ARP Announcements

Association of Rheumatology Professionals
2200 Lake Boulevard NE, Atlanta, Georgia 30319
www.rheumatology.org

ACR/ARP Annual Meeting

November 8–13, 2019, Atlanta

Download the New ACR Publications Mobile App

The brand-new ACR Publications app can be downloaded for free from the Apple store or Google Play. ACR members can log in for full-text access to all articles in A*rthritis Care & Research* and *Arthritis & Rheumatology*. Nonmembers can access abstracts of all *AC&R* and *A&R* articles, the full text of articles published more than one year ago, and select open-access articles published recently, as well as the full text of all articles from *ACR Open Rheumatology* and *The Rheumatologist*.

New Division Name

Rheumatology is truly a people specialty; We often develop lifelong relationships with our patients as well as our colleagues. We increasingly recognize that providing the best rheumatologic care requires a team effort. The collegial nature of our specialty is reflected in the ACR's mission statement: To empower rheumatology professionals to excel in their specialty.

In keeping with this mission, we are pleased to announce that our health professionals' membership division is changing its name to Association of Rheumatology Professionals (ARP). This name change highlights the dedication of the ACR to serve the entire rheumatology community. It also reflects our broadened base of interprofessional members (administrators, advanced practice nurses, health educators, nurses, occupational therapists, pharmacists, physical therapists, physician assistants, research teams, and more).

The name is new, but our commitment and promise remain the same: We are here for you, so you can be there for your patients.

ARP Membership

The Association of Rheumatology Professionals (ARP), a division of the American College of Rheumatology, appreciates your continued membership and looks forward to serving you another year. Membership costs range from $30 to $140. ARP welcomes nurse practitioners, nurses, physician assistants, office staff, researchers, physical therapists, occupational therapists, assistants, and students. Student membership is complimentary; the Annual Meeting registration fee is waived for students who submit the required student verification letter. For information, go to www.rheumatology.org and select "Membership" or call 404-633-3777 and ask for an ARP staff member.

New ACR Journal Twitter Account (@ACR_Journals) and Social Media Editor

The ACR journals are heightening our focus on social media, to benefit authors and readers. Among our first activities is the introduction of an official ACR Journals Twitter account: @ACR_Journals. Followers will enjoy special features and the opportunity to engage with authors and other fellow professionals about studies published in *Arthritis Care & Research, Arthritis & Rheumatology,* and *ACR Open Rheumatology*. Authors of published articles will have the opportunity to use @ACR_Journals to share their work and engage in dialogue with others interested in the research. The journals welcome Dr. Paul Sufka of Minneapolis as our first Social Media Editor.

Arthritis Care & Research

Aims and Scope

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

Volume 71, No. 9, September 2019

Arthritis Care & Research

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

Cover design: Sandra Pulmano

♾ This journal is printed on acid-free paper.

Arthritis Care & Research

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

VOLUME 71 • September 2019 • NO. 9

Special Articles

Systemic Sclerosis

Osteoarthritis

Rheumatoid Arthritis

Cover image: The figure on the cover (from Oatis et al, page 1171) shows the use of an inclinometer to measure knee flexion range of motion.

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1151–1153
DOI 10.1002/acr.23860

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

African Americans and Scleroderma: Examining the Root Cause of the Association

Nadia D. Morgan and Allan C. Gelber

Among the autoimmune rheumatic disorders, systemic sclerosis (scleroderma) too often signals an unfavorable prognosis. Even with the availability of angiotensin-converting enzyme inhibitor therapy to abort renal crisis, of mycophenolate mofetil and cyclophosphamide to ameliorate skin fibrosis and pulmonary alveolitis, and of stem cell transplantation for aggressive disease, the course of individuals who have scleroderma remains particularly severe (1–4).

In this context, are outcomes associated with a new diagnosis of scleroderma comparable among all affected individuals? The answer is a clear and unequivocal no! Not only do patients who manifest the diffuse cutaneous subset of disease experience a more severe course, but so do affected individuals of African American race (5). The next sequence of questions one may pose to understand this relationship further might include what evidence supports this association. Further, what are the factors and mechanisms of disease that contribute to the elevated risk for greater morbidity and heightened mortality among African Americans who have scleroderma (6)?

In fact, a focus on race as a predictor of disease phenotype, serologic profile, and outcome has been a sustained area of research during the last 50 years. In contrast, when Sir William Osler published his experience in 1901 regarding 11 patients with scleroderma at Johns Hopkins, he recognized that women were predominantly affected, that extreme cyanosis of the hands was an important clinical feature, and that the lungs and kidneys were critically affected, but he made no comment about the racial composition of his patients (7). Subsequently, in a larger experience reported a half-century later, in 1954, which detailed 150 patients with scleroderma at New York University, again it was noted that scleroderma predominantly occurred in women; 108 of these patients (72%) were female (8). These authors noted a marked variability in the expression and severity of disease, yet only 3 of their patients were African American.

Thereafter, a 1967 report from Baltimore specifically focused on the racial and sex composition among fatal cases with scleroderma (9). This appears to have been the first report to furnish race-specific mortality rates. A total of 53 deaths occurred during the period of 1949 through 1963, among whom 20 patients (38%) were African American. In fact, the mortality rate among the African American women was 3-fold greater than among the white counterparts. In addition, the mean age at death was younger among the African American patients. Many years later, directly related to the study by Moore et al in this issue of *Arthritis Care & Research* (10), the question remains: why are African Americans disproportionately affected with greater morbidity and excess mortality once diagnosed with scleroderma?

In 1971, a population-based study from Shelby County, Tennessee examined the epidemiology of systemic sclerosis (11). In this community encompassing Memphis, the incidence of scleroderma was greater among African Americans compared to white residents. In particular, African American women experienced earlier onset of disease compared to white women and to men. More recently, in 1997, 514 women with scleroderma were identified from across the state of Michigan (12). In that report, half the African American women manifested the diffuse cutaneous subset of disease compared to one-fourth of the white residents. The investigators highlighted meaningful sex-race variability in disease incidence, with 22.5 cases of incident scleroderma among African American women compared to 12.8 cases per million among white women. The investigators further noted serologic differences between the groups and a 2-fold greater risk of mortality in Michigan among the African American female residents of the state compared to the white women. Population-based estimates of scleroderma incidence, prevalence, and mortality have similarly been furnished from South Carolina (13). There, the 5-year in-hospital mortality rate was greater at 23% among the African American residents compared to 16% of the white residents with scleroderma.

Dr. Morgan's work was supported by the Rheumatology Research Foundation's Scientist Development Award and Investigator Award, and by the Staurulakis Family Discovery Fund. Dr. Gelber's work was supported by the Maryland Chapter Arthritis Foundation Investigator Award.

[1]Nadia D. Morgan, MBBS, MHS, Allan C. Gelber, MD, MPH, PhD: Johns Hopkins University School of Medicine, Baltimore, Maryland.

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

Address correspondence to Allan C. Gelber, MD, MPH, PhD, Johns Hopkins University School of Medicine, 5200 Eastern Avenue, Mason F. Lord Building, Center Tower, Suite 4100, Baltimore MD 21224. E-mail: agelber@jhmi.edu.

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

Further, multivariate analyses integrated measures of educational attainment and median household income in the effort to discern which factors, environmental, genetic, or otherwise, contributed to the observed 70% increase in mortality among the African American patients with scleroderma.

Just a decade ago, genotyping of HLA class II alleles was studied, together with clinical and laboratory parameters, as predictors of scleroderma disease manifestations and outcome (14). When HLA II alleles (DRB1*0802 and DQA1*0501) were integrated into multivariate analyses in the Genetics versus Environment and Scleroderma Outcome Study, these genetic markers were predictive of mortality. In contrast, the risk estimate associated with race and ethnicity was no longer statistically significant. Consequently, the authors concluded that risk associated with ethnicity was mediated by HLA II gene allele status.

In this issue of the journal, Moore et al examine the contribution of socioeconomic determinants of health to explain excess mortality among African Americans with scleroderma, using data assembled from 2008 to 2016 at a single university medical center in the District of Columbia (10). Race and ethnicity were self-reported. Next, the patients were dichotomized as being African American or not. Within these 2 groups, all patients were matched by sex, age, date at first clinical evaluation at the scleroderma center, disease duration, and in the disease subset (limited versus diffuse cutaneous). Next, a series of analytical models, using Cox proportional hazards analysis, incorporated marital status, educational attainment, health insurance status, employment, and median household income, together with racial group, as predictors of outcome.

There were 402 patients in the analysis, among whom 202 were African American. They were predominantly women, in their early 40s at the onset of scleroderma, and with average disease duration of 8 years when first evaluated at the academic center. Further, notwithstanding matching by scleroderma disease subset, the prevalence of diffuse cutaneous disease was greater, at 48% versus 41%, in the African American patients. Notably, the frequency of seropositivity to anticentromere antibody was lower among the African American patients. These differences in clinical and serologic features of disease between groups are noteworthy in terms of their impact as predictors of outcome. Consequently, an imbalance in these features may influence the risk associated with race in relation to mortality. In addition, there was variability in the comorbidity profile, with a higher frequency of stroke and hypertension, and less frequent occurrence of malignancies, among the African American patients. In terms of clinical features of disease, renal crisis as well as cardiac and pulmonary disease occurred more commonly among the African American patients.

In terms of demographic and socioeconomic profile, the African American patients were more often single and less often married. In addition, the African American patients were more often disabled at the first visit to the scleroderma facility, were less likely to have private health insurance, and less frequently had attained a college degree. The median household income as assessed by zip code of residence was $23,000 lower among the African American patients, at $74,000.

Mortality in this cohort with scleroderma differed by race. Specifically, 43 of the African American patients (21%) died during the period of follow-up compared to 22 (11%) of the non–African American patients. Notably, mean follow-up was comparable in both groups. Perhaps the most critical table in the article is Table 5 (and the related figure), which highlights the risk of mortality associated with race. In univariate analysis, the African American patients experienced a 2-fold increased risk of mortality that was statistically significant. Yet when the socioeconomic determinants were included in the multivariate model, specifically integrating marital status, educational status, employment and health insurance status, the risk ratio for African American race was reduced to 1.8 and was no longer statistically significant. In the final multivariate model, in which household income was further integrated, this parameter was predictive of outcome, whereas the association with race was markedly reduced, approaching null.

Quite recently, a large multicenter cohort was assembled from 18 academic medical centers across the US to examine the demographic, clinical, socioeconomic, and genetic contribution to scleroderma susceptibility (15). In a detailed and systematic fashion, the Genome Research in African American Scleroderma Patients (GRASP) cohort was established in 2013 to evaluate phenotype, susceptibility, and outcome among African Americans with scleroderma. Interestingly, the participating patients were enrolled both retrospectively and prospectively, with disease onset commencing as early as 1987 and continuing with new recruitment through 2016. The majority of the 1,009 African American patients in the GRASP cohort manifested the diffuse cutaneous subset of disease, an observation evident among both women and men (15). Overall, 94% were antinuclear antibody positive, whereas 30% were seropositive for topoisomerase antibody, the latter finding was more often observed among men (38%) than among women (28%). In addition, 43% had radiographic evidence of pulmonary fibrosis on computed tomography imaging, with 30% manifesting pulmonary hypertension at echocardiographic or cardiac catheterization assessment. In all, 7% of the cohort experienced renal crisis.

Next, the GRASP investigators undertook whole-exome gene sequencing to identify functional variants previously associated with susceptibility to scleroderma among individuals of European American ancestry (16). They compared genetic measures, particularly coding and deleterious variants related to fibrosis, among 379 participants from the GRASP cohort to 411 healthy controls derived from the Howard University Family Study, a population-based African American cohort in Washington, DC. Among the various candidate genes they examined, the single pathway that retained statistical significance, after adjustment for multiple comparisons, was the hepatic fibrosis/hepatic stellate cell gene activation pathway. Further, this association was observed specifically

among those with the diffuse cutaneous subset of scleroderma in comparison to healthy controls (16).

Overall, and in the context of these published reports that underscore the disproportionate and adverse impact of scleroderma among African Americans, and in light of the ongoing efforts of the GRASP study, the current article by Moore et al emphasizes the importance of socioeconomic status and of socioeconomic determinants of health, to account for differences in clinically relevant outcomes (10). These reports emphasize to the reader, each in its own way, the value of an optimal study design, which can concomitantly examine the independent contribution of socioeconomic status, and of clinical, serologic, and genetic determinants to outcomes of interest in scleroderma. Such opportunities, made possible by large multicenter collaborations with these key parameters integrated into the study design, are the ultimate goal.

AUTHOR CONTRIBUTIONS

Drs. Morgan and Gelber drafted the editorial. Dr. Gelber revised it critically for important intellectual content, and approved the final version to be published.

REFERENCES

1. Lopez-Ovejero JA, Saal SD, D'Angelo WA, Cheigh JS, Stenzel KH, Laragh JH. Reversal of vascular and renal crises of scleroderma by oral angiotensin-converting-enzyme blockade. N Engl J Med 1979;300:1417–9.
2. Tashkin DP, Roth MD, Clements PJ, Furst DE, Khanna D, Kleerup EC, et al. Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLS II): a randomised controlled, double-blind, parallel group trial. Lancet Respir Med 2016;4:708–19.
3. White B, Moore WC, Wigley FM, Xiao HQ, Wise RA. Cyclophosphamide is associated with pulmonary function and survival benefit in patients with scleroderma and alveolitis. Ann Intern Med 2000;132:947–54.
4. Van Laar JM, Farge D, Sont JK, Naraghi K, Marjanovic Z, Larghero J, et al. Autologous hematopoietic stem cell transplantation vs intravenous pulse cyclophosphamide in diffuse cutaneous systemic sclerosis: a randomized clinical trial. JAMA 2014;311:2490–8.
5. Gelber AC, Manno RL, Shah AA, Woods A, Le EN, Boin F, et al. Race and association with disease manifestations and mortality in scleroderma: a 20-year experience at the Johns Hopkins Scleroderma Center and review of the literature. Medicine (Baltimore) 2013;92:191–205.
6. Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. N Engl J Med 2009;360:1989–2003.
7. Osler W. The principles and practice of medicine. 4th ed. New York: Appleton; 1901.
8. Leinwand I, Duryee AW, Richter MN. Scleroderma (based on a study of over 150 cases). Ann Intern Med 1954;41:1003–41.
9. Masi AT, D'Angelo WA. Epidemiology of fatal systemic sclerosis (diffuse scleroderma): a 15-year survey in Baltimore. Ann Intern Med 1967;66:870–83.
10. Moore DF, Kramer E, Eltaraboulsi R, Steen VD. Increased morbidity and mortality of scleroderma in African Americans compared to non–African Americans. Arthritis Care Res (Hoboken) 2019;71:1154-63.
11. Medsger TA Jr, Masi AT. Epidemiology of systemic sclerosis (scleroderma). Ann Intern Med 1971;74:714–21.
12. Laing TJ, Gillespie BW, Toth MB, Mayes MD, Gallavan RH Jr, Burns CJ, et al. Racial differences in scleroderma among women in Michigan. Arthritis Rheum 1997;40:734–42.
13. Nietert PJ, Silver RM, Mitchell HC, Shaftman SR, Tilley BC. Demographic and clinical factors associated with in-hospital death among patients with systemic sclerosis. J Rheumatol 2005;32:1888–92.
14. Assassi S, del Junco D, Sutter K, McNearney TA, Reveille JD, Karnavas A, et al. Clinical and genetic factors predictive of mortality in early systemic sclerosis. Arthritis Rheum 2009;61:1403–11.
15. Morgan ND, Shah AA, Mayes MD, Domsic RT, Medsger TA Jr, Steen VD, et al. Clinical and serological features of systemic sclerosis in a multicenter African American cohort: analysis of the Genome Research in African American Scleroderma Patients clinical database. Medicine (Baltimore) 2017;96:e8980.
16. Gourh P, Remmers EF, Boyden SE, Alexander T, Morgan ND, Shah AA, et al. Brief report: whole-exome sequencing to identify rare variants and gene networks that increase susceptibility to scleroderma in African Americans. Arthritis Rheumatol 2018;70:1654–60.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1154–1163
DOI 10.1002/acr.23861

Increased Morbidity and Mortality of Scleroderma in African Americans Compared to Non–African Americans

Duncan F. Moore, Elisabeth Kramer, Rami Eltaraboulsi, and Virginia D. Steen

Objective. African Americans with scleroderma have more severe disease and higher mortality than non–African Americans. Differences in rates of diffuse disease, autoantibody status, or socioeconomic status have not completely explained this phenomenon. Our study evaluates these risks at our site.

Methods. A retrospective study comparing African American and non–African American patients with scleroderma seen from 2008 to 2016 was performed. Groups were matched by sex, age at first visit, date of first visit, disease duration at first visit, and limited versus diffuse cutaneous disease. Demographic, serologic, and clinical features were compared. Mortality risks were assessed by a Cox proportional hazards model with covariates of race, marital status, education, employment, insurance, and imputed household income.

Results. African Americans comprised 202 of 402 patients. They demonstrated reduced forced vital capacity and diffusing capacity for carbon monoxide, more severe lung fibrosis, a higher prevalence of pulmonary hypertension, and more severe cardiac involvement. The autoantibody profile statistically differed between the 2 groups. Death during follow-up was 21% in African Americans versus 11% in non–African Americans ($P = 0.005$). African American race demonstrated an unadjusted hazard ratio for death during follow-up of 2.061 ($P = 0.006$) that declined with adjustment for socioeconomic covariates to 1.256 ($P = 0.633$). The only significant covariate was median income in tens of thousands of dollars by zip code (hazard ratio 0.845; $P = 0.033$).

Conclusion. African American patients with scleroderma have more severe pulmonary disease and higher unadjusted mortality than matched non–African Americans. Following adjustment for socioeconomic factors, African American race was not a significant risk factor for mortality; however, independent of race, a lower median household income predicted increased mortality.

INTRODUCTION

African Americans with scleroderma (SSc; systemic sclerosis) experience more severe manifestations, a more aggressive disease course, and increased morbidity and mortality relative to non–African Americans. Prior studies have demonstrated that the incidence and prevalence of SSc are higher in African Americans (1–3). African Americans are younger at disease onset (4,5) and at time of diagnosis (2,5). They are more likely to have diffuse cutaneous disease (1,2,5–7) and demonstrate an increased prevalence (5,6) and severity of restrictive lung disease (5,8,9). Cardiac and renal involvement are also more prevalent (6), and there is an increased prevalence and severity of skeletal muscle involvement (5). African American patients are more likely than white patients to have anti–Scl-70 (anti–topoisomerase) and anti–U1 RNP antibodies and less likely to have anticentromere antibodies (5,6). Anti–Scl-70 is an independent risk factor for SSc lung disease (8). Younger African Americans with SSc are also hospitalized more frequently (10).

African American patients with SSc have increased mortality relative to non–African Americans (1,5,6,11–13). Mortality rates are higher among patients with diffuse skin disease or with pulmonary involvement (1). Steen et al adjusted mortality for age, sex, and diffuse disease status and found that African Americans were 1.68 times more likely to die during follow-up. This mortality difference persisted in a subgroup analysis of anti–Scl-70–positive patients. African Americans within this subgroup demonstrated increased prevalence and severity of pulmonary fibrosis and an increased hazard ratio for death (5). Gelber et al (6) adjusted their survival analyses for sex, duration, disease subtype, and either SSc-specific antibody status, educational attainment, or health insurance status and found that African American race remained an independent risk factor for mortality.

Duncan F. Moore, MD, Elisabeth Kramer, MD, Rami Eltaraboulsi, MD, Virginia D. Steen, MD: MedStar Georgetown University Hospital, Washington, DC.

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

Address correspondence to Virginia D. Steen, MD, Division of Rheumatology, Department of Medicine, MedStar Georgetown University Hospital, 3800 Reservoir Road, PHC 3004, Washington, DC 20007. E-mail: steenv@georgetown.edu.

Submitted for publication June 19, 2018; accepted in revised form February 19, 2019.

SIGNIFICANCE & INNOVATIONS

- As demonstrated in this previously undescribed cohort, African Americans with scleroderma have a distinct clinical and serologic profile, most notably more severe pulmonary disease and higher unadjusted mortality, relative to non–African Americans.
- African American patients had lower measures of socioeconomic status by all variables examined.
- After adjustment for available socioeconomic factors, African American race was not an independent risk factor for increased mortality. However, independent of race, lower household income (as imputed by zip code) was associated with increased mortality during follow-up.

While African American patients with SSc have a distinct phenotypic and serologic profile, their experience of and outcomes in the disease may be confounded by socioeconomic factors that correlate with race (14). Within SSc cohorts, African Americans have decreased educational attainment, increased Medicaid prevalence, decreased household income, and decreased vehicle ownership relative to non–African Americans (12,15,16). Marital status, employment, educational attainment, insurance status, and median income by household have all been examined as covariates in SSc mortality research (6,12,15,16). Effect sizes have varied, and socioeconomic factors have not fully explained racial disparities in any studies to date.

In our own study, we sought to augment the existing literature by describing a large set of African American patients with SSc compared with matched non–African American patients with SSc. Patients were matched in order to control for sex, age, disease duration, and diffuse disease status. We also sought to examine the interplay of race and socioeconomic status in SSc survival via a robust series of regression analyses.

PATIENTS AND METHODS

Study population. The study population was drawn retrospectively from the outpatient panel of the senior author (VDS) at a single academic center. Patients with SSc seen between 2008 and 2016 were eligible for inclusion. The clinical diagnosis of SSc was made by the senior author. At inclusion, African American and non–African American patients were matched by the senior author by the following 5 characteristics: sex, age (within 5 years) at first visit within the study period, date (either 2008–2012 or 2013–2016) of first visit within the study period, disease duration (either ≤4 years or >4 years) at first visit within the study period, and limited versus diffuse cutaneous disease.

Data collection. All data were abstracted via a manual review of printed and electronic charts according to a standardized form. The study was carried out in compliance with the Helsinki Declaration and was approved by the Georgetown University Institutional Review Board. All authors performed abstraction. Race and ethnicity were self-reported. Our institution's electronic health record allows for the designation of "black or African American," and hereafter all such patients are referred to as "African American." Reported marital status, educational attainment, employment, and insurance status refer to values collected at the index visit. If the patient's index visit occurred before 2008, the date of the index visit was changed to January 1, 2008. Clinical comorbidities were abstracted from all available clinical data.

Median household income was imputed onto patients by the zip code of their residence at the initial visit, per data from the US Census Bureau collected from 2006 to 2010 (17). Income imputed by zip code has been used widely as a socioeconomic status surrogate (18), including as an independent variable in the SSc literature (12,16) and also as a component of a composite score of relative socioeconomic deprivation (19,20).

Disease duration was noted from the first symptom attributable to SSc, including Raynaud's phenomenon. Clinical features, such as modified Rodnan skin score, were recorded by most abnormal value or status. Comorbid diseases, such as hypertension, diabetes mellitus, and malignancy were recorded as positive if present at any point during the disease course. Severe cardiac involvement was defined as symptomatic pericardial effusion, arrhythmias requiring treatment, or heart failure requiring treatment. Laboratory measures and objective clinical measurements, such as pulmonary function tests (PFTs) or values from right heart catheterization, were also recorded by most abnormal value or status. Mortality during follow-up was obtained by review of patient records and the Social Security Death Index.

