and Aging 1992; 7:331.
23. Hendricks J. Considering life course concepts. J Gerontol B Psychol Sci Soc Sci 2012;67:226–31.
24. Gignac MA, Bowring J, Jetha A, Beaton DE, Breslin FC, Franche RL, et al. Disclosure, privacy and workplace accommodation of episodic disabilities: organizational perspectives on disability communication-support processes to sustain employment. J Occup Rehabil 2021; 31:153–65.
25. Sacks JJ, Harrold LR, Helmick CG, Gurwitz JH, Emani S, Yood RA. Validation of a surveillance case definition for arthritis. J Rheumatol 2005;32:340–7.
26. Gignac MA, Jetha A, Bowring J, Beaton DE, Badley EM. Management of work disability in rheumatic conditions: a review of non-pharmacological interventions. Best Pract Res Clin Rheumatol 2012;26:369–86.
27. Jetha A, Tucker L, Bowring J, Backman CL, Proulx L, Kristman V, et al. Casting a wide net: comparing strategies for recruiting

18-35-year-olds with rheumatic disease as study participants [abstract]. Annual Scientific Meeting of the American College of Rheumatology, Atlanta, GA, November 2011.

28. Reilly MC, Gooch KL, Wong RL, Kupper H, Van der Heijde D. Validity, reliability and responsiveness of the Work Productivity and Activity Impairment Questionnaire in ankylosing spondylitis. Rheumatology (Oxford) 2010;49:812–9.
29. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: visual analog scale for pain (VAS pain), numeric rating scale for pain (NRS pain), McGill Pain Questionnaire (MPQ), short-form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 bodily pain scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S240–52.
30. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). J Psychosom Res 2005;58: 163–71.
31. Gignac MA, Cao X, Tang K, Beaton DE. Examination of arthritis-related work place activity limitations and intermittent disability over four-and-a-half years and its relationship to job modifications and outcomes. Arthritis Care Res (Hoboken) 2011;63:953–62.
32. Yu C, Yao W. Robust linear regression: a review and comparison. Commun Stat Simul Comput 2017;46:6261–82.
33. SAS 9.3. Version 9. 3rd ed. Cary (NC): SAS Institute; 2015.
34. Sawyer SM, Afifi RA, Bearinger LH, Blakemore SJ, Dick B, Ezeh AC, et al. Adolescence: a foundation for future health. Lancet 2012;379:1630–40.
35. Stuth S, Jahn K. Young, successful, precarious? Precariousness at the entry stage of employment careers in Germany. J Youth Studies 2020;23:702–25.
36. Lindsay S, Cagliostro E, Carafa G. A systematic review of workplace disclosure and accommodation requests among youth and young adults with disabilities. Disabil Rehabil 2018;40:2971–86.
37. Tang K, Escorpizo R, Beaton DE, Bombardier C, Lacaille D, Zhang W, et al. Measuring the impact of arthritis on worker productivity: perspectives, methodologic issues, and contextual factors. J Rheumatol 2011;38:1776–90.
38. Mancuso CA, Paget SA, Charlson ME. Adaptations made by rheumatoid arthritis patients to continue working: a pilot study of workplace challenges and successful adaptations. Arthritis Care Res (Hoboken) 2000;13:89–99.
39. Hoving JL, van Zwieten MC, van der Meer M, Sluiter JK, Frings-Dresen MH. Work participation and arthritis: a systematic overview of challenges, adaptations and opportunities for interventions. Rheumatology (Oxford) 2013;52:1254–64.
40. Bonaccio S, Connelly CE, Gellatly IR, Jetha A, Ginis KA. The participation of people with disabilities in the workplace across the employment cycle: employer concerns and research evidence. J Bus Psychol 2020;35:135–58.