Statistical analysis. Patients' demographic, laboratory, and clinical features were compared with Student's *t*-test, chi-square test, or Fisher's exact test, as appropriate. Mortality risks were assessed by a Cox proportional hazards model with sociodemographic and socioeconomic covariates of race, marital status, educational attainment, insurance status, employment status, and median household income imputed by zip code (in tens of thousands of dollars). Analyses were performed with SPSS statistics software, version 19.0.

RESULTS

Baseline characteristics and matching. African American patients comprised 202 of the 402 patients analyzed. The 200 non–African American patients included 193 (48%) who identified as white or Caucasian, 4 (1%) who identified as Asian, and 3 (1%) who identified as "other." Of the 12 patients (3%) who identified as Hispanic or Latino, 1 was African American, and 11 were non–African American.

Table 1. Characteristics for matching African American and non–African American patients at study inclusion*

Characteristics	African American (n = 202)	Non–African American (n = 200)	*P*
Female	175 (87)	174 (87)	0.914
Age at first visit, mean ± SD years	47.5 ± 13.2	48.4 ± 13.1	0.481
Disease duration at first visit, mean ± SD years	7.7 ± 8.1	8.3 ± 9.6	0.512
Date of first visit			0.372
2008–2012	133 (66)	140 (70)	
2013–2016	69 (34)	60 (30)	
Scleroderma type			0.103
Diffuse	97 (48)	81 (41)	0.129
Limited	101 (50)	118 (59)	0.070
Unclassified†	4 (2)	1 (1)	0.181

* Values are the number (%) unless indicated otherwise.
† Patients not followed early in the disease course and for whom the presence of limited cutaneous disease could not be confirmed (as in patients with diffuse disease that later appears limited due to softening of the skin).

As a group, African Americans were similar to non–African Americans in all factors by which they were matched (Table 1). Both groups were majority female (87%), in the fifth decade of life on average at first visit, and with disease duration at first visit of approximately 8 years. Age at symptom onset was approximately 40 ± 14 years (mean ± SD). The first visit occurred between 2008 and 2012 in approximately two-thirds of both groups. African American patients trended toward a higher prevalence of diffuse disease (48% versus 41%) than non–African American patients, but this group difference was not statistically significant.

There were no group differences in frequency of follow-up. The mean number of follow-up appointments was 7.4 in African Americans and 7.1 in non–African Americans ($P = 0.657$). The number of visits per year was 3.6 in African Americans and 3.4 in non–African Americans ($P = 0.912$).

The 2013 American College of Rheumatology/European League Against Rheumatism classification criteria for SSc (21) were fulfilled in 186 (92%) and 184 (92%) of African American and non–African American patients, respectively. Baseline non-SSc characteristics are shown in Table 2. African American patients had a higher prevalence of stroke and hypertension, and non–African American patients had a higher prevalence of cancer, primarily non-melanoma skin cancer. The prevalence of coronary artery disease, diabetes mellitus, and chronic obstructive pulmonary disease did not differ among the 2 groups. A similar prevalence of smoking was observed in both groups.

Socioeconomic status. There were differences among the 2 groups in all socioeconomic domains examined (Table 2). African American patients were more likely to be single and less likely to be married. African American patients were more likely to be disabled at the first visit and less likely to be a homemaker. African American patients were less likely to have private insurance and more likely to have Medicaid, although only 8% had Medicaid. Even though fewer African American patients had a college degree, 58% of all patients whose education status was known had college or postgraduate degrees (29% overall). African American patients were more likely to be in the fourth quintile of median household income by zip code and less likely to be in the fifth quintile (highest level). The median household income by zip code for African American patients was $74,000, which was $23,000 lower than that of non–African American patients ($P < 0.001$). In comparison, the nationwide mean of the median income of all zip codes was $63,000.

Laboratory markers. Serologic and inflammatory markers differed among the 2 groups (Table 3). Both groups had a similar prevalence of antinuclear antibody (ANA) positivity (defined as >1:80, or 1:80 with a positive SSc-specific autoantibody), with approximately 94% of patients tested. Not all patients were tested for all SSc-specific antibodies. Because the vast majority of patients with SSc have only 1 SSc-specific antibody, additional testing for SSc-specific antibodies was not performed after 1 was positively identified in a patient. Isolated nucleolar ANA was defined as a positive ANA with an isolated nucleolar pattern without a positive test for an SSc-specific autoantibody. "Other ANA" referred to a positive ANA without an isolated nucleolar pattern in a patient without any other SSc-specific antibodies or to patients with multiple autoantibodies (determined by the senior author [VDS] not to be low-titer false positives). All such autoantibody categories were mutually exclusive. African American patients were more likely to have an isolated nucleolar ANA, anti–U1 RNP antibody, or other positive ANA without an

Table 2. Baseline medical and socioeconomic characteristics, by race*

Characteristics	African American (n = 202)	Non–African American (n = 200)	*P*
Tobacco smoking			0.633
Never	130 (64)	123 (62)	0.553
Former	47 (23)	52 (26)	0.525
Present	14 (7)	10 (5)	0.414
Not recorded	11 (5)	15 (8)	0.402
Clinical comorbidities			
Coronary artery disease	13 (6)	9 (5)	0.394
Stroke	6 (3)	0 (0)	0.014†
Hypertension	88 (43)	28 (14)	< 0.001†
Diabetes mellitus	19 (9)	10 (5)	0.088
COPD	5 (3)	4 (2)	0.747
Malignancy	18 (9)	37 (19)	0.005†
Marital status			0.001†
Single	66 (33)	33 (17)	< 0.001†
Married	86 (43)	121 (61)	< 0.001†
Separated	5 (3)	4 (2)	0.747
Divorced	14 (7)	19 (10)	0.348
Widowed	9 (5)	7 (4)	0.624
Not recorded	22 (11)	16 (8)	0.322
Employment at first visit			0.007†
Working	109 (54)	109 (55)	0.914
Retired	18 (9)	30 (15)	0.060
Homemaker	2 (1)	12 (6)	0.006†
Student	6 (3)	9 (5)	0.418
Disabled	21 (10)	9 (5)	0.024†
Unemployed	13 (6)	9 (5)	0.394
Not recorded	33 (16)	22 (11)	0.120
Insurance			0.001†
Private	149 (74)	164 (82)	0.047†
Medicare‡	32 (16)	25 (13)	0.337
Medicaid§	16 (8)	1 (1)	< 0.001†
Self-pay/assistance	2 (1)	6 (3)	0.149
Not recorded	3 (1)	4 (2)	0.693
Education (highest attained)			0.013†
Grade school	7 (3)	4 (2)	0.368
High school	28 (14)	15 (8)	0.039†
Some college¶	7 (3)	14 (7)	0.111
College	23 (11)	43 (22)	0.006†
Post-college#	26 (13)	25 (13)	0.911
Not recorded	111 (55)	99 (50)	0.274
National quintile of median household income of zip code at first visit**			< 0.001†
First (lowest)	24 (12)	6 (3)	0.001†
Second	4 (2)	6 (3)	0.512
Third	2 (1)	9 (5)	0.031†
Fourth	51 (26)	16 (8)	< 0.001†
Fifth (highest)	119 (60)	160 (81)	< 0.001†
Median income of zip code at first visit, mean ± SD tens of thousands of dollars**	7.41 ± 2.65	9.74 ± 3.74	< 0.001†

* Values are the number (%) unless indicated otherwise. COPD = chronic obstructive pulmonary disease.
† Statistically significant.
‡ Medicare is the US Federal government health insurance program for people age ≥65 years, certain younger people with disabilities, and people with end-stage renal disease.
§ Medicaid is a joint Federal and state program in the US that provides health insurance to low-income people.
¶ Trade school, technical education, college courses not leading to degree.
# Includes graduate and professional school.
** Data analyzed from 200 African Americans and 197 non–African Americans.

Table 3. Serologic and laboratory characteristics, by race*

Characteristics	African American (n = 202)	Non–African American (n = 200)	*P*
Autoantibody			< 0.001†
Negative ANA, no. (%)	12 (6)	10 (5)	0.678
Anticentromere	14/120 (7)	44/153 (22)	0.001†
Anti-Scl-70	43/166 (21)	41/163 (21)	0.876
Anti–U1 RNP	26/114 (13)	10/119 (5)	0.002†
Isolated nucleolar ANA	49/154 (24)	32/151 (16)	0.036†
Anti–RNA polymerase III	7/78 (3)	30/118 (15)	0.004†
Other ANA, no. (%)‡	45 (22)	25 (13)	0.010†
No result, no. (%)	6 (3)	8 (4)	0.573
SSA positive	44/135 (22)	13/124 (7)	< 0.001†
SSB positive	11/130 (5)	3/123 (2)	0.036†
ESR, mean ± SD	36.5 ± 31.0	26.4 ± 25.5	0.002†
CRP, mean ± SD	10.0 ± 18.8	8.3 ± 12.5	0.399
CPK, mean ± SD	591 ± 1302	229 ± 403	0.005†

* Values are the number/total (%) unless indicated otherwise. Percentages are of the total racial group. Not all tests were performed in all patients: erythrocyte sedimentation rate (142 African Americans, 159 non–African Americans), C-reactive protein (116 African Americans, 129 non–African Americans), creatine phosphokinase (114 African Americans, 114 non–African Americans). ANA = antinuclear antibody; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; CPK = creatine phosphokinase.
† Statistically significant.
‡ Includes patients with positive ANA but no scleroderma-specific autoantibody or patients with multiple autoantibodies (determined not to be low-titer false positives).

SSc-specific antibody and less likely to have anticentromere or anti–RNA polymerase III antibodies. The prevalence of anti–Scl-70 antibodies was the same among the 2 groups. African American patients were more likely to have positive SSA and SSB antibodies but had sicca symptoms less frequently (Table 4). The C-reactive protein level was similar in both groups, but the erythrocyte sedimentation rate and creatine phosphokinase (CPK) level were higher in African American patients.

Clinical features. Numerous clinical features differed among the 2 groups (Table 4). Raynaud's phenomenon and digital pitting scars were less likely to be present in African Americans. The prevalence of gastric antral vascular ectasia and sicca symptoms were higher among non–African American patients. Although anti–RNA polymerase III antibody positivity was more common in non–African Americans, renal crisis occurred more frequently in African Americans (8% versus 3% in non–African Americans; $P = 0.019$). There were no group differences in the presence or severity of proximal muscle weakness (despite higher CPK levels in African Americans) or the prevalence of neuropathy symptoms requiring gabapentinoids. There was no categorical group difference in severity of gastrointestinal disease; however, non–African Americans were more likely to have gastroesophageal reflux disease treated with medications. On examination, there were no group differences in mean skin score, prevalence of inflammatory arthritis, or prevalence of contractures. Calcinosis, telangiectasias, and tendon friction rubs were more common among non–African American patients.

African Americans patients had more severe pulmonary and cardiac disease. The mean lowest forced vital capacity and diffusing capacity for carbon monoxide were lower in African Americans, and correspondingly, the prevalence of fibrosis and severe fibrosis per computed tomography scan, home oxygen use, and pulmonary hypertension by right heart catheterization were higher among African American patients. Pulmonary hypertension was present in 21% of African American patients. Of those 42 patients, 33% had pulmonary arterial hypertension, 16% had pulmonary venous hypertension, and 20% had pulmonary hypertension secondary to interstitial lung disease (PH-ILD). Of the 25 non–African Americans (18%) who had pulmonary hypertension, 42% had pulmonary arterial hypertension, 14% had pulmonary venous hypertension, and 11% had PH-ILD. Overall, there was no categorical difference among the 2 groups in the type of pulmonary hypertension present ($P = 0.591$). On echocardiography, African American patients had higher pulmonary artery systolic pressure or right ventricular systolic pressure and lower left ventricular ejection fraction. Overall, severe cardiac involvement was more prevalent in African Americans.

Table 4. Clinical features, by race*

Features	African American (n = 202)	Non–African American (n = 200)	*P*
Raynaud's phenomenon			0.003†
None	11 (5)	3 (2)	0.031†
Raynaud's phenomenon only	99 (49)	93 (47)	0.614
With digital pitting scars	4 (2)	19 (10)	0.001†
With digital ulcerations	72 (36)	75 (38)	0.699
With digital gangrene	16 (79)	10 (5)	0.234
Sicca symptoms	30 (15)	64 (32)	< 0.001†
Gastric antral vascular ectasia	7 (3)	18 (9)	0.022†
Renal crisis	17 (8)	6 (3)	0.019†
Gastrointestinal involvement			0.304
No gastrointestinal meds	38 (19)	30 (15)	0.308
GERD with meds	130 (64)	148 (74)	0.036†
Antibiotics for bacterial overgrowth, or abnormal gastric emptying study	14 (7)	8 (4)	0.196
Pseudo-obstruction	6 (3)	3 (2)	0.319
Hospitalization for gastrointestinal disease	9 (5)	9 (5)	0.983
Requiring total parenteral nutrition	5 (3)	2 (1)	0.258
Examination features			
Skin score, mean ± SD	14.2 ± 13.2	14.4 ± 12.6	0.881
Inflammatory arthritis	33 (16)	43 (22)	0.186
Contractures	66 (33)	73 (37)	0.420
Calcinosis	9 (4)	34 (17)	< 0.001†
Telangiectasia	39 (19)	101 (51)	< 0.001†
Tendon friction rubs	24 (12)	41 (21)	0.019†
Pulmonary features			
Forced vital capacity, mean ± SD	68.4 ± 20.4	80.7 ± 19.4	< 0.001†
DLco, mean ± SD	45.8 ± 19.8	63.7 ± 20.9	< 0.001†
Fibrosis per CT scan			0.002†
None	38 (19)	60 (30)	0.009†
Mild/moderate	77 (38)	67 (34)	0.334
Severe	22 (11)	6 (3)	0.002†
Not performed	65 (32)	67 (34)	0.778
Home oxygen use	48 (24)	18 (9)	< 0.001†
Pulmonary hypertension (any type) by right heart catheterization	42 (21)	25 (18)	0.026†
Cardiac features			
PASP or RVSP, mean ± SD mm Hg	39.3 ± 17.2	32.8 ± 14.2	0.001†
LVEF, mean ± SD %	56.3 ± 11.2	59.5 ± 6.0	0.002†
Severe cardiac involvement‡	35 (17)	17 (9)	0.008†

* Values are the number (%) of the total racial group, unless indicated otherwise. Not all tests were performed in all patients. Forced vital capacity: 161 African Americans, 178 non–African Americans; diffusing capacity for carbon monoxide (DLco): 161 African Americans, 175 non–African Americans; pulmonary hypertension (any type) by right heart catheterization: 55 African Americans, 36 non–African Americans; pulmonary artery systolic pressure (PASP) or right ventricular systolic pressure (RVSP): 142 African Americans, 160 non–African Americans; left ventricular ejection fraction (LVEF); 142 African Americans, 160 non–African Americans. GERD = gastroesophageal reflux disease; CT = computed tomography.
† Statistically significant.
‡ Symptomatic pericardial effusion, arrhythmias requiring treatment, or heart failure requiring treatment.

African American patients were more likely to have received prednisone at any point during their illness (48% versus 33%; $P = 0.002$) and were more likely to have received prednisone doses of 15 mg or more (19% versus 11%; $P = 0.011$). There was no difference among the 2 groups in prevalence of treatment with cyclophosphamide, mycophenolate mofetil, methotrexate, azathioprine, or D-penicillamine (data not shown).

African American patients had increased mortality during the follow-up period. The mean follow-up duration was similar among the 2 groups (mean ± SD 4.1 ± 2.8 years in African Americans and 4.4 ± 3.1 years in non–African Americans; $P = 0.207$). Forty-three (21%) of the African American patients died, and 22 (11%) of the non–African American patients died ($P = 0.005$). SSc-related causes of death (composed of pulmonary fibrosis, pulmonary hypertension, cardiac disease, renal disease, and multisystem SSc) accounted for 31 deaths (72%) in African Americans and 17 deaths (77%) in non–African Americans. Cancer accounted for 7 deaths (16%) in African Americans and 5 deaths (23%) in non–African Americans ($P = 0.570$).

Survival analyses. A Cox proportional hazards model was performed to adjust mortality outcomes for socioeconomic status (Table 5). The proportional hazards assumption was met for all such analyses per Schoenfeld residuals (22). In Model A, African American status was the only covariate. African American status showed an unadjusted hazard ratio for death during follow-up of 2.061 (95% confidence interval [95% CI] 1.232–3.449; $P = 0.006$). Figure 1 shows a plot of the unadjusted survival function. For Model B, the covariates of the patient-specific socioeconomic factors of marital status, educational attainment, employment status, and insurance type were added. In Model B, the hazard ratio of African American status was reduced to 1.778 (95% CI 0.727–4.350; $P = 0.207$). None of the other socioeconomic covariates demonstrated statistically significant effects. For Model C, the covariate of household income in tens of thousands of dollars (imputed by zip code) was added to Model B. In Model C, the hazard ratio of African American status was reduced to 1.256 (95% CI 0.494–3.191; $P = 0.633$), the patient-specific socioeconomic covariates remained statistically insignificant, and the hazard ratio of median income was 0.845 (95% CI 0.723–0.986; $P = 0.033$). Thus, for every additional $10,000 of household income, independent of race, the hazard of death during follow-up declined by 15.5%.

Table 5. Survival analyses, by race*

Model and variables	Hazard ratio (95% CI)	*P*
A†	2.061 (1.232–3.449)	0.006‡
B§	1.778 (0.727–4.350)	0.207
C¶		
African American race	1.256 (0.494–3.191)	0.633
Household income	0.845 (0.723–0.986)	0.033‡

* Cox proportional hazard models were used to calculate the hazard ratio for death during follow-up. In models B and C, none of the socioeconomic covariates other than household income in tens of thousands of dollars (imputed by zip code) was statistically significant. 95% CI = 95% confidence interval.
† Model A: African American race.
‡ Statistically significant.
§ Model B: African American race, marital status, educational attainment, employment status, insurance type. Reported hazard ratio is that of African American race.
¶ Model C: African American race, marital status, educational attainment, employment status, insurance type, household income in tens of thousands of dollars (imputed by zip code).

DISCUSSION

We have demonstrated in a large, retrospective, matched cohort of African Americans and non–African Americans with SSc that African Americans have a different autoantibody profile, more severe pulmonary and cardiac disease, and a higher unadjusted risk of mortality than their non–African American counterparts. African Americans had lower measures of socioeconomic status by all examined variables, including the patient-specific measures of marital status, employment, insurance type, and educational attainment as well as imputed household income. After adjustment for these factors, African American race was not a significant risk factor for mortality, but lower median household income by zip code was an independent risk factor for increased mortality.

The baseline characteristics, disease manifestations, and outcomes within our cohort differed in several ways from prior cohorts. Unlike prior studies (1,5,6), there was no observed racial difference in age at the first visit, duration of disease (by first symptom, including Raynaud's phenomenon) at first visit, or prevalence of diffuse disease, which specifically reflected the matching by these features in our study design.

Similar to prior studies, African Americans in our study were more likely to have anti–U1 RNP antibodies and less likely to have anticentromere antibodies and anti–RNA polymerase III antibodies. Unlike some prior studies (5,6), our cohort did not demonstrate an increased prevalence of anti-Scl-70 antibodies among African American patients. Because anti–Scl-70 strongly correlates with diffuse skin disease, matching by diffuse versus limited cutaneous disease eliminated this difference.

Sicca symptoms were more common among non–African Americans, despite a higher prevalence of positive SSA and SSB antibodies among African Americans. This finding in African Americans could reflect an often seen overlap disease with systemic lupus erythematosus, rather than Sjögren's syndrome. CPK levels were higher in African American patients despite the absence of an increase in prevalence of proximal muscle weakness. This finding likely reflects the fact that CPK levels are higher in healthy asymptomatic African Americans when compared to non–African Americans (23).

African American patients had worse restrictive lung disease by PFTs, as seen previously (6,8,9). Independent of any disease

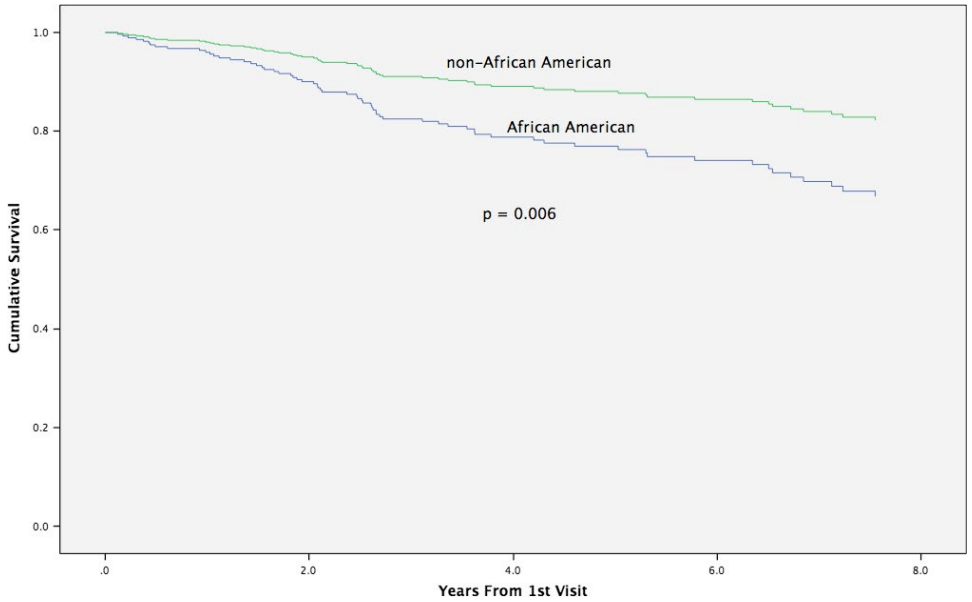

Figure 1. Unadjusted cumulative survival plot, stratified by race, not adjusted for socioeconomic covariates. African American status demonstrated an unadjusted hazard ratio for death during follow-up of 2.061 (95% confidence interval 1.232–3.449; $P = 0.006$).

process, African American patients have previously been noted to have lower lung volumes at baseline (24). PFT values reported in this study came from a variety of laboratories, so we are unable to determine what degree of race-based adjustment of lung volumes had occurred already; however, most PFTs were performed in our institution's laboratory, which does adjust for race. Regardless, some of our observed difference in patients with SSc was possibly explained by race alone. African American patients also had worse fibrosis on imaging, as seen previously (5). There was no racial difference in smoking history, even though prior cohorts have demonstrated higher (5) or lower (6) smoking rates among African American patients. In contrast to a prior adjusted analysis (6), pulmonary hypertension by right heart catheterization was slightly more prevalent among African Americans; however, right heart catheterization was performed in our study only as clinically indicated. Of note, many African American patients with pulmonary arterial hypertension in our study had both diffuse SSc and antinucleolar antibody positivity. This phenotype, consistent with prior descriptions (25), is important to recognize because it differs from the usual phenotype of a patient with limited cutaneous SSc with anticentromere antibody. As in prior studies, cardiac involvement was more severe among African Americans. However, the marked increase in underlying hypertension among African Americans in our study may have been contributory. Both groups received similar exposure to treatment drugs and were followed for similar durations. There was no group difference in the number or frequency of follow-up visits.

Our findings of increased unadjusted mortality are consistent with the aforementioned prior studies. No racial difference in survival was noted by Laing et al (1) after adjusting for age at diagnosis and diffuse versus limited disease status. Analyses of prior cohorts that controlled for age, sex, diffuse disease status, and autoantibody status showed a persistent elevation of mortality risk in African Americans (5,6). Additionally, a study of 1 cohort sought to control different individual socioeconomic factors in a series of individual analyses and showed that African American race remained an independent risk factor (6). In our final adjusted mortality analysis, age, sex, disease duration, diffuse disease status, and anti–Scl-70 status were controlled by matching, while the socioeconomic variables of marital status, insurance type, educational attainment, employment status, and imputed household income were all included as covariates. This model demonstrated that African American race was not a statistically significant independent mortality risk factor and that a lower household income increased the risk of death during follow-up. Although a history of hypertension or stroke was more common among African Americans in our study, the addition of these variables to our survival analyses did not change our results, nor was either comorbidity a significant predictor of mortality.

There were several limitations to this study. The matching process may have introduced unmeasured selection bias. The matches also demonstrated a trend toward more diffuse disease and less limited disease in African Americans, although actual mean skin scores were essentially identical. Because disease duration was defined as time from first manifestation attributable to SSc, including Raynaud's phenomenon, the disease durations reported in our study may not be directly comparable to studies in which disease onset was established from the first non–Raynaud's phenomenon manifestation. In particular, this distinction would affect analysis of patients with limited cutaneous disease. Only index socioeconomic variables were recorded, but we recognize that these variables may change over time as pro-

gressive disability occurs with SSc. The fact that the manual data abstraction was performed by multiple abstractors could represent a source for bias but was mitigated by the use of a standardized form. Generalizability may be limited by the fact that patients saw a single provider at a single academic center; however, this circumstance likely resulted in more uniform treatment of patients, even though every evaluation and treatment was individualized and there were no rote protocols. A low number of non–African American, non-white patients also limits generalizability.

In our study, household income by zip code is an area-based socioeconomic measure that is used as a proxy for the individual characteristic of household income. In this approach, some authors have found a tendency of the aggregate variable to exaggerate the effect of the microlevel variable (26), particularly when the aggregate variable represents a broader construct than the microlevel variable (27). In a literal sense, our analysis of imputed income demonstrates that living with SSc in a less advantaged area correlates with an increased risk of mortality during follow-up. A lower-income zip code may also correlate with other relative deprivations that could influence the disease course. Such deprivations may include structural causes of increased difficulty in getting to follow-up (e.g., increased distance to the medical center, diminished access to transit), decreased access to primary care, increased pollution, etc. The Washington, DC area, in which most of the patients in this study lived, is notable for its overall affluence, prevalence of insurance coverage, and high educational attainment. And yet the District of Columbia itself is "noticeably segregated by ward" (28). There are 3 predominantly African American zip codes east of the Anacostia River (20032, 20020, and 20019) that have socioeconomic indicators (including preventable hospitalizations) which "are among the worst in the nation" (28). In our study, 17 African Americans (but no non–African Americans) hailed from these zip codes, and 6 (35%) died during follow-up.

In summary, we have demonstrated the unique clinical and serologic profile and increased morbidity and mortality of SSc in African Americans, relative to non–African Americans, in a large and previously undescribed cohort. In the US, race is largely a social construct, rather than a biologic one, and it is confounded by relative economic deprivation (24). Thus, in robustly controlling for socioeconomic status, we have demonstrated a relatively diminished magnitude and significance of the mortality effect conferred by race. Nonetheless, race and ethnicity do correlate with differences in fibrosis-related gene expression (29,30) and also with specific HLA haplotypes and single nucleotide polymorphisms. A growing body of recent research is exploring racial differences in the genetic basis of SSc (18,31,32). Higher socioeconomic status may blunt the effects of intrinsic racial differences. Regardless of the relative magnitudes of the contributory socioeconomic versus genetic factors, African Americans with SSc clearly merit more intensive efforts to facilitate timely diagnosis and access to continued evaluation and suppressive treatment, particularly with respect to cardiopulmonary involvement.

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. Moore 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. Moore, Steen.
Acquisition of data. Moore, Kramer, Eltaraboulsi, Steen.
Analysis and interpretation of data. Moore, Steen.

REFERENCES

1. Laing TJ, Gillespie BW, Toth MB, Mayes MD, Gallavan RH Jr, Burns CJ, et al. Racial differences in scleroderma among women in Michigan. Arthritis Rheum 1997;40:734–42.
2. Mayes MD, Lacey JV, Beebe-Dimmer J, Gillespie BW, Cooper B, Laing TJ, et al. Prevalence, incidence, survival, and disease characteristics of systemic sclerosis in a large US population. Arthritis Rheum 2003;48:2246–55.
3. Steen VD, Oddis CV, Conte CG, Janoski J, Casterline GZ, Medsger TA Jr. Incidence of systemic sclerosis in Allegheny County, Pennsylvania: a twenty-year study of hospital-diagnosed cases, 1963–1982. Arthritis Rheum 1997;40:441–5.
4. Nietert PJ, Mitchell HC, Bolster MB, Shaftman SR, Tilley BC, Silver RM. Racial variation in clinical and immunological manifestations of systemic sclerosis. J Rheumatol 2006;33:263–8.
5. Steen V, Domsic RT, Lucas M, Fertig N, Medsger TA Jr. A clinical and serologic comparison of African American and Caucasian patients with systemic sclerosis. Arthritis Rheum 2012;64:2986–94.
6. Gelber AC, Manno RL, Shah AA, Woods A, Le EN, Boin F, et al. Race and association with disease manifestations and mortality in scleroderma: a 20-year experience at the Johns Hopkins Scleroderma Center and review of the literature. Medicine (Baltimore) 2013;92:191–205.
7. Beall AD, Nietert PJ, Taylor MH, Mitchell HC, Shaftman SR, Silver RM, et al. Ethnic disparities among patients with pulmonary hypertension associated with systemic sclerosis. J Rheumatol 2007;34:1277–82.
8. Greidinger EL, Flaherty KT, White B, Rosen A, Wigley FM, Wise RA. African-American race and antibodies to topoisomerase I are associated with increased severity of scleroderma lung disease. Chest 1998;114:801–7.
9. Steen VD, Conte C, Owens GR, Medsger TA. Severe restrictive lung disease in systemic sclerosis. Arthritis Rheum 1994;37:1283–9.
10. Nietert PJ, Silverstein MD, Silver RM. Hospital admissions, length of stay, charges, and in-hospital death among patients with systemic sclerosis. J Rheumatol 2001;28:2031–7.
11. Kuwana M, Kaburaki J, Arnett FC, Howard RF, Medsger TA Jr, Wright TM. Influence of ethnic background on clinical and serologic features in patients with systemic sclerosis and anti-DNA topoisomerase I antibody. Arthritis Rheum 1999;42:465–74.
12. Nietert PJ, Silver RM, Mitchell HC, Shaftman SR, Tilley BC. Demographic and clinical factors associated with in-hospital death among patients with systemic sclerosis. J Rheumatol 2005;32:1888–92.
13. Mendoza F, Derk CT. Systemic sclerosis mortality in the United States: 1999-2002 implications for patient care. J Clin Rheumatol 2007;13:187–92.
14. Kaufman JS, Cooper RS, McGee DL. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race. Epidemiology 1997;8:621–8.
15. Reveille JD, Fischbach M, McNearney T, Friedman AW, Aguilar MB, Lisse J, et al. Systemic sclerosis in 3 US ethnic groups: a comparison

of clinical, sociodemographic, serologic, and immunogenetic determinants. Semin Arthritis Rheum 2001;30:332–46.

16. Nietert PJ, Silver RM. Patterns of hospital admissions and emergency room visits among patients with scleroderma in South Carolina, USA. J Rheumatol 2003;30:1238–43.
17. Population Studies Center, Institute for Social Research, University of Michigan. Zip code characteristics: mean and median household income. URL: https://www.psc.isr.umich.edu/dis/census/Features/tract2zip/.
18. Mayes MD. Race, scleroderma, and survival: why is there a difference? [editorial] J Rheumatol 2005;32:1873–4.
19. Ward MM. Socioeconomic status and the incidence of ESRD. Am J Kidney Dis 2008;51:563–72.
20. Feldman CH, Hiraki LT, Liu J, Fischer MA, Solomon DH, Alarcón GS, et al. Epidemiology and sociodemographics of systemic lupus erythematosus and lupus nephritis among US adults with Medicaid coverage, 2000–2004. Arthritis Rheum 2013;65:753–63.
21. 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.
22. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69:239–41.
23. Black HR, Quallich H, Gareleck CB. Racial differences in serum creatine kinase levels. Am J Med 1986;81:479–87.
24. Berry CE, Wise RA. Interpretation of pulmonary function test: issues and controversies. Clin Rev Allergy Immunol 2009;37:173–80.
25. Steen VD, Lucas M, Fertig N, Medsger TA. Pulmonary arterial hypertension and severe pulmonary fibrosis in systemic sclerosis patients with a nucleolar antibody. J Rheumatol 2007;34:2230–5.
26. Geronimus AT, Bound J, Neidert LJ. On the validity of using census geocode characteristics to proxy individual socioeconomic characteristics. J Am Stat Assoc 1996;91:529–37.
27. Geronimus AT. Invited commentary: using area-based socioeconomic measures. Think conceptually, act cautiously. Am J Epidemiol 2006;164:835–40.
28. King CJ. The health of the African American community in the District of Columbia: disparities and recommendations. URL: https://www.georgetown.edu/sites/www/files/The%20Health%20of%20the%20African%20American%20Community%20in%20the%20District%20of%20Columbia.pdf.
29. Silver RM, Bogatkevich G, Tourkina E, Nietert PJ, Hoffman S. Racial differences between blacks and whites with systemic sclerosis. Curr Opin Rheumatol 2012;24:642–8.
30. Reese C, Perry B, Heywood J, Bonner M, Visconti RP, Lee R, et al. Caveolin-1 deficiency may predispose African Americans to systemic sclerosis–related interstitial lung disease. Arthritis Care Res (Hoboken) 2014;66:1909–19.
31. Gorlova OY, Li Y, Gorlov I, Ying J, Chen WV, Assassi S, et al. Gene-level association analysis of systemic sclerosis: a comparison of African-Americans and white populations. PloS One 2018;13:e0189498.
32. Morgan ND, Shah AA, Mayes MD, Domsic RT, Medsger TA, Steen VD, et al. Clinical and serological features of systemic sclerosis in a multicenter African American cohort: analysis of the genome research in African American scleroderma patients clinical database. Medicine (Baltimore) 2017;96:e8980.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1164–1170
DOI 10.1002/acr.23763

BRIEF REPORT

Anti–RNPC-3 (U11/U12) Antibodies in Systemic Sclerosis in Patients With Moderate-to-Severe Gastrointestinal Dysmotility

Zsuzsanna H. McMahan,[1] Robyn T. Domsic,[2] Lei Zhu,[2] Thomas A. Medsger,[2] Livia Casciola-Rosen,[1] and Ami A. Shah[1]

Objective. To examine the association of anti–RNPC-3 antibodies in patients with systemic sclerosis (scleroderma or SSc) with selected gastrointestinal (GI) tract complications.

Methods. Sera from patients with SSc with or without severe GI dysfunction (total parenteral nutrition dependence) from the Johns Hopkins Scleroderma Center were screened for anti–RNPC-3 antibodies. We then examined anti–RNPC-3–positive cases and negative SSc controls from the University of Pittsburgh and the University of Pittsburgh Medical Center (UPMC) scleroderma cohort to confirm our findings and to examine whether specific GI features were associated with anti–RNPC-3 antibodies.

Results. In the discovery cohort, patients with SSc with severe GI dysfunction (n = 37) and without GI dysfunction (n = 38) were screened for anti–RNPC-3 antibodies. The former were more likely to have anti–RNPC-3 antibodies (14% versus 3%; $P = 0.11$). In the Pittsburgh cohort, moderate-to-severe GI dysfunction (Medsger GI score ≥2) was present in 36% of anti–RNPC-3–positive patients versus 15% of anti–RNPC-3–negative patients ($P \leq 0.01$). Anti–RNPC-3–positive patients were more likely to be male (31% versus 15%; $P = 0.04$), African American (18% versus 6%; $P = 0.02$), have esophageal dysmotility (93% versus 62%; $P < 0.01$), and interstitial lung disease (ILD) (77% versus 35%; $P < 0.01$). After adjusting for relevant covariates and potential confounders, moderate-to-severe GI disease was associated with anti–RNPC-3 antibodies (odds ratio [OR] 3.8 [95% confidence interval (95% CI) 1.0–14.3]), and ILD trended toward significance (OR 2.8 [95% CI 1.0–8.2]).

Conclusion. Patients with SSc and anti–RNPC-3 antibodies are more likely to be male and African American and to have moderate-to-severe GI disease and ILD. Further studies on larger patient cohorts may be helpful in further defining subsets of patients with SSc at risk for severe GI involvement.

INTRODUCTION

Gastrointestinal (GI) dysfunction is the most common internal complication of systemic sclerosis (SSc), affecting ≥90% of patients. The heterogeneity among patients with GI dysfunction is striking, because some patients have upper GI dysmotility, others have lower GI dysmotility, and still others have both (1).

Small nuclear ribonucleoproteins (RNPs) are recognized targets of the autoimmune response in SSc. While the protein portion of the complex is the most common target of the autoimmune response, distinct RNPs (e.g., U3 RNP, U1 RNP), are also well recognized. Recent reports (2,3) suggest that an association between anti–RNPC-3 (i.e., anti–U11/U12 RNP) antibodies and GI dysmotility in SSc may exist. However, one of these studies focused on a selected patient group (patients with SSc with cancer), limiting the generalizability of the findings (2,3). Furthermore, neither study assessed the association with distinct GI outcomes (2,3).

Supported by the Scleroderma Research Foundation. The Rheumatic Diseases Research Core Center is supported by the NIH (grants P30-AR-053503 and P30-AR-070254). Dr. McMahan's work was supported by a Jerome L. Greene Scholar award, the Rheumatology Research Foundation, a Johns Hopkins Clinician Scientist Career Development award, and the NIH (grant K23-AR-071473). Dr. Domsic's work was supported by the NIH (grant R21-AR-066305-02). Dr. Casciola-Rosen's work was supported by the Donald B. and Dorothy L. Stabler Foundation. Dr. Shah's work was supported by the NIH (grant K23-AR-061439) and the Donald B. and Dorothy L. Stabler Foundation.

[1]Zsuzsanna H. McMahan, MD, MHS, Livia Casciola-Rosen, PhD, Ami A. Shah, MD, MHS: Johns Hopkins University School of Medicine, Baltimore, Maryland; [2]Robyn T. Domsic, MD, MPH, Lei Zhu, MS, Thomas A. Medsger, MD: University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

Drs. Casciola-Rosen and Shah contributed equally to this work.

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

Address correspondence to Zsuzsanna McMahan, MD, MHS, 5200 Eastern Avenue, MFL Center Tower, Suite 5300, Baltimore, MD 21224. E-mail: Zmcmaha1@jhmi.edu.

Submitted for publication June 11, 2018; accepted in revised form September 18, 2018.

SIGNIFICANCE & INNOVATIONS

- Anti–RNPC-3 antibody–positive patients are significantly more likely to have moderate-to-severe gastrointestinal (GI) dysfunction, even after adjustment for relevant covariates.
- Esophageal dysmotility is more prevalent among anti–RNPC-3 antibody–positive patients with systemic sclerosis (SSc).
- Antibody status may inform GI risk stratification and associate with specific GI clinical complications in patients with SSc.
- Very high rates of interstitial lung disease in anti–RNPC-3 antibody–positive patients are further confirmed in this study.

In the current study, we sought to determine whether anti–RNPC-3 antibodies in SSc associate with severe GI dysmotility and with specific GI dysmotility complications. We initially compared patients on total parenteral nutrition (TPN) with asymptomatic patients from the Johns Hopkins Scleroderma Center and found that anti–RNPC-3 antibodies are more prevalent among the former group. We then sought to confirm and expand this finding by comparing GI severity and examining the prevalence of specific GI complications in anti–RNPC-3–positive and anti–RNPC-3–negative patients from the University of Pittsburgh and UPMC scleroderma cohort.

PATIENTS AND METHODS

Patients. The discovery cohort included all patients with SSc with severe GI dysfunction (requiring TPN) and patients with SSc without symptoms of GI dysfunction (modified Medsger severity score of 0) in the Johns Hopkins Scleroderma Center database (4). All patients meeting these GI criteria were included if they had both clinical data and banked serum, and met the 1980 American College of Rheumatology (ACR) criteria, the 2013 ACR/European League Against Rheumatism criteria, or at least 3 of 5 features of CREST syndrome (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasia) (5–7). Clinical charts of the cases and controls were reviewed to obtain details on SSc GI signs, symptoms, and severity, as well as to review all available objective GI tests. Because this study was specifically focused on GI dysmotility, patients with gastric antral vascular ectasia were excluded. All study patients were evaluated as part of routine clinical care at the Johns Hopkins Scleroderma Center.

As our initial analysis suggested an association between anti–RNPC-3 antibodies and severe SSc GI dysmotility, we subsequently performed a case–control study to confirm these findings using the University of Pittsburgh scleroderma cohort. All anti–RNPC-3–positive patients (cases) in the Pittsburgh database, first evaluated between 1980 and 2015, were identified and then matched to the next 3 consecutive anti–RNPC-3–negative patients with SSc (controls) evaluated in the clinic. The most extreme points in the Pittsburgh database were used to capture phenotype. GI severity (moderate to severe; Medsger severity score of ≥2) and the prevalence of specific GI characteristics were compared between groups. All cases and controls met the SSc classification criteria described above. Written informed consent was obtained from all patients at both sites. The institutional review boards of Johns Hopkins University and University of Pittsburgh approved this study.

Clinical phenotyping. *The Johns Hopkins Scleroderma Center (discovery cohort)*. The center's database captures demographic and detailed clinical data at the first encounter and every 6 months thereafter at follow-up visits. Disease duration was defined as the time from the first symptom (Raynaud's or non-Raynaud's) that was attributed to SSc by the treating physician, to the date of serum sample collection (sample tested for anti–RNPC-3 antibodies). Patients were classified as having diffuse or limited SSc, based on the extent of skin involvement. Cutaneous thickening proximal to the elbows and knees or involving the trunk at any time during the illness is considered diffuse SSc, and thickening that is always located only distal to the elbows and knees is considered limited SSc. Objective evidence of severe GI dysmotility was determined by physician documentation in the clinical notes and/or the presence of ≥1of the following: 1) esophageal dysmotility as determined by hypomotility or abnormal lower esophageal sphincter on esophageal manometry, esophageal dilation on esophagogastroduodenoscopy, dilation of the esophagus on fluoroesophagopharyngogram, esophageal transit delay on scintigraphy-based whole gut transit study, patulous esophagus on computed tomography (CT) chest scan, or esophageal hypomotility identified on barium swallow; 2) gastroparesis as determined by delayed gastric emptying on a scintigraphy-based gastric emptying study or whole gut transit study; 3) small bowel dysmotility as determined by distention, dilation, pseudo-obstruction, or pneumatosis intestinalis and/or air fluid levels on an abdominal radiograph or CT scan, dilation and/or markedly delayed small bowel transit on upper GI small bowel series, hydrogen breath test documenting small bowel bacterial overgrowth, small bowel follow-through confirming the presence of dilated intestinal loops with features of pseudo-obstruction, or hypomotility in the small bowel as determined by scintigraphy-based whole gut transit study; or 4) colonic dysmotility as defined by abnormal motility on a sitz marker study, abdominal radiograph or CT scan demonstrating dilated loops of colon, or a barium enema with dilated air-filled colon, or colonic hypomotility as determined by scintigraphy-based whole gut transit study.

Cardiac involvement was defined by a score of ≥1 on the Medsger severity scale (4,8). Skin involvement was scored with the maximum modified Rodnan skin score (range 0–51). Skeletal myopathy was considered present when patients had an abnormal

creatinine phosphokinase and muscle weakness and/or abnormal electromyography test, magnetic resonance imaging (MRI) result, or muscle biopsy result (8). Pulmonary function was determined based on findings on pulmonary function tests (minimum forced vital capacity [FVC] and minimum single breath diffusing capacity for carbon monoxide at any visit) (9,10). Estimated right ventricular systolic pressure was measured by echocardiogram and obtained as part of routine clinical screening for pulmonary hypertension; the maximum value at any visit was used for analysis. Sicca symptoms were defined as previously described (11). Renal crisis was confirmed by renal biopsy in the context of an acute symptomatic increase in blood pressure. All antibody data, outside of anti–RNPC-3 status, were obtained from the immunoblot assay, systemic sclerosis profile (EUROIMMUN), which was performed on the baseline serum sample.

University of Pittsburgh Scleroderma Center (confirmatory cohort). The Pittsburgh database contains demographic and clinical data, including SSc subtype, organ system symptoms, and objective testing at baseline. The approach to clinical phenotyping of the Pittsburgh patients was consistent with those described for the Johns Hopkins scleroderma cohort, with the exception of small differences in the GI, interstitial lung disease (ILD), and myopathy outcomes and calculation of disease duration. In this cohort, moderate-to-severe GI dysmotility was defined by distal esophageal aperistalsis, antibiotics for bacterial overgrowth, the presence of malabsorption syndrome, episodes of pseudo-obstruction, or the requirement for hyperalimentation (Medsger GI severity score ≥2). Given the heterogeneity of SSc GI findings, specific upper and lower GI outcomes were also recorded. The outcomes collected in the Pittsburgh database include patient-reported symptoms of acid reflux (heartburn), patient-reported symptoms of distal dysphagia for solid foods, esophageal dysmotility on imaging, hypomotility or the presence of small bowel dilatation on radiographic studies, the initiation of antibiotics for small intestinal bacterial overgrowth, or physician-judged malabsorption syndrome and/or the presence of dilated loops of bowel on a radiographic study. The presence of ILD was defined by radiographic evidence of bibasilar fibrosis and/or FVC <70% without obstructive findings. FVC was also analyzed as a longitudinal variable, using the lowest recorded percent predicted FVC in each patient. Muscle involvement was defined by the presence of proximal muscle weakness on physical examination (Medsger score ≥1), and creatine kinase levels >2 times the upper limit, and/or abnormal electromyography, MRI result, or muscle biopsy result consistent with myopathy. Disease duration was calculated from the date of any first symptom to the time of the first visit.

Anti–RNPC-3 antibody assay. In the discovery cohort, anti–RNPC-3 antibodies were assayed by immunoprecipitation of ^{35}S-methionine–labeled protein generated by in vitro transcription/translation (IVTT) from cDNA encoding full length RNPC-3 (Origene Technologies) (12). These antibodies were assayed in the Pittsburgh cohort using serum samples from the first visit as described previously (3). Because the assays used by the 2 centers were different, all Pittsburgh cases with an available banked serum sample (n = 41 of 49) and a sample of 15 randomly selected Pittsburgh controls were re-assayed by immunoprecipitation using RNPC-3 generated by IVTT at the Johns Hopkins site. Anti–RNPC-3 antibodies were confirmed in 39 of 41 of the Pittsburgh cases, and were not present in any of the 15 controls tested. Pittsburgh cases in which the anti–RNPC-3 antibody status was not confirmed at Johns Hopkins (n = 2) or when sera were unavailable for re-assay (n = 8), and their corresponding controls were excluded from the analysis.

Statistical analysis. In the first case–control analysis (Johns Hopkins), the outcome of interest was severe GI dysmotility (dependent variable) defined as severe GI dysmotility (e.g., requiring TPN; Medsger GI severity score of 4) versus no symptoms of GI dysmotility (modified Medsger GI severity score of 0). In the confirmatory case–control analysis, we examined the association between anti–RNPC-3–positive status (dependent variable) and the severity of SSc GI dysmotility (GI Medsger severity score ≥2), specific GI complications, and other non-GI clinical features. Pearson's correlation tests for parametric continuous variables and Spearman's correlation tests for nonparametric continuous variables were conducted. Evaluation for associations between dichotomous variables was done using chi-square or Fisher's exact tests. Comparisons of continuous variables were performed using Student's *t*-test (parametric data) and Wilcoxon's signed-rank test (nonparametric data) of matched samples. The association of GI severity and specific GI complications with anti–RNPC-3–positive status was evaluated using conditional logistic regression models consisting of the anti–RNPC-3 status indicator and potential covariates. We then constructed models to explore the association between moderate-to-severe GI disease and anti–RNPC-3 antibodies in SSc and included the following: unadjusted model with only the GI severity variable; a simple adjusted conditional logistic model adjusting for age, race, and GI severity; and an adjusted conditional logistic model using covariates included by backwards selection. For the backwards elimination, we compared the *P* value with a preselected significance level, 0.2. If the value was statistically nonsignificant, then the variable got dropped. Akaike's information criterion and likelihood ratio tests were used for selecting the best-fitted model. Statistical computing was conducted using Stata software, version 14.0, and SAS software, version 9.4. *P* values less than 0.05 were considered significant.

RESULTS

Discovery study evaluating the association between severe SSc GI dysmotility and anti–RNPC-3 antibodies in the Johns Hopkins cohort. SSc sera from 37 patients with severe GI dysmotility (requiring TPN) and 38 patients without symp-

toms and/or objective findings of GI dysmotility (modified Medsger GI severity score of 0) were assayed for antibodies to RNPC-3. All cases and none of the controls were confirmed to have severe GI dysmotility requiring TPN, as documented in the physician notes. In addition, 78% of cases (29 of 37) had objective testing reports available for review, also supporting the presence of GI dysmotility. The symptoms associated with TPN initiation were progressive weight loss, dysphagia, malabsorption, and/or recurrent pseudo-obstruction, which occurred in the context of severe GI dysmotility.

Table 1 shows the clinical features of these 2 groups. Anti–RNPC-3 antibodies were more frequently detected in patients with severe GI dysmotility compared to controls (14% versus 3%; $P = 0.11$). Patients in the severe GI group were significantly more likely to be male (38% versus 16%; $P = 0.031$), African American (43% versus 13%; $P \leq 0.01$), have diffuse disease (65% versus 34%; $P \leq 0.01$), myopathy (24% versus 5%; $P = 0.05$), and anti–U3 RNP antibodies (12% versus 0%; $P = 0.05$). Patients with severe GI symptoms were significantly

Table 1. Demographic disease and autoantibody characteristics of the 37 systemic sclerosis patients with severe gastrointestinal (GI) dysfunction (total parenteral nutrition dependence), and the 38 patients without symptoms of GI dysfunction in the discovery cohort (Johns Hopkins Scleroderma Center)*

Variable	Severe GI (n = 37)	No GI (n = 38)	*P*
Age, mean ± SD years (range 23–89)	53.7 ± 15.3	53.9 ± 16.1	0.95
Male	37.8†	15.8†	0.03†
African American	43.2†	13.2†	<0.01†
Ever smoker	46.0	52.6	0.56
Cutaneous subtype, diffuse‡	64.9†	34.2†	<0.01†
Disease duration, median (IQR) years (range 0.4–46)§	8.6 (4–20)	6.8 (3–16)	0.70
Modified Rodnan skin score, median (IQR) (range 0–47)‡	12.0 (4–30)	5.5 (3–22)	0.06
Raynaud's phenomenon >1‡	67.6	50.0	0.12
Lung involvement ≥1‡	86.5	68.4	0.10
Cardiac involvement ≥1‡	43.2	26.3	0.12
Skeletal myopathy, no. (%)‡	13 (35.1)†	5 (13.2)†	0.03†
Tendon friction rub, no. (%)‡	35 (20.0)	36 (16.7)	0.72
Renal crisis, no./total (%)‡	3/34 (8.8)	0 (0.0)	0.10
Sicca complex‡	59.5	39.5	0.08
Pulmonary function, mean ± SD			
Minimum FVC, % predicted	62.8 ± 21.4 (n = 35)	72.2 ± 23.1 (n = 35)	0.08
Minimum DLco, % predicted	60.1 ± 29.6 (n = 34)	67.7 ± 29.0 (n = 34)	0.29
RVSP, median (IQR) mm Hg	39.3 (35–44) (n = 34)	35.0 (28–50) (n = 27)	0.38
Autoantibodies, no./total (%)			
Anti–topoisomerase-1	4/34 (11.8)	7/36 (19.4)	0.52
Anticentromere	6/34 (17.7)	11/36 (30.6)	0.27
Anti–RNA polymerase 3	1/34 (2.9)	9/36 (25.0)	0.01†
Anti–Ro 52	10/34 (29.4)	12/36 (33.3)	0.80
Anti-Ku	1/34 (2.9)	2/36 (5.6)	1.00
Anti–Pm-Scl	0/34 (0.0)	2/36 (5.6)	0.49
Anti-ThTo	1/34 (2.9)	3/36 (8.3)	0.62
Anti–U3 RNP	4/34 (11.8)	0/36 (0.0)	0.05†
Anti–RNPC-3	5/37 (13.5)	1/38 (2.6)	0.11

* Values are the percentage, unless indicated otherwise. IQR = interquartile range; FVC = forced vital capacity; DLco = diffusing capacity for carbon monoxide; RVSP = right ventricular systolic pressure.
† Statistically significant.
‡ Maximum Medsger severity score ever recorded in the database; FVC and DLco are represented by the minimum values ever recorded.
§ Disease duration from any first symptom to the date of the serum sample collection.

less likely to have anti–RNA pol 3 antibodies (3% versus 25%; $P = 0.01$). Two patients in the severe GI group were double-positive for antibodies, having both anti–RNPC-3 antibodies, and antibodies to either Ro 52 or PM-Scl.

Confirmatory study defining specific GI characteristics associated with anti–RNPC-3–positive patients with SSc in the Pittsburgh cohort. Since the number of anti–RNPC-3 antibody–positive patients in the Johns Hopkins discovery study was small, but anti–RNPC-3 antibodies were >4 times more frequent than expected in the severe GI group, we pursued additional analyses to understand this association using the current Pittsburgh scleroderma cohort. This cohort is larger than the original published cohort that demonstrated anti–RNPC-3 (anti–U11/U12 RNP) antibodies as an important specificity in SSc (3), and included 39 anti–RNPC-3 antibody–positive cases and their 3:1 matched anti–RNPC-3–negative controls (n = 117) (see Patients and Methods).

Age, disease duration, and disease subtype were not significantly different between anti–RNPC-3 antibody–positive and –negative patients in the Pittsburgh cohort (Table 2). Anti–RNPC-3 antibody–positive patients were more likely to be African American (18% versus 6%; $P = 0.02$) and male (31% versus 15%; $P = 0.04$). Likewise, they were more likely to have moderate-to-severe GI dysfunction (36% versus 15%; $P \leq 0.01$). Twenty-four of the 31 patients (77%) with significant GI dysmotility (Medsger score ≥2) had confirmatory objective testing available in the database. Anti–RNPC-3 antibody–positive patients were also more likely to have ILD (77% versus 35%; $P < 0.01$), and the FVC was significantly lower in anti–RNPC-3–positive cases compared to controls (67% predicted versus 76% predicted; $P = 0.03$). The distribution of other clinical features did not differ between groups (Table 2).

We then examined the specific GI features associated with anti–RNPC-3 antibodies in the Pittsburgh cohort (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23763/abstract). Anti–RNPC-3–positive patients were significantly more likely to have esophageal dysmotility (92.6% versus 62.3%; $P < 0.01$), as defined by evidence of esophageal dysmotility on imaging or manometry, although the presence of distal dysphagia and heartburn were not significantly different between groups. There was no significant difference between the presence of other features of GI disease and the presence of anti–RNPC-3 antibodies. Although the number of patients studied was small, all comparisons favored higher frequencies of objective small intestinal involvement in anti–RNPC-3 antibody–positive patients.

We then sought to determine whether the association between clinical variables and anti–RNPC-3 antibodies from the bivariate analysis remained after adjusting for relevant covariates and potential confounders. In the unadjusted model (model 1), moderate-to-severe GI disease was associated with a 3.8-times increased odds of having anti–RNPC-3 antibodies (95% confidence interval [95% CI] 1.5–9.8). In the simple adjusted model (model 2), which was adjusted for age, race, and GI severity covariates, moderate-to-severe GI disease was again associated with a 3.8-times increased odds of having anti–RNPC-3 antibodies (95% CI 1.4–10.0). However, there was no significant association for age (odds ratio [OR] 1.0 [95% CI 0.95–1.0]) or African American race (OR 2.4 [95% CI 0.7–8.5]). In the fully adjusted model (model 3; covariates selected by backwards selection), patients with moderate-to-severe GI disease continued to have a 3.8-times increased risk of having anti–RNPC-3 antibodies (95% CI 1.0–14.3) (Table 3). There was no detectable change in the risk of having anti–RNPC-3 antibodies per year increase in age at first visit (OR 1.0 [95% CI 0.94–1.01]), or with African American race (OR 5.6 [95% CI 0.6–48.7]), diffuse cutaneous disease (OR 1.9 [95% CI 0.8–4.8]), or myopathy (OR 0.1 [95% CI 0.0–0.8]). An association with ILD trended toward significance (OR 2.8 [95% CI 1.0–8.2]).

Table 2. Clinical characteristics of anti–RNPC-3–positive and –negative patients in the confirmation study (University of Pittsburgh scleroderma cohort)*

Variable	Anti–RNPC-3 positive (n = 39)	Anti–RNPC-3 negative (n = 117)	*P*
Age, mean ± SD years	47.1 ± 13.4†	52.4 ± 14.8†	0.05†
Male	31†	15†	0.04†
African American	18†	6†	0.02†
Cutaneous subtype, diffuse‡	56	44	0.17
Disease duration, median (IQR) years§	3.1 (2.0–5.7)	4.4 (1.1–14.0)	0.91
Moderate-to-severe GI disease	36†	15†	<0.01†
Interstitial lung disease	77†	35†	<0.01†
Skeletal myopathy‡	5	14	0.15
Minimum FVC, % predicted, mean ± SD	66.8 ± 22.2 (n = 38)	76.3 ± 21.0 (n = 82)	0.03†

* Values are the percentage, unless indicated otherwise. IQR = interquartile range; GI = gastrointestinal; FVC = forced vital capacity.
† Statistically significant.
‡ Maximum ever Medsger severity score in the database.
§ From any first symptom to first visit.

DISCUSSION

This study evaluated the association between the severity of GI dysmotility and anti–RNPC-3 antibodies and whether specific GI complications were more frequent in connection

Table 3. Statistical models evaluating the association between anti–RNPC-3–positive patients with systemic sclerosis and the presence of moderate-to-severe gastrointestinal (GI) disease in the confirmatory study (University of Pittsburgh scleroderma cohort)*

Characteristic	Unadjusted conditional logistic: model 1	Conditional logistic adjusted for age and race: model 2	Conditional logistic adjusted for significant covariates: model 3†
Significant GI dysfunction			
Medsger GI score <2	1.00	1.00	1.00
Medsger GI score ≥2	3.84 (1.50–9.83)	3.79 (1.43–10.02)	3.81 (1.02–14.28)
Race			
Non–African American	1.00	1.00	1.00
African American	3.58 (1.11–11.48)	2.36 (0.65–8.54)	5.59 (0.64–48.7)
Diffuse cutaneous disease	1.61 (0.80–3.25)	–	1.90 (0.75–4.78)
Interstitial lung disease	5.85 (2.33–14.65)	–	2.79 (0.95–8.19)
Myopathy	0.36 (0.08–1.61)	–	0.07 (0.01–0.75)
Age at visit 1	0.97 (0.95–1.00)	0.97 (0.95–1.00)	0.97 (0.94–1.01)

* Values are the odds ratio (95% confidence interval).
† Adjusted conditional logistic regression model, adjusted for age and race, interstitial lung disease, diffuse cutaneous disease, and myopathy. Those covariates were selected by backwards selection.

with this autoantibody subset. In our initial discovery analysis, we screened for anti–RNPC-3 antibodies and compared their prevalence among patients with SSc with severe GI dysmotility (TPN dependence) and those patients without symptoms of GI dysmotility. We found that the frequency of anti–RNPC-3 antibodies is increased in the severe GI SSc population compared with patients with SSc without symptoms of GI dysmotility (14% versus 3%), consistent with findings in the published literature (3). We then further explored the association between anti–RNPC-3 antibodies and GI dysmotility using a second cohort (Pittsburgh) and demonstrated an association between these antibodies and the presence of moderate-to-severe GI disease and esophageal dysmotility, confirmed by objective testing. These data suggest that anti–RNPC-3 antibodies are a marker of clinically important GI dysmotility in SSc.

The association between anti–RNPC-3 antibodies and both pulmonary fibrosis and esophageal dysmotility in SSc is interesting. High rates of ILD are reported in association with anti–RNPC-3 antibodies in SSc, with anti–RNPC-3 antibody–positive patients having an estimated 70% prevalence of ILD (3). In addition, recent studies suggest that microaspiration in patients with SSc with uncontrolled reflux could contribute to the development of pulmonary fibrosis (13–15). Anti–RNPC-3 antibodies may identify a specific subset of patients at higher risk for microaspiration who would benefit from more aggressive gastroesophageal reflux disease management.

Our study confirms and extends observations made in 2 earlier reports. These earlier studies were limited by single-center assessments (3) and cancer bias in sample selection (2), and they did not examine GI complications as the primary outcome measure (2,3). The current study used the power of 2 large, carefully phenotyped SSc observational cohorts to examine the association between anti–RNPC-3 antibody–positive status, GI tract involvement severity, and specific GI complications. Narcotic use prior to the initiation of TPN was not widely available across the cohort and thus limited our analysis in this regard. Our study was limited by its retrospective design, because not all patients had complete GI assessments with objective testing (usually due to lack of symptoms warranting clinical testing). Prospective longitudinal data are needed to confirm our findings.

An association between anti–RNPC-3 antibodies, GI tract involvement severity, and specific GI dysmotility characteristics exists in patients with SSc. This association occurs alongside a very high rate of ILD. Further studies examining the use of anti–RNPC-3 antibodies as biomarkers for risk stratification of GI dysmotility in patients, and detailed studies of association of GI severity and ILD in patients with SSc should be performed.

ACKNOWLEDGMENTS

The authors thank Adrianne Woods and Margaret Sampedro for their involvement in the data set generation and sample acquisition from Johns Hopkins, and Maureen Laffoon for her work with the University of Pittsburgh Scleroderma Center database.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. McMahan 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. Domsic, Casciola-Rosen, Shah.
Acquisition of data. Medsger.
Analysis and interpretation of data. McMahan, Domsic, Zhu, Shah.

REFERENCES

1. Jaovisidha K, Csuka ME, Almagro UA, Soergel KH. Severe gastrointestinal involvement in systemic sclerosis: report of five cases and review of the literature. Semin Arthritis Rheum 2005;34:689–702.
2. Shah AA, Xu G, Rosen A, Hummers LK, Wigley FM, Elledge SJ, et al. Anti–RNPC-3 antibodies as a marker of cancer-associated scleroderma. Arthritis Rheumatol 2017;69:1306–22.
3. Fertig N, Domsic RT, Rodriguez-Reyna T, Kuwana M, Lucas M, Medsger TA Jr, et al. Anti–U11/U12 RNP antibodies in systemic sclerosis: a new serologic marker associated with pulmonary fibrosis. Arthritis Rheum 2009;61:958–65.
4. Medsger TA Jr, Bombardieri S, Czirjak L, Scorza R, Della Rossa A, Bencivelli W. Assessment of disease severity and prognosis. Clin Exp Rheumatol 2003;21 Suppl 29:S42–6.
5. Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum 1980;23:581–90.
6. 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.
7. LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. J Rheumatol 2001;28:157–36.
8. Medsger TA Jr, Silman AJ, Steen VD, Black CM, Akesson A, Bacon PA, et al. A disease severity scale for systemic sclerosis: development and testing. J Rheumatol 1999;26:2159–67.
9. Hankinson JL, Odencrantz JR. Fedan KB. Spirometric reference values from a sample of the general U.S. population. Am J Respir Crit Care Med 1999;159:179–87.
10. Knudson RJ, Kaltenborn WT, Knudson DE, Burrows B. The single-breath carbon monoxide diffusing capacity: reference equations derived from a healthy nonsmoking population and effects of hematocrit. Am Rev Respir Dis 1987;135:805–11.
11. McMahan ZH, Paik JJ, Wigley FM, Hummers LK. Determining the risk factors and clinical features associated with severe gastrointestinal dysmotility in systemic sclerosis. Arthritis Care Res (Hoboken) 2018;70:1385–92.
12. Fiorentino DF, Chung LS, Christopher-Stine L, Zaba L, Li S, Mammen AL, et al. Most patients with cancer-associated dermatomyositis have antibodies to nuclear matrix protein NXP-2 or transcription intermediary factor 1γ. Arthritis Rheum 2013;65:2954–62.
13. Zhang XJ, Bonner A, Hudson M, Canadian Scleroderma Research Group, Baron M, Pope J. Association of gastroesophageal factors and worsening of forced vital capacity in systemic sclerosis. J Rheumatol 2013;40:850–8.
14. Savarino E, Bazzica M, Zentilin P, Pohl D, Parodi A, Cittadini G, et al. Gastroesophageal reflux and pulmonary fibrosis in scleroderma: a study using pH-impedance monitoring. Am J Respir Crit Care Med 2009;179:408–13.
15. Christmann RB, Wells AU, Capelozzi VL, Silver RM. Gastroesophageal reflux incites interstitial lung disease in systemic sclerosis: clinical, radiologic, histopathologic, and treatment evidence. Semin Arthritis Rheum 2010;40:241–9.

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1171–1177
DOI 10.1002/acr.23761

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Characteristics of Usual Physical Therapy Post-Total Knee Replacement and Their Associations With Functional Outcomes

Carol A. Oatis,[1] Joshua K. Johnson,[2] Traci DeWan,[3] Kelly Donahue,[4] Wenjun Li,[5] and Patricia D. Franklin[5]

Objective. Although total knee replacement (TKR) surgery is highly prevalent and generally successful, functional outcomes post-TKR vary widely. Most patients receive some physical therapy (PT) following TKR, but PT practice is variable and associations between specific content and dose of PT interventions and functional outcomes are unknown. Research has identified exercise interventions associated with better outcomes but studies have not assessed whether such evidence has been translated into clinical practice. We characterized the content, dose, and progression of usual post-acute PT services following TKR, and examined associations of specific details of post-acute PT with patients' 6-month functional outcomes.

Methods. Post-acute PT data were collected from patients who were undergoing primary unilateral TKR and participating in a clinical trial of a phone-based coaching intervention. PT records from the terminal episode of care were reviewed and utilization and exercise content data were extracted. Descriptive statistics and linear regression models characterized PT treatment factors and identified associations with 6-month outcomes.

Results. We analyzed 112 records from 30 PT sites. Content and dose of specific exercises and incidence of progression varied widely. Open chain exercises were utilized more frequently than closed chain (median 21 [interquartile range (IQR) 4–49] versus median 13 [IQR 4–28.5]). Median (IQR) occurrence of progression of closed and open chain exercise was 0 (0–2) and 1 (0–3), respectively. Shorter timed stair climb was associated with greater total number of PT interventions and use and progression of closed chain exercises.

Discussion. Data suggest that evidence-based interventions are underutilized and dose may be insufficient to obtain optimal outcomes.

INTRODUCTION

Over 690,000 primary total knee replacement (TKR) surgeries were performed in the US in 2012 to relieve pain and restore physical function in patients with advanced knee arthritis, according to the Centers for Disease Control and Prevention. However, studies suggest that more than one-third of patients receiving TKR report little or no improvement in function (1). Following hospital discharge post-TKR, rehabilitation is a routine intervention, with 75–85% of patients receiving physical therapy (PT) (2). Guidelines consistently include post-acute rehabilitation (3–5), yet there is no accepted standard program of PT care and little is known about the contributions of rehabilitation to long-term outcomes.

Wide variation exists in the amount and form of PT following TKR (2,6–8). Little information exists about when or if PT should be provided, and which PT components are most beneficial. Evidence-based guidance is needed to decrease unwarranted treatment variation and optimize outcomes.

Consensus exists on the need to increase knee strength and range of motion (ROM) (3,5), but little agreement exists on the kinds and amount of exercises used. In a recent study, no consensus was found among patients, therapists, and surgeons on PT

ClinicalTrials.gov identifier: NCT00566826.
Supported in part by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant 1-RO1-AR-054479-01A1).
[1]Carol A. Oatis, PT, PhD: Arcadia University, Glenside, Pennsylvania; [2]Joshua K. Johnson, PT, DPT, PhD, Cleveland Clinic, Cleveland, Ohio [3]Traci DeWan, DPT: US Army, Fort Benning, Georgia; [4]Kelly Donahue, DPT: MossRehab, Einstein Healthcare Network, Elkins Park, Pennsylvania; [5]Wenjun Li, PhD, Patricia D. Franklin, MD, MBA, MPH: Northwestern University Feinberg School of Medicine, Chicago, Illinois.
No potential conflicts of interest relevant to this article were reported.
Address correspondence to Carol A. Oatis, PT, PhD, Department of Physical Therapy, Arcadia University, 450 South Easton Road, Glenside, PA 19038. E-mail: oatisc@arcadia.edu.
Submitted for publication March 6, 2018; accepted in revised form September 11, 2018.

SIGNIFICANCE & INNOVATIONS

- Although physical therapy (PT) is widely used following total knee replacement (TKR), treatment is extremely variable, lacking any standardized approach in practice. The contribution of "usual physical therapy" to functional outcomes is unknown.
- In the absence of evidence identifying optimal content and dose of PT, new reimbursement models may reduce the use of PT postoperatively without an understanding of the impact of such changes on functional outcomes. Pragmatic studies to identify optimal PT practice may have significant impact on both PT practice and public policy.
- To our knowledge, this is the first study to report the specific details of PT interventions provided in a clinical setting and preliminary analyses of associations with functional outcomes following TKR.
- Our data suggest that evidence-based PT interventions may be underutilized in clinical practice and that the dose of interventions may be insufficient to achieve optimal outcomes.

treatment duration, intensity, or frequency in rehabilitation post-TKR (3). Clinical trials demonstrate improved functional outcomes with quadriceps strengthening following TKR (9), but assessment of strengthening regimens suggest that many PT exercise interventions following TKR lack sufficient intensity to produce physiologic benefits (4,10). Our own data demonstrated that the strengthening exercises documented in PT records following TKR varied widely and approximately 25% of records had no documentation of progressive quadriceps strengthening (7).

In the absence of clear evidence for the contributions of post-acute PT following TKR, new reimbursement models may incentivize a reduction in PT services. Studies suggest that patients following total hip replacement may not benefit from PT (11). Anecdotal data suggest that a similar pattern is emerging post-TKR.

Given the variability in amount, content, and dose of PT following TKR, and trends toward reduction in PT services, the purpose of our study was to describe the content, dose, and progression of post-acute PT services across multiple PT facilities and to identify associations between specific details of the PT services and patients' 6-month self-reported functional outcomes and performance measures. We focused on the type and intensity of exercise content delivered and the number and timing of PT visits during the terminal episode of PT care following TKR surgery.

PATIENTS AND METHODS

This was a cross-sectional observational study of the usual course of post-acute PT provided to patients after hospital discharge for TKR. Data were obtained from PT records of individuals enrolled in the Joint Action Randomized Clinical Trial study at the University of Massachusetts Medical School (UMMS) between 2008 and 2011 (NCT00566826). The trial was designed to examine the effects on 6-month functional outcomes following TKR of a behavioral intervention that consisted of up to 12 telephone-delivered coaching sessions focused on at-home self-management strategies to enhance post-TKR recovery. Neither the patient's surgeon nor PT providers were aware of the patients' random allocation. The intervention did not influence PT care. The methods have been described elsewhere (12). The study was approved by the Institutional Review Boards of UMMS and Arcadia University.

Participants were consecutively enrolled from all individuals ages 21 years or older who were scheduled for primary TKR surgery at the Arthritis and Joint Replacement Center of the UMass Memorial Medical Center. Subjects were eligible if they had a primary unilateral TKR for osteoarthritis. Exclusion criteria included inflammatory arthritis, coexisting conditions preventing functional improvement, and cognitive impairments. More than 95% of eligible patients enrolled. Participants signed releases allowing review of their health records during the study period. Participants in the study received usual operative and rehabilitation care according to their own personal preferences and those of clinicians.

Baseline variables were collected prior to surgery, including age, sex, physical comorbid conditions, and body mass index (BMI). Mental and physical health status and function were collected at baseline and 6-months postsurgery using the mental component scores (MCS) and the physical component scores (PCS) of the Short Form 36 health survey and the joint specific Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).

Outcome measures at 6 months included the functional subscore of the WOMAC and knee performance measures consisting of the timed stair climb (TSC) and knee flexion ROM. The TSC was measured as the time in seconds required to ascend and descend a standard flight of stairs of 10 steps. Instructions to the patient were standardized and the patient was allowed to use the railing or assistive device as needed. ROM was measured in degrees using an inclinometer. With the patient lying supine, the hip on the side of the index knee was flexed to 90°. The patient actively flexed the knee as far as tolerated. The inclinometer was placed on the medial surface of the mid shaft of the tibia (Figure 1). Flexion ROM equaled the inclinometer reading plus 90°.

All performance measures were collected by 1 of 2 physical therapists, and interrater reliability of the performance measures was assessed using 7 patients post-TKR or with knee osteoarthritis for the TSC reliability testing, and 9 patients for the ROM test. ROM reliability was reassessed regularly throughout the study. Interrater reliability (2,1) for the TSC was 0.90. Interrater reliability of flexion ROM was 0.94 and remained >0.85 for the entire study.

Figure 1. Use of an inclinometer to measure knee flexion range of motion (ROM).

At least 3 dedicated arthroplasty surgeons at 1 high-volume center performed all TKR surgeries, using a consistent perioperative protocol for inpatient care. Patients chose their PT provider and received usual care PT as prescribed by their health care provider. At their 6-month study assessment, participants listed the facilities where they had received PT.

The trial enrolled 180 eligible patients. Figure 2 illustrates the disposition of the patients for this analysis. Sixteen patients were excluded due to no, insufficient, or invalid PT facility information. PT records for the remaining 164 subjects were requested from the facility where they had received their terminal episode of care (defined as care provided in the setting where the participant completed rehabilitation associated with the TKR). We focused on the terminal episode of care because studies demonstrate that quadriceps muscle strength decreases immediately following surgery and is less than preoperative levels at 4 weeks post-TKR (13). Rehabilitation directed toward functional improvement likely occurs in the final rehabilitation setting.

We performed a retrospective review of each record to determine the number of PT visits postsurgery and the type, frequency, and dose of each exercise provided over the entire episode of care. Each exercise was listed, and exercises were grouped into 3 categories, including open chain, closed chain, and passive. In open chain exercises, the patient actively moved the joint while the limb was non–weight-bearing. In closed chain exercises, the patient actively moved the joint while the limb was weight bearing. In passive interventions, the therapist moved the limb or joint. Intervention frequency was the number of times that an intervention was delivered over the course of care. Intervention dose was the number of times that an intervention was progressed. Progression was any increase in the level of difficulty of an exercise, by changing the form of the exercise or by increasing the resistance.

Investigators with advanced PT training extracted treatment data from the PT record. These extractors were trained to ensure that each exercise or progression was documented and classified in a consistent manner. Two investigators independently reviewed each record (JKJ, TD, KD). Differences were resolved through discussion and if necessary adjudicated by the lead PT investigator (CO).

Linear regression models assessed associations of the number of PT visits, PT intervention content, and frequency with 6-month outcomes, with and without adjusting for sex, age, baseline PCS, and baseline WOMAC function. Descriptive statistics were used to describe patient sociodemographic and clinical characteristics, and the utilization and characteristics of PT.

RESULTS

We received 159 of the 164 PT records (97%). Records for 112 patients (70% of records received) contained sufficient intervention detail to analyze (Figure 2). Records that were considered to be lacking sufficient detail were those that were missing initial or discharge evaluations or daily notes, referred to protocols not included in the record, or were illegible. Of the 112 records analyzed, 91 records were from 27 outpatient facilities and 21 records were from 3 home care facilities.

Baseline characteristics of the 112 patients whose records were reviewed were similar to the national average of patients undergoing primary TKR (14) (Table 1). Seventy percent of participants were women. The average age of patients was 64 years and average BMI was 32.8. Average PCS and MCS scores were 33.3 and 52.8, respectively (scores can range from 0–100, with higher scores indicating better health). Average WOMAC pain, stiffness, and function were each approximately 5.0. WOMAC scores can range from 0–10, with 10 being the worst (14).

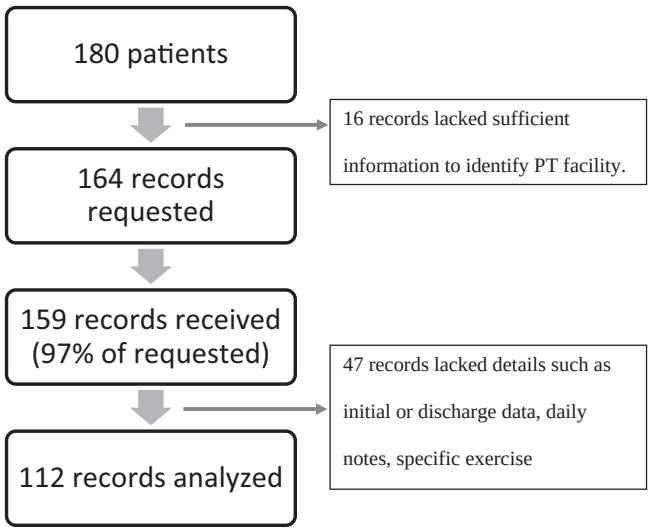

Figure 2. Physical therapy (PT) records requested, obtained, and analyzed for participants in clinical trial following total knee replacement.

Table 1. Baseline characteristics of patients whose records were reviewed and those of the entire sample*

Variable	No. of patients	Mean ± SD
Women	112	70.1
Age (years)	112	64.2 ± 8.4
BMI	112	32.8 ± 5.2
MCS score	112	52.8 ± 10.0
PCS score	112	33.3 ± 8.0
WOMAC†		
Pain	96‡	4.9 ± 2.0
Stiffness	92‡	5.3 ± 2.5
Function	90‡	4.9 ± 2.0

* BMI = body mass index; MCS = mental component scores; PCS = physical component scores; Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC).
† Scores ranged from 0–10 (10 being the worst).
‡ Some subjects lacked baseline WOMAC scores.

From the PT records, we identified a total of 34 different interventions, including 16 closed chain, 14 open chain, and 4 passive exercises. Four additional exercises (biking for ROM or endurance, and hamstring or plantar flexor stretches) were identified but lacked sufficient detail to categorize or quantify and were not included in the analysis.

Over the course of PT care, on average, patients had 14.5 PT visits and 12.8 different exercise interventions, including 4.5 closed chain exercises, 5.1 open chain exercises, and 1.8 passive interventions (Table 2). Open chain exercises were utilized more frequently than closed chain (median 21 [interquartile range (IQR) 4–49) versus median 13 [IQR 4–28.5]). Over the entire episode of care, on average, progression of open chain exercises was documented 2.6 times and 1.4 times for closed chain exercises. The median number of progressions documented for open chain exercises was 1 (IQR 0–3) and was 0 (IQR 0–2) for closed chain exercises. Timing of the initiation of either closed or open chain exercises varied widely.

Considerable variation existed in the specific exercises documented in the PT records. Of the 16 closed chain exercises identified, only squats (or wall slides) and step-ups were documented in >50% of the records (71% and 63%, respectively). Of the 14 open chain exercises identified, only straight leg raises, quadriceps sets, and short arc quadriceps exercises were documented in more than half the records (63%, 57%, and 55%, respectively). Of the 4 passive exercises, passive ROM for knee extension (57%) and knee flexion (59%) were documented in the records.

Shorter TSC was associated with greater total numbers of PT interventions, closed chain interventions, PT visits in which closed chain exercises were performed, and closed chain progressions (Table 3). Better post-TKR knee flexion ROM was seen among patients who had larger total numbers of closed chain progressions. Worse ROM was associated with total number of passive interventions and duration between surgery and the first postoperative day on which a passive intervention was provided.

Adjusting for baseline characteristics did not substantially alter the results, although some associations were no longer statistically significant at the 5% significance level. Adjustment for sex, age, baseline PCS, and WOMAC function did not alter the associations between the number of closed chain exercise progressions and TSC. We observed an approximately 1 second decrease in the TSC for every closed chain progression. In contrast, the total number of PT visits was not associated with knee performance outcomes or with WOMAC function scores.

Table 2. Dose of PT interventions by content*

Dosage	Mean ± SD	Median (IQR)
No. of PT visits	14.5 ± 8.4	12 (9.3–17.0)
No. of PT interventions	12.8 ± 5.3	13 (9–17)
Closed chain (CC)		
Total no. of CC interventions	4.5 ± 2.5	4 (3–7)
Frequency of CC exercises performed	20.7 ± 23.2	13 (4–28.5)
Total no. of CC progressions	1.4 ± 2.5	0 (0–2)
First day of CC intervention initiation	33.7 ± 23.7	36 (19–43.5)
Average no. of CC exercises per visit	1.5 ± 1.7	1.2 (0.4–2.3)
Open chain (OC)		
Total no. of OC interventions	5.1 ± 3.8	5 (2–8)
Frequency of OC exercises performed	30.45 ± 30.9	21 (4–49)
Total no. of OC progressions	2.6 ± 3.9	1 (0–3)
First day of OC intervention initiation	30.2 ± 15.0	32 (17.3–42)
Average no. of OC exercises per visit	2.5 ± 2.8	1.8 (0.3–4.0)
Passive interventions (PS)		
Total no. of PS interventions	1.8 ± 1.2	2 (1–3)
Frequency of PS performed	13.9 ± 20.2	7 (1–19)
Total no. of PS progressions	0.05 ± 0.3	0 (0–0)
First day of PS intervention initiation	28.12 ± 17.5	32 (16–40.8)
Average no. of PS per visit	0.85 ± 0.86	0.6 (0.1–1.5)

* PT = physical therapy; IQR = interquartile range.

Table 3. Associations of overall and content-specific dose of PT interventions with functional improvement at 6 months post TKR*

	Timed Stair Climb		WOMAC Function Subscale		Knee Flexion ROM	
Predictors	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
No. of PT visits	−0.11 (−0.38, 0.16)	−0.14 (−0.38, 0.09)	0.14 (−0.13, 0.42)	0.18 (−0.11, 0.46)	−0.28 (−0.64, 0.08)	−0.27 (−0.62, 0.08)
No. of PT interv.	−0.49 (−0.87, −0.10)†	−0.20 (−0.57, 0.17)	−0.05 (−0.49, 0.39)	−0.04 (−0.51, 0.44)	0.10 (−0.44, 0.63)	0.03 (−0.53, 0.59)
Closed chain (CC)						
Total CC interv.	−0.97 (−1.81, −0.14)†	−0.46 (−1.25, 0.32)	−0.36 (−1.29, 0.57)	−0.47 (−0.147, 0.53)	0.23 (−0.95, 1.40)	−0.06 (−1.25, 1.13)
Total visits completing CC	−0.09 (−0.18, −0.01)†	−0.05 (−0.13, 0.03)	−0.03 (−0.13, 0.07)	−0.02 (−0.13, 0.08)	0.06 (−0.05, 0.18)	0.03 (−0.10, 0.15)
Total CC progressions	−0.99 (−1.78, −0.20)†	−0.70 (−1.38, −0.02)†	−0.83 (−1.78, 0.12)	−0.96 (−1.97, 0.05)	1.49 (0.44, 2.54)‡	1.53 (0.48, 2.58)‡
1st day of CC interv. initiation	0.06 (−0.10, 0.22)	0.04 (−0.10, 0.18)	0.08 (−0.02, 0.19)	0.11 (−0.01, 0.22)	−0.14 (−0.34, 0.06)	−0.18 (−0.39, 0.04)
Ratio of CC visits to total visits	−1.45 (−3.08, 0.17)	−0.54 (−2.07, 0.99)	0.009 (−1.40, 1.42)	−0.03 (−1.54, 1.48)	1.99 (−0.17, 4.13)	1.57 (−0.65, 3.78)
Open Chain (OC)						
Total OC interv.	−0.30 (−0.87, 0.27)	−0.06 (−0.59, 0.47)	0.08 (−0.55, 0.70)	0.09 (−0.56, 0.74)	0.24 (−0.54, 1.03)	0.30 (−0.49, 1.09)
Total visits completing OC	−0.02 (−0.08, 0.05)	−0.01 (−0.07, 0.05)	0.04 (−0.03, 0.12)	0.05 (−0.03, 0.13)	0.03 (−0.06, 0.12)	0.04 (−0.05, 0.13)
Total OC progressions	−0.16 (−0.75, 0.42)	−0.40 (−0.89, 0.08)	0.04 (−0.60, 0.68)	0.02 (−0.64, 0.69)	0.39 (−0.36, 1.13)	0.46 (−0.30, 1.23)
1st day of OC interv. initiation	0.08 (−0.08, 0.24)	0.07 (−0.07, 0.22)	0.16 (−0.01, 0.34)	0.16 (−0.01, 0.34)	−0.06 (−0.27, 0.15)	−0.12 (−0.34, 0.11)
Ratio of OC visits to total visits	−0.09 (−1.08, 0.90)	0.03 (−0.84, 0.91)	0.41 (−0.41, 1.24)	0.43 (−0.42, 1.28)	0.67 (−0.66, 2.00)	0.89 (−0.39, 2.17)
Passive interv. (PS)						
Total PS interv.	−1.27 (−3.07, 0.53)	−0.97 (−2.61, 0.67)	−0.53 (−2.43, 1.38)	−0.32 (−2.31, 1.66)	−2.73 (−5.14, −0.33)†	−2.94 (−5.32, −0.56)†
Total visits completing PS	−0.05 (−0.16, 0.07)	−0.02 (−0.13, 0.079)	0.02 (−0.09, 0.14)	0.03 (−0.09, 1.50)	−0.02 (−0.18, 0.14)	−0.03 (−0.19, 0.12)
Total PS progressions	1.28 (−6.58, 9.13)	0.68 (−6.06, 7.43)	−4.01 (−14.3, 6.25)	−3.30 (−13.9, 7.31)	1.50 (−8.65, 11.65)	2.75 (−7.73, 13.2)
1st day of PS interv. initiation	−0.005 (−0.15, 0.14)	−0.015 (−0.14, 0.11)	0.07 (−0.07, 0.21)	0.07 (−0.08, 0.22)	−0.23 (−0.41, −0.06)‡	−0.24 (−0.42, −0.06)‡
Ratio of PS visits to total visits	−1.62 (−4.40, 1.17)	−0.41 (−2.90, 2.01)	−0.96 (−3.69, 1.77)	−0.99 (−3.85, 1.88)	0.29 (−3.51, 4.09)	−0.38 (−4.08, 3.32)

* Values are the regression coefficients (95% confidence interval (95% CI]). PT = physical therapy; TKR = total knee replacement; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index; ROM = range of motion; interv. = intervention.
† Significant at $P < 0.05$.
‡ Significant at $P < 0.01$.

DISCUSSION

Previous studies have found little difference in effectiveness between "usual" PT and other interventions (15). Our data show wide variation in the exercise treatment that patients receive during "usual" PT and suggest that the details of the intervention may be important determinants of the functional outcomes following TKR. To our knowledge, this is the first study to examine specific details of usual PT intervention and their effect on functional outcomes 6 month after TKR.

A previous study by Westby et al demonstrated little consensus among rehabilitation specialists regarding the timing, content, or quantity of PT interventions for patients post-TKR surgery although there is general agreement that exercise to increase ROM and quadriceps strength is important (3). There is considerable data supporting the effectiveness of progressive quadriceps strengthening to improve long-term function following TKR (9). Studies also support the use of weight-bearing exercises to enhance functional outcomes following TKR. Medicare Current Procedural Terminology data from a sample of Medicare TKR patients demonstrate that strength and ROM exercises and mobilization were the most commonly provided PT interventions post hospital discharge, but these data offer no details about exercise content or dose (6). The lack of such detail hampers the ability to assess the quality and effectiveness of routine PT practice.

The data presented in the current study confirm the wide variations in timing, content, and amount of PT in current practice and support the use of weight-bearing exercises and progressions to improve 6-month knee performance outcomes follow-

ing TKR. Indeed, our study demonstrated an approximately 1 second decrease in the time required to ascend and descend a flight of stairs for every closed chain progression made. In contrast, open chain exercises showed little or no association with performance outcomes.

Several plausible explanations exist for the associations of PT characteristics with the functional outcomes observed in this study. Perhaps only physically fit patients perform closed chain exercises, and they would naturally have the best outcomes. However, after adjustment for sex, age, baseline PCS, and WOMAC function, the associations between closed chain exercise progressions and TSC persisted. Total number of passive interventions and the days from surgery to onset of passive interventions were associated with worse ROM outcomes. These data may suggest that patients struggling with pain or ROM received more passive treatments such as passive stretching. The data may also suggest that, with limited treatment time, spending time on passive interventions decreases the amount of time available for active exercises that would increase function. The small sample size and cross-sectional nature of our study limits our ability to adequately adjust for all patient baseline characteristics.

Despite published evidence supporting progressive quadriceps strengthening and weight-bearing exercises for improved outcomes following TKR, the data from our study show that 106 of the records (95%) documented the use of weight-bearing exercises but only 35 of the records (31%) documented progression of those exercises. Muscle strengthening requires progressively increased resistance to adequately overload the muscle. Our findings suggest many patients may have exercised at a level insufficient to produce strengthening.

Our study has several limitations as well as strengths. First, because we collected data from a single geographic area and participants were participating in a clinical trial, the study results may not be generalizable. However, the participants of the clinical trial represented over 95% of all subjects undergoing TKR who met the inclusion criteria, and baseline characteristics of subjects whose PT records were reviewed did not differ from the total sample. We believe selection bias is unlikely. Second, our data were derived from retrospective chart reviews, and not all facilities require the same documentation. Only 70% of records contained sufficient detail for review. However, our review of over 100 records from 30 facilities provides considerable insight into what is meant by "usual" PT care. Third, although it is possible that the records reviewed did not include all exercises or progressions actually provided, we only reviewed those records that provided notes for every visit and detail about exercises and dose. While some details may be lacking, we believe that the records are generally representative of the care provided. Finally, the independent assessors of knee performance measures assured independent end points across all treating PTs.

Our data collection system accounted for every exercise intervention and created a classification system that provides a theoretical framework to evaluate the content of PT interventions. Virtually no record contained enough detail to completely characterize the volume and intensity of the exercise intervention, and many records lacked details to explain discharge decisions. Despite the limitations of the current study, it offers the first known insights into the actual details of usual PT provided to patients following TKR and the associations found between the details of PT practice and long-term functional outcomes.

Given the limitations of this study, the next step necessary to determine the value of PT in post-TKR recovery is to create a PT documentation system capable of prospectively detailing the content, frequency, and intensity of each PT visit and intervention. This documentation system is currently being deployed in a pragmatic study with a sufficient sample size of therapists and patients in order to identify the most effective treatment strategies within PT practice while accounting for individual patient characteristics.

In conclusion, the use of weight-bearing exercises and the frequency of progression of these exercises following TKR surgery are associated with positive functional outcomes at 6 months post-surgery. Evidence suggests that these interventions are underutilized in routine practice. Additional research is needed to fully understand the characteristics of PT interventions that contribute to positive functional outcomes and to identify "best practice" adjustments for patient characteristics that lead to optimal functional outcomes following TKR surgery.

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. Oatis 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. Oatis, Li, Franklin.

Acquisition of data. Oatis, Johnson, DeWan, Donahue, Li.

Analysis and interpretation of data. Oatis, Johnson, DeWan, Donahue, Li, Franklin.

REFERENCES

1. Franklin PD, Li W, Ayers DC. The Chitranjan Ranawat Award: functional outcome after total knee replacement varies with patient attributes. Clin Orthop Relat Res 2008;466:2597–604.
2. Freburger JK, Holmes GM, Ku LJ, Cutchin MP, Heatwole-Shank K, Edwards LJ. Disparities in post–acute rehabilitation care for joint replacement. Arthritis Care Res (Hoboken) 2011;63:1020–30.
3. Westby MD, Brittain A, Backman CL. Expert consensus on best practices for post–acute rehabilitation after total hip and knee arthroplasty: a Canada and United States Delphi study. Arthritis Care Res (Hoboken) 2014;66:411–23.
4. Westby M, Kazemi S, Jones D. Do exercise interventions for total knee arthroplasty have therapeutic validity? A sensitivity analysis of trials included in a Cochrane systematic review [abstract]. J Rheumatol 2014;41:1537–8.

5. Ebert JR, Munsie C, Joss B. Guidelines for the early restoration of active knee flexion after total knee arthroplasty: implications for rehabilitation and early intervention. Arch Phys Med Rehabil 2014;95:1135–40.
6. Ong KL, Lotke PA, Lau E, Manley MT, Kurtz SM. Prevalence and costs of rehabilitation and physical therapy after primary TJA. J Arthroplasty 2015;30:1121–6.
7. Oatis CA, Li W, DiRusso JM, Hoover MJ, Johnston KK, Butz MK, et al. Variations in delivery and exercise content of physical therapy rehabilitation following total knee replacement surgery: a cross-sectional observation study. Int J Phys Med Rehabil 2014;Suppl 5:002.
8. Fritz JM, Hunter SJ, Tracy DM, Brennan GP. Utilization and clinical outcomes of outpatient physical therapy for medicare beneficiaries with musculoskeletal conditions. Phys Ther 2011;91:330–45.
9. Bade MJ, Stevens-Lapsley JE. Early high-intensity rehabilitation following total knee arthroplasty improves outcomes. J Orthop Sports Phys Ther 2011;41:932–41.
10. Zech A, Hendrich S, Pfeifer K. Association between exercise therapy dose and functional improvements in the early postoperative phase after hip and knee arthroplasty: an observational study. PM R 2015;7:1064–72.
11. Austin MS, Urbani BT, Fleischman AN, Fernando ND, Purtill JJ, Hozack WJ, et al. Formal physical therapy after total hip arthroplasty is not required: a randomized controlled trial. J Bone Joint Surg Am 2017;99:648–55.
12. Rosal MC, Ayers D, Li W, Oatis C, Borg A, Zheng H, et al. A randomized clinical trial of a peri-operative behavioral intervention to improve physical activity adherence and functional outcomes following total knee replacement. BMC Musculoskelet Disord 2011;12:226.
13. Bade MJ, Kohrt WM, Stevens-Lapsley JE. Outcomes before and after total knee arthroplasty compared to healthy adults. J Orthop Sports Phys Ther 2010;40:559–67.
14. Li W, Ayers DC, Lewis CG, Bowen TR, Allison JJ, Franklin PD. Functional gain and pain relief after total joint replacement according to obesity status. J Bone Joint Surg Am 2017;99:1183–9.
15. Rajan R, Pack Y, Jackson H, Gillies C, Asirvatham R. No need for outpatient physiotherapy following total knee arthroplasty: a randomized trial of 120 patients. Acta Orthop Scand 2004;75:71–3.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1178–1185
DOI 10.1002/acr.23831

Factors Associated With Opioid Use in Presurgical Knee, Hip, and Spine Osteoarthritis Patients

J. Denise Power,[1] Anthony V. Perruccio,[2] Rajiv Gandhi,[2] Christian Veillette,[2] J. Roderick Davey,[2] Stephen J. Lewis,[2] Khalid Syed,[2] Nizar N. Mahomed,[2] and Y. Raja Rampersaud[2]

Objective. To evaluate rates of prescription opioid use among patients with presurgical knee, hip, and spine osteoarthritis (OA) and associations between use and sociodemographic and health status characteristics.

Methods. Participants were patients with presurgical, end-stage OA of the knee (n = 77), hip (n = 459), and spine (n = 168). Data were collected on current use of opioids and other pain medications, as well as measures of sociodemographic and health status variables and depression and pain (0–10 numeric rating scale). Rates of opioid use were calculated by sex, age, and surgical site. Multivariable logistic regression was used to examine associations between opioid use (sometimes/daily versus never) and other study variables.

Results. The mean age of participants was 65.6 years; 55.5% were women, 15% of patients reported "sometimes" using opioids, and 15% reported "daily use." Use of opioids was highest among patients with spine OA (40%) and similar among patients with knee and hip OA (28% and 30%, respectively). Younger women (ages <65 years) reported the greatest use of opioids overall, particularly among patients with spine OA. From multivariable logistic regression, greater likelihood of opioid use was significantly associated with spine OA (versus knee OA), obesity, being a current or former smoker, higher symptomatic joint count, greater depressive symptoms, greater pain, and current use of other prescription pain medication.

Conclusion. Nearly one-third of patients with presurgical OA used prescription opioid medication. Given the questionable efficacy of opioids in OA and risk of adverse effects, higher opioid use among younger individuals and those with depressive symptoms is of concern and warrants further investigation.

INTRODUCTION

Osteoarthritis (OA) is a chronic, degenerative condition characterized primarily by pain. OA ranks among the top 10 causes of disability worldwide and is associated with substantial social and economic costs (1). Due to the aging of the population, the prevalence and impact of this disease is projected to greatly increase (2,3).

OA treatment focuses on symptom management, with end-stage disease leading to surgical total joint replacement (TJR) for hip and knee OA and decompression (with or without fusion procedures) for lumbar spine OA. Patients typically live with OA for many years before surgery (4) making presurgical pain management both a critical and often long-term clinical undertaking. Unfortunately, effect sizes for typical first-line analgesics for OA, such as acetaminophen and nonsteroidal antiinflammatory drugs (NSAIDs), are small overall and decrease with long-term use (5,6). The appropriate role of opioids in OA pain management is not clear and treatment guidelines have varied widely. The most recent American College of Rheumatology guidelines (7) for the management of hip and knee OA "conditionally recommend" the use of opioids in patients who had an inadequate response to initial therapy. Current Osteoarthritis Research Society International guidelines (8) for knee OA list opioids as a treatment of "uncertain appropriateness," and guidelines from the American Academy of Orthopedic Surgeons (9) state that they are "unable to recommend for or against the use of opioids" in knee OA. Specific guidelines for pharmacologic management of lumbar spine OA, or spinal stenosis, are lacking due to sparse evidence (10).

Supported by the Toronto General and Western Hospital Foundation.

[1]J. Denise Power, PhD: Arthritis Program and Krembil Research Institute, University Health Network, Toronto, Ontario, Canada; [2]Anthony V. Perruccio, PhD, Rajiv Gandhi, MD, MSc, Christian Veillette, MD, MSc, J. Roderick Davey, MD, Stephen J. Lewis, MD, MSc, Khalid Syed, MD, MEd, Nizar N. Mahomed, MD, ScD, Y. Raja Rampersaud, MD: Arthritis Program and Krembil Research Institute, University Health Network and University of Toronto, Toronto, Ontario, Canada.

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

Address correpondence to Y. Raja Rampersaud, MD, Toronto Western Hospital, University Health Network, 399 Bathurst Street, East Wing, 1st Floor, Room 1-441, Toronto, Ontario, Canada M5T 2S8. E-mail: raja.rampersaud@uhn.ca.

Submitted for publication July 18, 2018; accepted in revised form January 8, 2019.

SIGNIFICANCE & INNOVATIONS

- Opioid use and abuse is a critically important public health issue in North America. From a societal perspective, use of opioids in osteoarthritis (OA) has the potential to have a significant impact on the numbers of opioids in circulation, given the substantial and growing population prevalence of OA.
- While mounting evidence suggests that presurgical opioid use may negatively impact surgical outcomes, nearly one-third of patients with presurgical hip, knee and spine OA reported using prescription opioid medications for their OA pain, with one-half of them reporting daily usage.
- Higher rates of opioid use were found among younger women and lumbar spine OA patients, and use was associated with greater depressive symptoms.
- There is currently a lack of consensus from professional medical organizations regarding the appropriate use of opioid medications in OA. Our findings raise questions as to the appropriateness of current patterns of use among patients with hip, knee, and lumbar spine OA and highlight the need for more specific guidance.

With the emergence of opioid abuse as a critically important public health issue in North America, greater clarity regarding the appropriate use of opioids in OA is of heightened importance. Rates of use of prescription opioid analgesics in North America are more than double those of the European Union and Australia/New Zealand (11) and have increased in parallel with opioid-related morbidity and mortality (12). Not surprisingly, the use of prescription opioids to manage chronically painful conditions like OA has been under greater scrutiny due to concerns about the potential for misuse, dependency, and addiction. From a societal perspective, widespread use of opioids in OA has the potential to have a significant impact on the numbers of opioids in circulation, given the substantial and growing population prevalence of OA (2,3). From an OA surgical treatment perspective, it is also of concern given mounting evidence that suggests that presurgical opioid use may negatively impact surgical outcomes leading to less improvement in pain and function, greater postsurgical opioid use, and higher rates of adverse events (13–16).

The current study represents the first stage of a larger project aimed at examining the use and impact of opioid medications in surgical OA patients. The specific objectives were to examine the rates of prescription opioid use among patients with presurgical knee, hip, and spine OA and investigate associations between opioid use and a range of sociodemographic and health status variables. Given the questionable efficacy and potential individual and societal impacts of opioid use in OA, it is important to understand current usage patterns in order to assess their appropriateness and potential impacts. Elucidation of the factors associated with opioid use in patients with OA may also prove informative in terms of potentially targeting efforts aimed at modifying patterns of opioid use.

PATIENTS AND METHODS

This cross-sectional study is a retrospective analysis of baseline data from an ongoing prospective study (Longitudinal Evaluation in the Arthritis Program [LEAP-OA]). Patients with OA scheduled for orthopedic surgery are consecutively recruited from the Toronto Western Hospital in Toronto, Ontario, Canada. Eligibility criteria include age ≥18 years of age and the ability to read and comprehend English. Individuals undergoing revision procedures and those with posttraumatic or inflammatory forms of arthritis are excluded. For the present analysis, 1,126 patients recruited between November 2013 and January 2017 were included, and 86.8% of eligible patients agreed to participate. Seventy-eight patients were excluded due to missing opioid data. Participants included 539 patients with knee OA and 436 with hip OA who were scheduled for primary, unilateral TJR and 151 patients with lumbar spine OA scheduled for decompression surgery with or without fusion. The study was approved by the University Health Network Research Ethics Board. Written informed consent was obtained from all patients.

Participants completed a questionnaire within the 3 weeks prior to surgery. Patients were asked about their current use (response options of "never," "sometimes" or "daily") of medications for their arthritis/joint pain. Our variable of interest was derived from the response to the item inquiring about use of "Narcotic/Opioid Pain Medications e.g., Dermerol, MS Contin, Morphine, Oxycontin, Percocet, Talwin, Tylenol #3." Two additional dichotomous (never versus sometimes/daily) pain medication variables were created; 1 for the use of other prescription arthritis medications (NSAIDs, antidepressants, neuroleptics/anticonvulsants), and 1 for the use of over-the-counter medications.

Data were also collected on sociodemographic characteristics including sex, age, and highest level of education. Data on measured height and weight were used to compute body mass index (BMI), categorized as underweight (<18.5 kg/m^2), normal (18.5 to <25 kg/m^2), overweight (25 to <30 kg/m^2), or obese (≥30 kg/m^2) (17). Smoking status was grouped as never smoker, former smoker, or current smoker. A comorbidity count variable was derived from yes/no responses to an extended list of 20 conditions based on the American Academy of Orthopedic Surgeon's comorbidity scale (18). Fibromyalgia was considered separately as a single dichotomous variable (present versus absent). Participants indicated (on a homunculus) any joints/sites that were painful on most days for at least a month. A summed count score of painful joints was derived, excluding the surgical joint.

Depressive symptoms were measured using the 7-item depression subscale of the Hospital Anxiety and Depression Scale (19). This measure has been found to be a reliable and valid measure for assessing emotional distress in medical populations (20).

Data on pain intensity were derived from a 0–10 numeric rating scale for average hip/knee/spine pain in the last week, with responses of 0 representing "no pain" and 10 representing "worst pain imaginable." Numeric rating scales for pain have been found to be reliable and valid measures of pain intensity in OA and a variety of other chronic pain populations (21).

Descriptive statistics were generated for all variables, including mean and standard deviations for continuous variables and frequencies and percentages for categorical variables. These were generated overall and separately for daily/sometimes and never opioid use. Differences between these latter groups were assessed using *t*-tests and chi-square tests, as appropriate. Rates (percentages and 95% confidence intervals [95% CIs]) of opioid use were also calculated by surgical site. Rates of any opioid use (sometimes or daily use) were further examined by sex and age (<65 years and ≥65 years). Exact 95% CIs were computed for all percentages. Multivariable logistic regression was used to examine the associations between current opioid use (daily/sometimes versus never) and all of the above noted study variables, including surgical site. All analyses were conducted using SAS, version 9.4. *P* values ≤0.05 (2-tailed) were considered significant.

Table 1. Sample characteristics*

Variable	Overall sample (n = 1,126)	Opioid users (n = 339)	Opioid nonusers (n = 787)	*P*†
Surgical site				
Knee	539 (47.9)	149 (44.0)	390 (49.6)	0.016
Hip	436 (38.7)	130 (38.4)	306 (38.9)	
Spine	151 (13.4)	60 (17.7)	91 (11.6)	
Women	617 (54.8)	210 (62.0)	407 (51.7)	0.002
Men	509 (45.2)	129 (38.1)	380 (48.3)	
Age				
Mean ± SD	65.5 ± 9.2	63.9 ± 9.4	66.2 ± 9.1	<0.001
<65 years	503 (44.7)	175 (51.6)	328 (41.7)	0.002
≥65 years	623 (55.3)	164 (48.4)	459 (58.3)	
Education				
High school	309 (28.6)	114 (34.2)	195 (26.0)	0.006
>High school	773 (71.4)	219 (65.8)	554 (74.0)	
BMI				
Underweight/normal	233 (22.9)	52 (16.8)	181 (25.6)	<0.001
Overweight	369 (36.3)	106 (34.2)	263 (37.3)	
Obese	414 (40.8)	152 (49.0)	262 (37.1)	
Smoking status				
Never	549 (49.6)	144 (43.2)	405 (52.3)	<0.001
Former	455 (41.1)	136 (40.8)	319 (41.2)	
Current	103 (9.3)	53 (15.9)	50 (6.5)	
Fibromyalgia‡	149 (13.2)	74 (21.8)	75 (9.5)	<0.001
Comorbidity count, mean ± SD	1.7 ± 1.6	2.1 ± 1.6	1.6 ± 1.6	<0.001
SJC, mean ± SD	5.6 ± 9.5	8.0 ± 11.8	4.6 ± 8.1	<0.001
Depressive symptoms (0–21 scale), mean ± SD	5.3 ± 3.6	6.5 ± 4.0	4.8 ± 3.4	<0.001
Pain intensity (0–10 scale), mean ± SD	6.1 ± 2.2	6.9 ± 1.8	5.7 ± 2.2	<0.001
Prescription pain medication§	612 (54.5)	245 (72.7)	367 (46.6)	<0.001
OTC pain medication use	907 (81.9)	262 (80.1)	645 (82.6)	0.332

* Values are the frequency (%) unless indicated otherwise. Comparison of patients who used opioids (daily or sometimes) versus those who did not use opioids. BMI = body mass index; SJC = swollen joint count.
† All values statistically significant (except over-the-counter [OTC] pain medication use).
‡ Percentage of patients who report having fibromyalgia.
§ Percentage of patients reporting any current use (excluding opioids).

Table 2. Presurgical opioid use by surgical site and frequency of use*

	Sometimes use	Daily use	Sometimes or daily
Knee OA	15.8 (12.8–19.1)	11.9 (9.3–14.9)	27.6 (23.9–31.6)
Hip OA	12.8 (9.9–16.4)	17.0 (13.6–20.8)	29.8 (25.6–34.4)
Spine OA	19.2 (13.3–26.4)	20.5 (14.4–27.9)	39.7 (31.9–48.0)
Overall sample	15.1 (13.1–17.3)	15.0 (13.0–17.2)	30.1 (27.4–32.9)

* Percentage (95% confidence interval [95% CI]) of patients reporting indicated frequency of opioid use. OA = osteoarthritis.

RESULTS

Table 1 describes the study sample and compares the 339 participants (30%) who reported currently taking opioids (daily or sometimes) with the 787 participants (70%) who reported no current use. About one-half of the sample was scheduled for knee TJR (48%), 39% for hip TJR, and 13% for surgery for lumbar spine OA (Table 1). The mean age of the sample was 65.5 years, 55% of participants were women and most were overweight (36%) or obese (41%). Participants reported a mean of 5.6 symptomatic joints and 1.7 comorbid conditions, with 13% of participants reporting fibromyalgia specifically. The mean pain intensity score for the sample was 6.1 out of a possible score of 10 and the mean depression score was 5.3 out of a possible 21.

Those who used opioids comprised a significantly higher proportion of spine patients, women, and younger individuals (ages <65 years) (Table 1). Higher proportions of individuals taking opioids also had less than a high school education, were obese, and were current smokers than those who were not taking opioids. In addition, individuals taking opioids reported significantly more symptomatic joints and comorbidities, including fibromyalgia. Those taking opioids were more likely to report taking other prescription pain medications and had significantly higher mean depression and pain scores.

Overall, an equal proportion of study participants reported sometimes (15%) and daily (15%) opioid use (Table 2). Patients with hip and knee OA used opioids sometimes or daily in similar proportions. Patients with lumbar spine OA used opioids in higher proportions (39.7% [95% CI 31.9–48.0%]) than those with knee OA (27.6% [95% CI 23.9–31.6%]).

Approximately one-third of women undergoing surgery for OA reported using opioids sometimes or daily for their arthritis/joint pain (34.0% [95% CI 30.3–37.9%]), compared to about one-fourth of men (25.3% [95% CI 21.6–29.4%]). Overall, younger men (ages <65 years) and older men (ages ≥65 years) reported taking opioids in similar proportions (Table 3). Younger women, however, reported higher use than older women, and used opioids in the highest proportion across the age and sex strata at 41.4% (95% CI 35.7–47.3%). Younger women with spine OA specifically had the highest point estimate for opioid use, at 61.3%, although the associated 95% CI (42.2–78.2%) overlapped with those for the other age, sex, and surgical site strata.

Results from the multivariable logistic regression model (current daily/sometimes opioid use versus no use) are shown in Table 4. Of note, relative to the descriptive findings, the effect of sex was not statistically significant and that of age was marginally nonsignificant in the fully adjusted model, with an odds ratio (OR) of 0.72 (95% CI 0.52–1.02) for those ages ≥65 years versus <65 years. Individuals with lumbar spine OA had 1.73 times greater odds than those with knee OA for opioid use. Patients who were obese had significantly greater odds than those who were of normal weight or underweight. There was a strong effect of smoking status; current and former smokers had odds 2.76 and 1.58 times greater than never smokers for use of opioids, respectively. Individuals reporting more symptomatic joints also had significantly higher odds of opioid use, such that the odds of use were 1.03 times greater for each additional joint. Patients with OA who reported using other prescription pain medications were also more likely to report current opioid use for their arthritis/joint pain. Greater depressive symptoms and higher pain intensity

Table 3. Presurgical daily or sometimes opioid use by surgical site, sex, and age group*

	Men		Women	
	Ages <65 years	Ages ≥65 years	Ages <65 years	Ages ≥65 years
Knee OA	27.7 (18.9–37.9)	18.1 (12.1–25.6)	40.5 (32.6–48.9)	23.9 (17.5–31.3)
Hip OA	22.6 (14.6–32.4)	31.1 (22.3–40.9)	36.9 (28.0–46.6)	27.9 (20.4–36.5)
Spine OA	30.8 (14.3–51.8)	30.9 (19.1–44.8)	61.3 (42.2–78.2)	41.0 (25.6–57.9)
Overall sample	25.8 (20.1–32.3)	25.0 (20.2–30.3)	41.4 (35.7–47.3)	27.5 (22.8–32.7)

* Percentage (95% CI) of patients reporting any opioid use (sometimes or daily). OA = osteoarthritis.

Table 4. Multivariable logistic regression results: associations with presurgical opioid use (current daily/sometimes use versus no use)*

Variable	OR (95% CI)	*P*
Surgical site (vs. knee)†		
Hip	1.33 (0.93–1.90)	0.115
Spine	1.73 (1.05–2.85)‡	0.030‡
Sex (female vs. male)	1.28 (0.91–1.81)	0.152
Age (≥65 years vs. <65 years)	0.72 (0.52–1.02)	0.061
Education (≤high school vs. >high school)	1.26 (0.88–1.81)	0.200
BMI (vs. underweight/normal)		
Overweight	1.43 (0.90–2.29)	0.133
Obese	1.83 (1.16–2.91)‡	0.010‡
Smoking status (vs. never smoker)		
Current smoker	2.76 (1.61–4.73)‡	<0.001‡
Former smoker	1.58 (1.12–2.23)‡	0.009‡
Fibromyalgia (present vs. absent)	1.56 (0.98–2.48)	0.060
Comorbidity count	1.06 (0.95–1.17)	0.315
Symptomatic joint count	1.03 (1.01–1.04)‡	0.007‡
Depressive symptoms§	1.05 (1.01–1.10)‡	0.030‡
Pain intensity¶	1.25 (1.14–1.37)‡	<0.001‡
Prescription pain medication (any use vs. no use)	2.32 (1.65–3.25)‡	<0.001‡
OTC pain medication (any use vs. no use)	0.67 (0.44–1.02)	0.061

* Values are the odds ratios (OR) (95% confidence intervals [95% CIs]) unless indicated otherwise. BMI = body mass index; OTC = over-the-counter.
† The estimate for spine versus hip is OR 1.30 (95% CI 0.79–2.14); *P* = 0.296.
‡ Values are statistically significant.
§ Scores are derived from the Hospital Anxiety and Depression Scale and range from 0 to 21, with higher scores indicating greater depressive symptoms. ORs are for a 1-unit increase in score.
¶ Ratings range from 0 to 10 for average joint-specific pain in the past week, with higher scores indicating greater pain intensity. ORs are for a 1-unit increase in rating.

were also significantly associated with an increased likelihood of opioid use.

DISCUSSION

Our study adds to a growing body of literature examining opioid use in chronic pain populations and is one of few that focus specifically on examining factors associated with use in patients with OA. Although guidelines for opioid use in OA are unclear, we found that nearly one-third of patients with presurgical hip, knee, and spine OA reported using prescription opioids for their OA pain, with one-half reporting daily usage. Higher rates of opioid use were found among patients with lumbar spine OA and younger women (ages <65 years), and use was associated with greater depressive symptoms. These findings may prove useful in targeting both future research to better understand opioid use in OA as well as interventions to potentially reduce use.

There have only been a limited number of studies that have shown rates of opioid use before OA-related surgical procedures specifically. Our rate for knee replacement patients of 27.6% (95% CI 23.9%–31.6%) is similar to that reported in the US study by Franklin et al (22). The authors reported that 24% of knee replacement patients had at least 1 opioid prescribed prior to the procedure. In contrast, Hansen et al (23) reported in an Australian administrative data study that 38.6% of patients with presurgical knee replacement had at least 1 opioid prescription filled in the year prior to surgery. The longer study timeframe and the focus on dispensed prescriptions, rather than actual opioid use, may have contributed to this comparatively higher estimate. Opioid use prior to spine surgery has been reported to vary between 20% and 55% (24,25), with some differences between studies in terms of the specific diagnoses and procedures that were considered. Walid et al (25) reported comparable data to the findings of the current study on patients with lumbar decompression and fusion surgical, and found that 46.5% were taking opioids presurgery, a slightly higher proportion than the 40% we found in our sample. Select baseline data from the Spine Patient Outcomes Research Trial (SPORT) (26) has also been reported separately for lumbar spinal stenosis surgical candidates with and without degenerative spondylolisthesis, and findings were also similar to those of our study in that 34% of participants reported current opioid usage.

Published findings on the associations of age and sex with opioid use in OA have varied. In an Ontario, Canada population-based cohort study of older adults with hip and knee OA, age and sex were not significantly associated with opioid use (27). In contrast, Wright et al (28) reported that opioid use correlated with female sex and younger age in those with knee OA specifically. A study by DeMik et al (29), in which OA of any joint was taken into consideration, corroborated the findings of Wright et al in terms of the effect of younger age, but found greater opioid use in men than women. Given these inconsistencies in reported findings, we specifically examined rates of opioid use by age and sex, and found that women had higher rates of use than men and it was among women particularly that the higher rates of use in younger individuals were most apparent. However, the effects of age and sex were not statistically significant in the adjusted analyses (*P* = 0.06 for age, *P* = 0.11 for sex), which suggests that the additionally considered factors may explain the observed variation in rates. It is likely that age and sex differences in analgesic use are influenced by multiple factors. This is an area of research in OA that warrants further exploration, particularly as more is known about the potential risks and negative impacts associated with opioid use.

From the multivariable regression analyses performed in the present study, we demonstrated that higher levels of pain intensity

were significantly associated with increased odds of opioid use. Although this finding may simply reflect the prescription of opioids as a last resort for patients who continue to report high pain, it could also reflect that these medications may not adequately control OA pain. In addition to the risk of adverse effects such as dependency, there is evidence to suggest that even strong opioids are not more effective than acetaminophen or NSAIDs in reducing pain due to musculoskeletal conditions (30–32). A 2016 systematic review focusing on knee OA estimated that the mean decrements in Western Ontario and McMaster Universities Osteoarthritis Index pain achieved by NSAIDs (–18 points), less potent opioids such as tramadol (–18 points), and more potent opioids such as oxycodone (–19 points) were highly comparable (31).

The association between opioid use and greater pain could also be influenced by issues with tolerance or opioid-induced hyperalgesia (33). Wasserman et al (34) reported that individuals who consistently report high pain despite taking opioids are also more likely to report other central symptoms, including greater neuropathic and fibromyalgia-like symptoms. However, in our multivariable analyses the effect of comorbid fibromyalgia on opioid use was marginally insignificant ($P = 0.06$). It may be that some of the effect of fibromyalgia was reflected in the significant estimate found for joint count; participants with fibromyalgia reported significantly higher joint counts than those without fibromyalgia (mean 10.7 joints versus 4.8 joints). A higher joint count likely further reflects the impact of a greater pain burden on opioid use.

Depression and pain are strongly linked, with each being identified as an important predictor of the other in what is likely a complex bidirectional relationship (35,36). The observed association between greater depressive symptoms and current opioid use that was demonstrated in the present study may reflect this close interrelationship, but could also represent some self-treatment of depressive symptoms with opioid medications (37). A number of studies focused on patients with chronic pain have similarly reported that those with depression or other mental health disorders are more likely to be prescribed opioid therapy, as well as to be prescribed a higher dose (38–40). This is concerning as it has also been reported that individuals with depression are more likely to develop clinically recognized opioid abuse and dependence (41). Our findings taken together with the literature suggest some caution may be warranted in prescribing opioids in individuals with OA and with symptoms of depression or other psychological comorbidities. Although we did not examine data on the use of antidepressants in patients with high depressive symptom scores reporting opioid use, it may be that for some of these individuals with multiple centrally mediated symptoms such as depression and sensitized pain, treatment with medications that target these clusters of symptoms (such as duloxetine) may warrant additional consideration. It has been reported that physicians underdiagnose depression in patients with OA (42) and that few patients receive help for dealing with mental health issues (43–45). Psychological comorbidities may represent a treatable cause of chronic opioid use for some patients with painful chronic conditions like OA.

Strengths of our study include the large sample size overall and the inclusion of patients with hip and knee OA, as well as those with lumbar spine OA. The smaller sample size of this latter group, however, likely contributed to the relatively large width of confidence interval estimates for the joint-specific opioid use rates in Table 3. Future work that includes more patients with spine OA is needed to confirm the higher point estimates that we observed for these patients, particularly among younger women. Our findings should be interpreted with consideration of our study population; participants were all patients with presurgical, end-stage OA. However, our overall rate of opioid use of 30% is in line with the rate of 33% for all OA patients identified in electronic medical records as part of the Canadian Primary Care Sentinel Surveillance Network (46). Our sample does derive from Ontario, Canada, where residents have access to publicly funded universal health care. Differences in access to medical care in other jurisdictions may influence the use of prescription medications. Our study was a retrospective analysis of baseline data from a larger cohort study. We did not have data on specific opioid medications used, dosage, or duration of use. It has been reported that orthopedic patients with hip and knee OA underreport narcotic use in comparison to electronic health records (47). Our findings in terms of the proportions of patients using opioids presurgery are likely to underestimate actual use due to a potential reporting bias from negative media attention around the opioid epidemic and as a result of the use of non-prescribed opioids. Our data were cross-sectional in nature and findings thus reflect associations with an increased likelihood of opioid use; caution should be exercised in assigning causality or directionality. For example, in addition to depression being associated with greater opioid use as discussed above, there is research to suggest that opioid use may also induce symptoms of depression (48,49). It may be informative in future work to investigate whether specific characteristics of OA pain such as its nature or its impacts on certain components of quality life, are more or less likely to be associated with opioid use. Though unavailable for this study, information on duration of OA symptoms, pain-related behaviors, and coping strategies may also be relevant factors to consider in this regard. As factors associated with opioid use may vary with the specific degree of use, additional research with more detailed drug utilization data would be informative.

Despite nearly one-third of patients with presurgical hip, knee, and spine OA currently using opioids in our study and the questionable efficacy and high risk of adverse effects with opioid use, there is a general lack of consensus from professional medical organizations around the appropriate use of these medications in OA. Our findings of greater use in younger individuals with OA and among those with greater depressive symptoms raise further questions as to the appropriateness of current opioid use patterns and highlights the need for more specific guidance. Translation of such recommendations to primary care physicians is of particular

importance, as this is where the vast majority of OA is currently managed. Given the current opioid crisis and the substantial and growing population prevalence of OA (2,3), appropriately limiting the use of opioids in OA may have a significant impact on the numbers of opioids in circulation and thus have a beneficial public health impact as well. Our findings suggest that addressing patient factors such as depression, smoking, and obesity may be beneficial to reducing opioid use, in addition to their positive impacts on patients' overall health. As evidence continues to build that opioid use may negatively impact surgical outcomes (13–16), optimized preoperative pain management and consideration of presurgical opioid use screening, including potential dependency, may also be warranted for patients undergoing surgery for OA.

ACKNOWLEDGMENT

We are grateful for the committed efforts of the clinical research team of the Arthritis Program.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Power 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. Power, Perruccio, Rampersaud.
Acquisition of data. Gandhi, Veillette, Davey, Lewis, Syed, Mahomed, Rampersaud.
Analysis and interpretation of data. Power, Perruccio, Rampersaud.

REFERENCES

1. Murray CJ, Lopez AD, editors. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Boston (MA): Harvard School of Public Health on behalf of the World Health Organization and the World Bank; 1996.
2. Hootman JM, Helmick CG, Barbour KE, Theis KA, Boring MA. Updated projected prevalence of self-reported doctor-diagnosed arthritis and arthritis-attributable activity limitation among US adults, 2015–2040. Arthritis Rheumatol 2016;68:1582–7.
3. 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.
4. Losina E, Paltiel AD, Weinstein AM, Yelin E, Hunter DJ, Chen SP, et al. Lifetime medical costs of knee osteoarthritis management in the United States: impact of extending indications for total knee arthroplasty. Arthritis Care Res (Hoboken) 2015;67:203–15.
5. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010;18:476–99.
6. Bjordal JM, Ljunggren AE, Klovning A, Slordal L. Non-steroidal anti-inflammatory drugs, including cyclo-oxygenase-2 inhibitors, in osteoarthritic knee pain: meta-analysis of randomised placebo controlled trials. BMJ 2004;329:1317.
7. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012;64:465–74.
8. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014;22:363–88.
9. Jevsevar DS. Treatment of osteoarthritis of the knee: evidence-based guideline. J Am Acad Orthop Surg 2013;21:571–6.
10. Lurie J, Tomkins-Lane C. Management of lumbar spinal stenosis. BMJ. 2016. E-pub ahead of print.
11. International Narcotics Control Board (INB). Narcotic Drugs Technical Report: Estimated World Requirments for 2013- Statistics for 2011. New York: United Nations; 2013.
12. Fischer B, Argento E. Prescription opioid related misuse, harms, diversion and interventions in Canada: a review. Pain Physician 2012;15 Suppl:ES191–203.
13. Zywiel MG, Stroh DA, Lee SY, Bonutti PM, Mont MA. Chronic opioid use prior to total knee arthroplasty. J Bone Joint Surg Am 2011;93:1988–93.
14. Pivec R, Issa K, Naziri Q, Kapadia BH, Bonutti PM, Mont MA. Opioid use prior to total hip arthroplasty leads to worse clinical outcomes. Int Orthop 2014;38:1159–65.
15. Smith SR, Katz JN, Collins JE, Solomon DH, Jordan JM, Suter LG, et al. Cost-Effectiveness of tramadol and oxycodone in the treatment of knee osteoarthritis. Arthritis Care Res (Hoboken) 2017;69:234–42.
16. Villavicencio AT, Nelson EL, Kantha V, Burneikiene S. Prediction based on preoperative opioid use of clinical outcomes after transforaminal lumbar interbody fusions. J Neurosurg Spine 2017;26:144–9.
17. Health Canada. Canadian guidelines for body weight classification in adults. Ottawa: Minister of Public Works and Government Services Canada; 2003.
18. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis Rheum 2003;49:156–63.
19. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–70.
20. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res 2002;52:69–77.
21. 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.
22. Franklin PD, Karbassi JA, Li W, Yang W, Ayers DC. Reduction in narcotic use after primary total knee arthroplasty and association with patient pain relief and satisfaction. J Arthroplasty 2010;25 Suppl:12–6.
23. Hansen CA, Inacio MCS, Pratt NL, Roughead EE, Graves SE. Chronic use of opioids before and after total knee arthroplasty: a retrospective cohort study. J Arthroplasty 2017;32:811–7 e1.
24. Armaghani SJ, Lee DS, Bible JE, Archer KR, Shau DN, Kay H, et al. Preoperative opioid use and its association with perioperative opioid demand and postoperative opioid independence in patients undergoing spine surgery. Spine (Phila Pa 1976) 2014;39:E1524–30.
25. Walid MS, Zaytseva NV. Prevalence of mood-altering and opioid medication use among spine surgery candidates and relationship with hospital cost. J Clin Neurosci 2010;17:597–600.

26. Krebs EE, Lurie JD, Fanciullo G, Tosteson TD, Blood EA, Carey TS, et al. Predictors of long-term opioid use among patients with painful lumbar spine conditions. J Pain 2010;11:44–52.
27. Fisher JE, Ballantyne PJ, Hawker GA. Older adults living with osteoarthritis: examining the relationship of age and gender to medicine use. Can J Aging 2012;31:323–33.
28. Wright EA, Katz JN, Abrams S, Solomon DH, Losina E. Trends in prescription of opioids from 2003–2009 in persons with knee osteoarthritis. Arthritis Care Res (Hoboken) 2014;66:1489–95.
29. DeMik DE, Bedard NA, Dowdle SB, Burnett RA, McHugh MA, Callaghan JJ. Are we still prescribing opioids for osteoarthritis? J Arthroplasty 2017;32:3578–82 e1.
30. Berthelot JM, Darrieutort-Lafitte C, Le Goff B, Maugars Y. Strong opioids for noncancer pain due to musculoskeletal diseases: not more effective than acetaminophen or NSAIDs. Joint Bone Spine 2015;82:397–401.
31. Smith SR, Deshpande BR, Collins JE, Katz JN, Losina E. Comparative pain reduction of oral non-steroidal anti-inflammatory drugs and opioids for knee osteoarthritis: systematic analytic review. Osteoarthritis Cartilage 2016;24:962–72.
32. Krebs EE, Gravely A, Nugent S, Jensen AC, DeRonne B, Goldsmith ES, et al. Effect of opioid vs nonopioid medications on pain-related function in patients with chronic back pain or hip or knee osteoarthritis pain: the SPACE randomized clinical trial. JAMA 2018;319:872–82.
33. Lee M, Silverman SM, Hansen H, Patel VB, Manchikanti L. A comprehensive review of opioid-induced hyperalgesia. Pain Physician 2011;14:145–61.
34. Wasserman RA, Brummett CM, Goesling J, Tsodikov A, Hassett AL. Characteristics of chronic pain patients who take opioids and persistently report high pain intensity. Reg Anesth Pain Med 2014;39:13–7.
35. Gureje O, Simon GE, Von Korff M. A cross-national study of the course of persistent pain in primary care. Pain 2001;92:195–200.
36. Vadivelu N, Kai AM, Kodumudi G, Babayan K, Fontes M, Burg MM. Pain and psychology: a reciprocal relationship. Ochsner J 2017;17:173–80.
37. Goesling J, Henry MJ, Moser SE, Rastogi M, Hassett AL, Clauw DJ, et al. Symptoms of depression are associated with opioid use regardless of pain severity and physical functioning among treatment-seeking patients with chronic pain. J Pain 2015;16:844–51.
38. Edlund MJ, Martin BC, Devries A, Fan MY, Braden JB, Sullivan MD. Trends in use of opioids for chronic noncancer pain among individuals with mental health and substance use disorders: the TROUP study. Clin J Pain 2010;26:1–8.
39. Braden JB, Sullivan MD, Ray GT, Saunders K, Merrill J, Silverberg MJ, et al. Trends in long-term opioid therapy for noncancer pain among persons with a history of depression. Gen Hosp Psychiatry 2009;31:564–70.
40. Merrill JO, Von Korff M, Banta-Green CJ, Sullivan MD, Saunders KW, Campbell CI, et al. Prescribed opioid difficulties, depression and opioid dose among chronic opioid therapy patients. Gen Hosp Psychiatry 2012;34:581–7.
41. Edlund MJ, Steffick D, Hudson T, Harris KM, Sullivan M. Risk factors for clinically recognized opioid abuse and dependence among veterans using opioids for chronic non-cancer pain. Pain 2007;129:355–62.
42. Memel DS, Kirwan JR, Sharp DJ, Hehir M. General practitioners miss disability and anxiety as well as depression in their patients with osteoarthritis. Br J Gen Pract 2000;50:645–8.
43. Sale JE, Gignac M, Hawker G. The relationship between disease symptoms, life events, coping and treatment, and depression among older adults with osteoarthritis. J Rheumatol 2008;35:335–42.
44. Vali FM, Walkup J. Combined medical and psychological symptoms: impact on disability and health care utilization of patients with arthritis. Med Care 1998;36:1073–84.
45. Gleicher Y, Croxford R, Hochman J, Hawker G. A prospective study of mental health care for comorbid depressed mood in older adults with painful osteoarthritis. BMC Psychiatry 2011;11:147.
46. Birtwhistle R, Morkem R, Peat G, Williamson T, Green ME, Khan S, et al. Prevalence and management of osteoarthritis in primary care: an epidemiologic cohort study from the Canadian Primary Care Sentinel Surveillance Network. CMAJ Open 2015;3:E270–5.
47. Hereford TE, Cryar KA, Edwards PK, Siegel ER, Barnes CL, Mears SC. Patients with hip or knee arthritis underreport narcotic usage. J Arthroplasty 2018;33:3113–7.
48. Martins SS, Fenton MC, Keyes KM, Blanco C, Zhu H, Storr CL. Mood and anxiety disorders and their association with non-medical prescription opioid use and prescription opioid-use disorder: longitudinal evidence from the National Epidemiologic Study on Alcohol and Related Conditions. Psychol Med 2012;42:1261–72.
49. Scherrer JF, Salas J, Lustman PJ, Burge S, Schneider FD; Residency Research Network of Texas I. Change in opioid dose and change in depression in a longitudinal primary care patient cohort. Pain 2015;156:348–55.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1186–1193
DOI 10.1002/acr.23740

Thresholds in the Relationship of Quadriceps Strength With Functional Limitations in Women With Knee Osteoarthritis

Kathryn L. Bacon,[1] Neil A. Segal,[2] Britt Elin Øiestad,[3] Cora E. Lewis,[4] Michael C. Nevitt,[5] Carrie Brown,[6] Michael P. LaValley,[1] Charles E. McCulloch,[5] and David T. Felson[7]

Objective. To investigate thresholds of strength below which individuals with knee osteoarthritis (OA) may have more difficulty carrying out physical functions of daily life. Individuals below such thresholds might benefit more from strengthening interventions than those with greater strength.

Methods. We studied individuals with symptomatic OA at baseline in the Multicenter Osteoarthritis Study who had knee extensor strength measured isokinetically at 60°/second. Participants underwent a 20-meter walk test and a sit-to-stand test and answered questions from the Western Ontario and McMaster Universities Osteoarthritis Index. Physical function results were plotted against measures of quadriceps strength (Nm) (and as strength:body weight) for the worst knee. Loess technique was used to examine inflection points. Nonlinear relationships were examined in piecewise linear regression models. Differences were tested using linear and logistic regression models.

Results. The study had 834 participants (65.8% women). The mean ± SD age of the participants was 62.9 ± 7.9 years. In women, there were thresholds of strength below which the slope of strength versus function was steeper: walking speed (<58 Nm), chair stand time (<32 Nm), and the McMaster Universities Osteoarthritis Index functions of rising from a chair and getting on/off the toilet (<38 Nm). We found no thresholds in men. Loess analyses using strength:weight showed similar results.

Conclusion. In individuals with symptomatic knee OA, thresholds in the strength function relationship may help identify individuals, especially women, at the brink of disability insofar as strength and capacity for daily tasks. In those with low strength, small increments in strength may be associated with improvement in function and greater ease with common daily life, emphasizing the importance of preventing loss of strength.

INTRODUCTION

Studies of individuals with knee osteoarthritis (OA) and those at risk of OA have shown that quadriceps weakness is strongly associated with functional limitations (1–3). For these individuals, addressing quadriceps strength often becomes a major target of rehabilitation interventions, including general aerobic and local strengthening exercises, with the hope of improving physical function, reducing pain and physical disability, and avoiding further progression of OA (4,5).

Although analyses may implicitly assume a linear association between strength and functional limitations (i.e., gains in strength lead to gains in function), daily tasks may require a specific amount of strength to be successfully accomplished (6,7). This fact suggests that the associations of quadriceps strength with measures of physical function may not be linear (8,9); there may be thresholds of strength necessary to accomplish some tasks, such as rising from a chair. In weaker adults, with strength below the required threshold for the task, training may have immediate benefits in improved function, while in stronger adults, strength training may yield little apparent improvement in physical performance because the individual may already be stronger than the threshold (8–10). For stronger individuals, strength training may assist in maintaining function or progressing to more strenuous activities.

Supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases Multidisciplinary Clinical Research Center (grant P60-AR-047785), the National Institute on Aging (grants U01-AG-18820, U01-AG-18832, U01-AG-18947, and U01-AG-19069), the Multicenter Osteoarthritis Study (MOST) Second Renewal, and a Rheumatology Training Grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant 5-T32-AR-7598-18).

[1]Kathryn L. Bacon, PhD, Michael P. LaValley, PhD: Boston University, Boston, Massachusetts; [2]Neil A. Segal, MD, MS: University of Kansas, Kansas City; [3]Britt Elin Øiesta, PT, PhD: Oslo Metropolitan University, Oslo, Norway; [4]Cora E. Lewis, MD, MSPH: University of Alabama, Birmingham; [5]Michael C. Nevitt, PhD, Charles E. McCulloch, PhD: University of California, San Francisco; [6]Carrie Brown, MS: Boston University School of Medicine, Boston, Massachusetts; [7]David T. Felson, MD, MPH: Boston University, Boston, Massachusetts, NIHR Manchester Musculoskeletal Biomedical Research Centre, Manchester University NHS Foundation Trust, and Manchester Academic Health Science Centre, Manchester, UK.

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

Address correspondence to Kathryn L. Bacon, PhD, 650 Albany Street, Suite X200, Clinical Epidemiology Research and Training Unit, Boston, MA 02118. E-mail: kbacon@bu.edu.

Submitted for publication October 18, 2017; accepted in revised form August 21, 2018.

SIGNIFICANCE & INNOVATIONS

- We explored quadriceps strength and physical function in individuals with knee osteoarthritis from the Multicenter Osteoarthritis Study.
- We found evidence of strength thresholds for functional tasks that involve the body acting against gravity, such as walking or rising from a chair, in women but not in men.
- These thresholds in strength may provide a means of identifying individuals at the brink of disability, for whom increments in strength may be associated with improvements in physical function or greater ease in carrying out common tasks in daily life; for individuals above the strength threshold, focus should be on maintaining strength.

Despite the many factors that may contribute to functional decline, including health-related behaviors, medical conditions, socioeconomic status, and psychological well-being, muscle strength remains of key importance to performance (7,9,11). In a study using performance-based tests of walking speed, chair stands, and standing balance in a sample of elderly women, the relationship of strength to performance was independent of age, weight, and height and was largely nonlinear (9). Although the variance in performance explained by strength alone was >20% for each measure tested, with respect to the disabling process, the role played by reduction of strength was of particular importance in the weaker subset of the population (9).

Individuals with painful knee OA are often referred to physical therapy to work on strengthening, with the goal of improving function and pain, because strengthening is feasible for all age groups and patients with or without comorbidities. Using data from the Multicenter Osteoarthritis Study (MOST), the purpose of the current analysis was to identify thresholds of quadriceps strength below which performance of basic physical function tasks is adversely affected. The basic functional tasks chosen for this analysis involve quadriceps strength to move the body in opposition to gravity: rising from a chair, going up stairs, getting on and off the toilet, and walk time. We included performance-based tests when possible (walking and rising from a chair) as well as self-reported measures of function, because these may represent different components of function in reflecting what individuals perceive they can do versus what they can actually do (12,13). We hypothesized that across the range of quadriceps strength values, the relationship of strength and functional tasks would be nonlinear. Further, we hypothesized that there are strength thresholds for each functional task, such that below the threshold, the relationship of strength and function would be steeper than the relationship above the threshold. Strengthening might achieve a greater improvement in function for those below the threshold.

PATIENTS AND METHODS

Study population. MOST is a cohort study of 3,026 men and women ages 50–79 years at baseline, at risk of knee OA (i.e., overweight, obese, with a history of knee injury, or with frequent knee pain) or with established knee OA. The study participants were from Birmingham, Alabama and Iowa City, Iowa (14). The study started in 2003 when study participants were interviewed by telephone and attended clinic visits. Further details of inclusion and exclusion criteria have been published previously (14,15). The initial visit included examination age, height, weight, self-reported physical activity level, knee extensor muscle strength, self-reported physical function and pain, radiographic evaluation, and magnetic resonance imaging.

For this analysis, we excluded those individuals who had total knee replacement at the baseline MOST visit and any individuals who had total hip replacement at any time during MOST follow-up, because these individuals may have had hip OA at baseline that affected the studied associations. We focused on individuals with symptomatic OA at the baseline visit who had knee extensor strength measurements (2). Frequent knee symptoms were assessed by questionnaire; radiographic knee OA was assessed with fixed flexion posteroanterior and lateral weight-bearing radiographs. We defined symptomatic knee OA as present when the individual reported frequent knee pain, aching, or stiffness on most days when asked about it during their clinic visit and whose radiograph demonstrated a Kellgren/Lawrence (K/L) grade ≥2 in any compartment in that knee (16). Participants with bilateral OA or using assistive devices were included. All measures are from the baseline MOST visit.

Exposure: knee extensor muscular strength. Concentric isokinetic strength of the quadriceps knee extensor was measured using a Cybex 350 isokinetic dynamometer at 60º/second. After instruction, and 3 practice trials, participants completed 4 repetitions; peak torque over 4 repetitions was used for concentric knee extensor strength (2). Strength from the worst knee by K/L grade was used as a continuous variable (strength in Nm). Secondary analyses used the ratio of strength (Nm) to body weight (kg) (17,18) as the exposure.

Measures of physical function. Measures of physical function described below and chosen for this analysis involve employment of quadriceps strength to move the body in opposition to gravity: rising from a chair, going up stairs, and getting on and off the toilet, and also included the major functional activity of walking.

20-meter walk test. We measured total time in seconds for study participants to walk at their usual walking pace from the starting point to the end. The test was then repeated, and we used the mean time of the 2 test trials, in seconds. Refusal or inability to do the test or use of walking aids were recorded.

5 times sit-to-stand test. Using a chair with a straight back, a flat, level, and firm seat, and seat height 45 cm at front, participants stood up from the chair 5 times as quickly as they

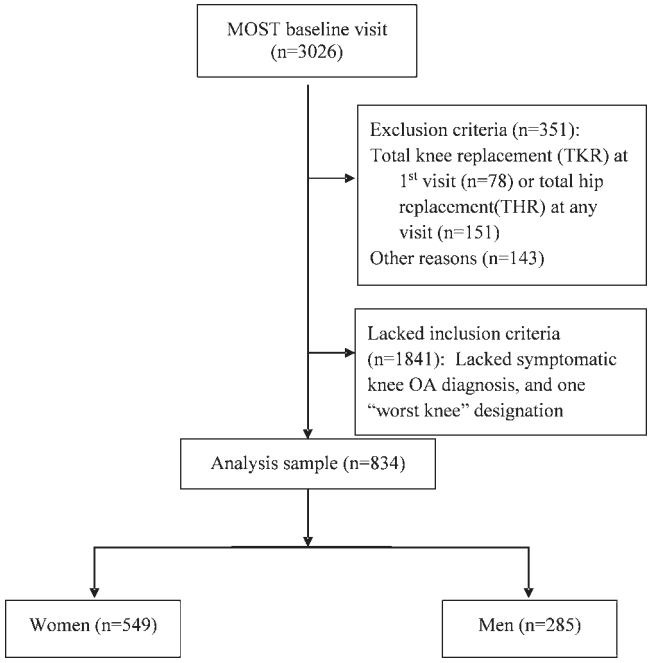

Figure 1. Analysis sample flow diagram. MOST = Multicenter Osteoarthritis Study.

could, keeping their arms folded across their chest. Walking aids were not allowed. Refusal or inability to do the test were recorded. We recorded total time in seconds using a stopwatch from start to finish of the test; the test was then repeated, and we used the mean time of the 2 test trials, in seconds.

Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). We chose 3 individual questions from the WOMAC survey that use quadriceps strength and are similar to the performance-based measures: going up stairs, rising from a chair, and getting on and off the toilet. For each of these outcomes, respondents answered a question "How much difficulty have you had…" performing the task. For each question, the possible responses were none (0 = best performance level), mild (score = 1), moderate (score = 2), severe (score = 3), or extreme (4 = worst performance level), yielding an ordinal value ranging from 0 to 4 for each WOMAC item outcome.

Additionally, for each WOMAC outcome, we created a dichotomous variable that classified individuals as having a high level of difficulty with the task (combining the scores 2, 3, and 4) or a low level of difficulty (combining the scores 0 and 1). The dichotomization of this measure made it easier to explain, and divided participants into high and low categories near the median number of respondents. Covariates were age in years as a continuous variable and weight in kilograms at baseline as a continuous variable.

Statistical analysis. We took an exploratory and graphic approach to identify inflection points, or thresholds, in the relationship between measures of physical function and quadriceps strength. We stratified by sex due to differences in the distributions of strength and height in men and women. Our primary analyses adjusted for age and weight, which have been linked to performance in other studies of function (6,8,9,19). Analyses were performed using SAS software, version 9.4.

Participant characteristics were summarized with frequencies and mean. For continuous outcomes (the 20-meter walk test and the 5 times sit-to-stand test), we graphically explored the shape of the relationship between quadriceps strength and each outcome, using nonparametric loess technique (20), separately for men and women, seeking possible inflection points that might indicate a threshold of strength that alters the relationship with physical function. A range of potential inflection points around the

Table 1. Sample baseline characteristics*

Characteristic	Men (n = 285)	Women (n = 549)
Age, years	62.5 ± 8.3	63.1 ± 7.7
Weight, kg	101.7 ± 21.7	89.1 ± 20.2
Body mass index, kg/m^2	32.2 ± 6.5	33.5 ± 7.3
K/L grade at baseline in worst knee, %		
Grade 2	26	36
Grade 3	42	43
Grade 4	32	21
Knee extensor muscle strength, Nm	105.3 ± 38.7	51.5 ± 22.8
Cutoffs for quintiles of strength, Nm		
Quintile 1	<73	<33
Quintile 2	73–92	33–43
Quintile 3	93–113	44–53
Quintile 4	114–133	54–66
Quintile 5	≥134	≥67
Normalized strength†	1.1 ± 0.4	0.6 ± 0.3
Performance-based outcomes		
Chair stand (5 chair stands, average time), seconds	11.9 ± 3.9	13.8 ± 5.2
Walking time (20 meters), seconds	17.2 ± 3.3	19.3 ± 5.7
WOMAC self-reported outcomes, %‡		
Going up stairs (high difficulty)	55	70
Rising from sitting position (high difficulty)	48	64
Getting on and off the toilet (high difficulty)	33	40

* Values are the mean ± SD unless indicated otherwise. Normalized strength (Nm:kg) was used only in secondary analyses. K/L = Kellgren/Lawrence.
† Strength:weight ratio in Nm:kg.
‡ Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) questions are for either knee, in the past 30 days, dichotomized as high level of difficulty versus low level of difficulty with this task.

value suggested by inspection of the graphs was then fit in separate piecewise models (21) and tested with nonlinear least squares regression for convergence to a solution that best fit the data. We limited points to those between the 20th and 80th percentiles of strength for stability of the models. Our primary models used strength as the exposure, adjusted for age and weight; major differences in height were already addressed by stratification by sex. Although we report *P* values examining the change in slope below and above the specific inflection points tested, these *P* values were considered exploratory aids to guide selection of points and not definitive tests of the exact locations. Using these selected inflection points as fixed quantities in a regression model, the change in slope was tested to see whether it was non-zero, using an F test.

For WOMAC outcomes, we used proportional odds logistic regression with restricted cubic splines to evaluate nonlinear relationships with quadriceps strength (22). A range of potential inflection points around the points suggested by the cubic spline models was fit using piecewise linear trends in the log odds using separate logistic regression models. The –2 log-likelihood was used to aid in the selection of an inflection point between the 20th and 80th percentiles of strength, and then the change in slope was tested using a chi-square test in the proportional odds regression model with the fixed inflection point.

In secondary analyses for the performance-based outcomes, we tested a set of models using the ratio of strength:weight (i.e., normalized strength) as the exposure, first stratified by sex and adjusted only for age; then using the sample of men and women together and adjusted for age and height. The ratio of strength:weight (Nm:kg) is less interpretable than strength but accommodates different ranges of strength and/or weight: men and women may have the same ratio but very different strength and weight values.

RESULTS

Among the 834 study participants (Figure 1), the mean ± SD age was 62.9 ± 7.9 years, and the mean ± SD body mass index was 33.1 ± 7.1 kg/m^2; 65.8% of the participants were women (Table 1). All participants had symptomatic OA in at least 1 knee, with 64% of both men and women having a maximal K/L score of 3 or 4 in the worse knee. The mean ± SD quadriceps strength in women (52 ± 23 Nm) was lower than that in men (105 ± 39 Nm).

Results in women. In women, the loess plot for the performance-based 20-meter walk test versus quadriceps strength (Figure 2) suggested the presence of an inflection point in strength near 60 Nm. In a piecewise linear model, the slope of strength versus walking time became steeper at quadriceps strength of 56.8 Nm (slope before 56.8 Nm = –0.11, slope after 56.8 Nm = –0.02; *P* = 0.001 for difference in slopes), indicating a steeper relation of strength with elapsed walking time below this threshold in women.

The loess plot for chair stand test results versus quadriceps strength (Figure 3) suggested an inflection point below 50 Nm; in

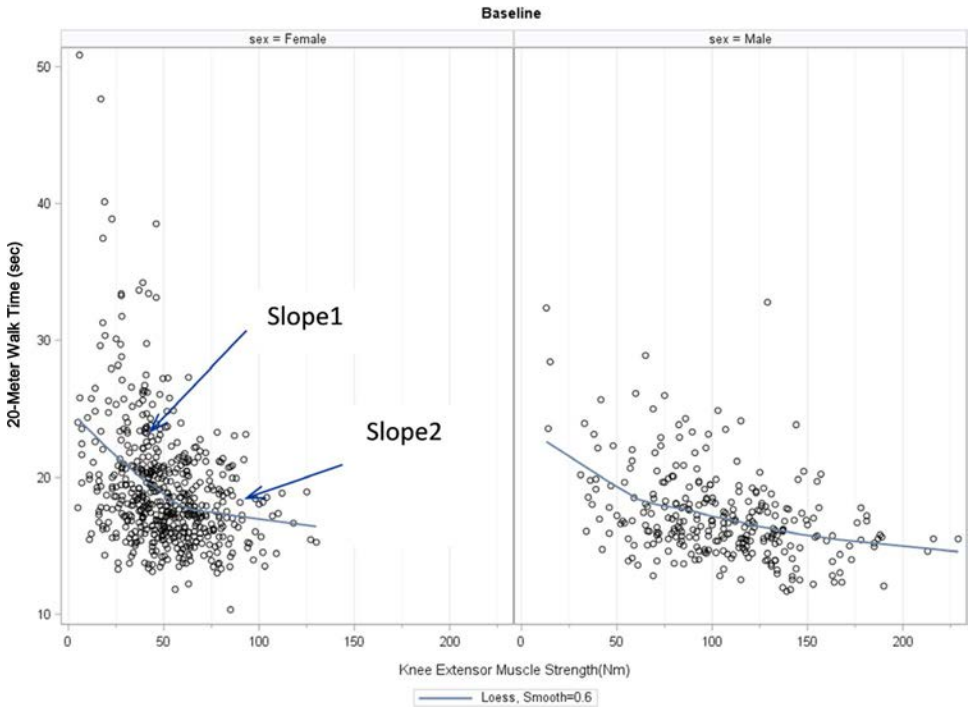

Figure 2. Knee extensor muscle strength versus 20-meter walk time, by sex. Nonparametric local weighted regression (LOESS) was used to generate the scatter plots and LOESS lines. In women, the piecewise linear model indicated a possible inflection point at 56.8 Nm (arrows indicate segments of the LOESS line: slope 1 = –0.11, slope 2 = –0.02; *P* = 0.001 for difference in slopes). In men, no inflection points were found in the 20–80% range of strength. sec = seconds.

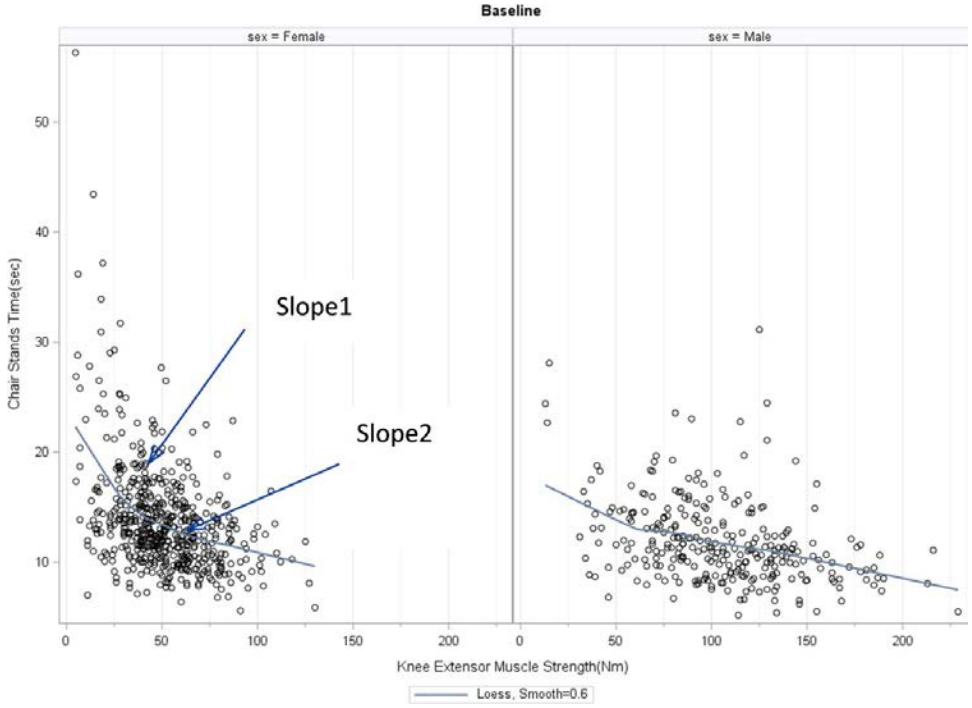

Figure 3. Knee extensor muscle strength versus chair stands time, by sex. Nonparametric local weighted regression (LOESS) was used to generate the scatter plots and LOESS lines. In women, the piecewise linear model indicated a possible inflection point at 32 Nm (arrows indicate segments of the LOESS line: slope 1 = –0.30, slope 2 = –0.05; $P < 0.0001$ for difference in slopes). In men, no inflection points were found in the 20–80% range of strength. sec = seconds. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23740/abstract.

a piecewise linear model we found the slope of strength versus chair stands time became steeper at strength 32 Nm (slope 1 = –0.30, slope 2 = –0.05; $P < 0.0001$ for difference in slopes). Loess analyses using strength:weight as the predictor showed similar graphic relationships.

In unadjusted proportional odds models with restricted cubic splines, we found evidence of nonlinearity in the association of strength and the WOMAC items rising from a chair and getting on and off the toilet in women (data not shown). The joint effect of nonlinear components in the restricted cubic spline models for rising from a chair was of borderline significance ($\chi^2 = 5.99$, degrees of freedom [df] = 2, $P = 0.05$), and became nonsignificant after adjustment for age and weight ($\chi^2 = 5.13$, df = 2, $P = 0.08$). For getting on and off the toilet, neither association was statistically significant (unadjusted $\chi^2 = 4.95$, df = 2, $P = 0.08$; adjusted $\chi^2 = 4.49$, df = 2, $P = 0.11$). For WOMAC going up stairs we found no significant evidence of nonlinearity in the restricted cubic spline models ($P = 0.23$ in age- and weight-adjusted models).

In explorations of piecewise models in individual WOMAC items in women (Figure 4), we found suggestions of an inflection point in quadriceps strength at approximately 38 Nm for both difficulty rising from a chair (slope 1 = –0.034, slope 2 = –0.024; $P = 0.12$ for difference in slopes) and getting on and off the toilet (slope 1 = –0.038, slope 2 = –0.029; $P = 0.05$ for difference in slopes). After adjustment for age and weight, the changes in slope were reduced (rising from a chair slope 1 = –0.031, slope 2 = –0.019; $P = 0.20$; getting on and off the toilet slope 1 = –0.035, slope 2 = –0.023; $P = 0.11$). The association of quadriceps strength and difficulty going up stairs (data not shown) appeared to be linear; we found no suggestion of inflection points between the 20th to 80th percentiles of strength in women (highly nonsignificant P values >0.37 for tests of inflection points at 33 Nm through 87 Nm of strength in women).

Results in men. Men had a wider range of strength values than women (Table 1). For men, the loess plots and models for walking time (Figure 2) and chair stands time (Figure 3) suggested approximately linear relationships between quadriceps strength and these outcomes. In men, piecewise linear models suggested no inflection points for walking time or chair stands time in the 20th to 80th percentile range of strength. Mean walking time for the 20-meter walk test was faster than 1.0 meter/second for all men in this sample.

In men, restricted cubic spline models yielded no suggestions of nonlinearity in the association of quadriceps strength and the WOMAC items (tests of joint effect of nonlinear components in adjusted models: going up stairs, rising from a chair, and getting on and off the toilet, $P = 0.20$, $P = 0.39$, and $P = 0.75$, respectively). In explorations of piecewise models for individual WOMAC items (Figure 4) we found no visual evidence of inflection points in the range 20th to 80th percentiles of strength in men, with nonsignificant P values for tests of inflection points at 73 through 134 Nm of strength.

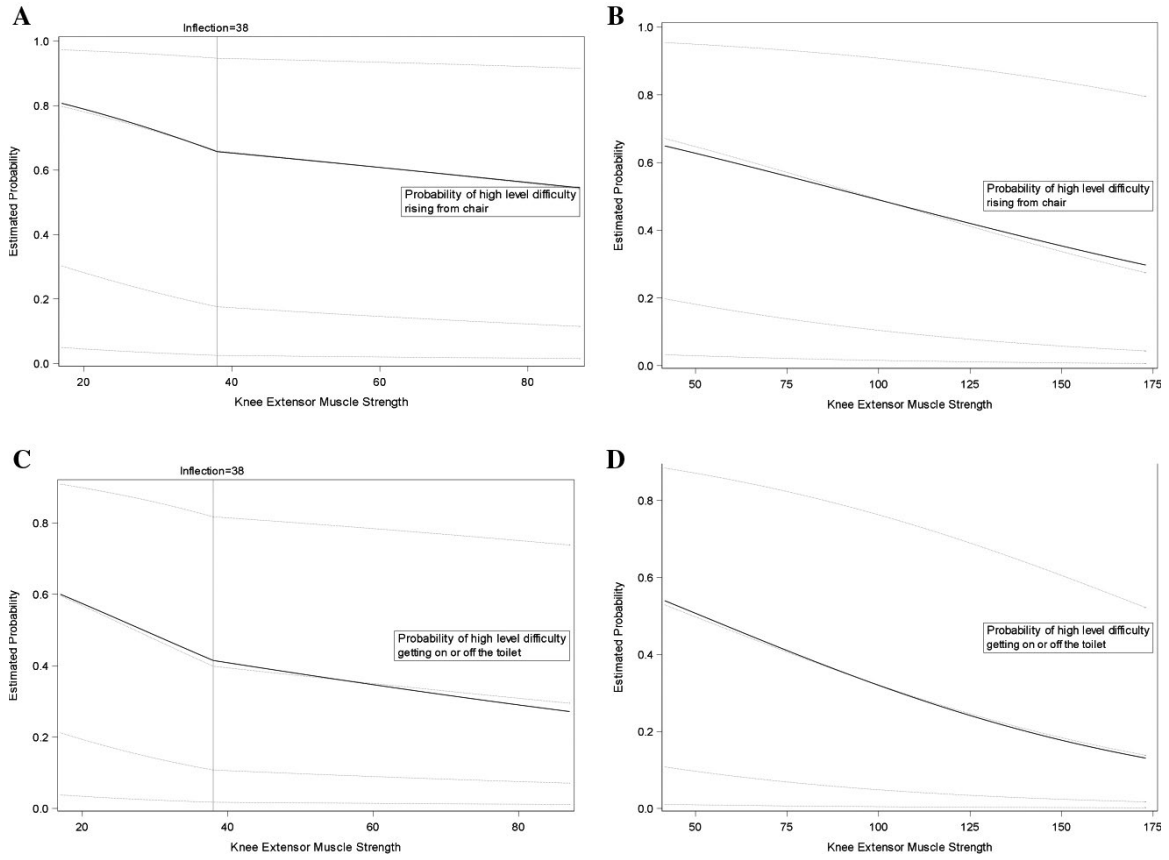

Figure 4. Estimated probabilities for association of quadriceps strength and selected Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) items from logistic piecewise regression models. Figures show estimated probabilities for the association of strength and each WOMAC item. Graphs combine both ordinal and dichotomous WOMAC outcomes variables, from unadjusted models. Plots are restricted to a 5–95% range of strength distribution by sex. Possible inflection points of strength are noted on graphs, where found. Dark lines indicate probability of "high level of difficulty" for dichotomous WOMAC outcomes (combining WOMAC extreme, severe, or moderate level of difficulty). Dotted lines indicate ordinal WOMAC variable probabilities. **A**, Difficulty rising from a chair (women); **B**, Difficulty rising from a chair (men); **C**, Difficulty getting on and off the toilet (women); **D**, Difficulty getting on and off the toilet (men).

Secondary analyses. In secondary analyses using normalized strength as the exposure (i.e., the ratio of strength: weight), in men and women together, for the 20-meter walk test the estimated inflection point was 0.74 Nm:kg (slope 1 = –8.761, slope 2 = 7.757; $P < 0.0001$ adjusted for age and height). Of 494 individuals with a strength:weight ratio below 0.7432 Nm:kg, 83% were women and 17% were men. Reviewing the characteristics of individuals with a strength:weight ratio of <0.74 Nm:kg, we found that women in that range had a mean strength of 42 Nm, well within the 20th to 80th percentiles of strength in women. Men had a mean strength of 58 Nm in that range, below the 20th percentile of strength in men. Only 22 men were below the inflection point for the ratio, but above the 20th percentile in strength for men.

For tests of chair stand time with normalized strength, in men and women together, the estimated inflection point was 0.30 Nm:kg (slope 1 = –37.274, slope 2 = 33.269; $P < 0.0001$ adjusted for age and height). However, the mean strength of individuals with that ratio of strength:weight is not above the 20th percentile of strength for either men or women.

DISCUSSION

In this study in individuals with knee OA, we found suggestions of strength inflection points in associations between quadriceps strength and physical function measures in women but not in men. If these cross-sectional data are applicable to treatment effects, they suggest that rehabilitation strategies that focus on strengthening are likely to be more effective in improving function in women who are weaker than those who are stronger, including men. For those above the inflection point, our results may provide an incentive to maintain those higher levels of strength. The results also suggest that a quantitative assessment of strength might be indicated prior to rehabilitation planning.

Analyses were focused on common functional tasks that require quadriceps strength, arising from a sitting position and going up stairs. Walk time was included because it is one of the most commonly evaluated measures of physical performance (23). Strength inflection points in women were in the vicinity of 57 Nm for the 20-meter walk test and 32 Nm for the 5 times sit-

to-stand test. In concordance with the sit-to-stand test results, we found suggestions of strength inflection points near 38 Nm for WOMAC items involving similar physical actions: rising from a chair and getting on and off the toilet. For women, these strength estimates were close to the median value of 52 Nm and the cut-off for the lowest quintile of 33 Nm, respectively. These values provide thresholds in the sense that for knee extensor strength values above the threshold, the relationship between function and strength increased more slowly, but we did not find thresholds where the relationship flattened out with no improvement in function above the threshold value. A comparison may serve to contrast relationships below or above the strength inflection point estimated for the 20-meter walk test. Our model estimates that for a 20% increase in strength for individuals whose strength fell below the inflection point, the improvement in walking time was nearly a second (for strength 40 Nm and 48 Nm, predicted walking times are 19.8 and 18.9 seconds, respectively), while for those above the inflection point, a 20% increase in strength resulted in a 0.23-second faster walk time (for strength 60 Nm and 71 Nm, predicted walking times were 17.7 and 17.5 seconds, respectively.).

We have estimated the location of inflection points using visual and exploratory methods. While we are confident that these inflections occur close to the strength values presented here, the exact placements are likely different from our values. The *P* values used for testing individual threshold values suffer from a multiple testing problem, and small differences in thresholds would be very challenging to detect even with very large observational data sets.

Our findings are similar to other studies exploring thresholds of strength and function, that have not focused on individuals with arthritis (6,8–10,24–26). In a population-based sample of older women with significant functional limitations, Ferrucci et al (9) found largely nonlinear associations between both hip flexor and knee extensor muscle strength and performance-based tests of walking speed, chair stands, and standing balance. Another study in community-dwelling women found thresholds of quadriceps strength below which performance of basic ambulatory tasks (gait speed, chair rise, stair ascent and descent) was likely to be compromised (26).

Our findings, while cross-sectional and not examining treatment effects, have important implications for rehabilitation strategies. Identification and use of functionally relevant thresholds of quadriceps strength enables both the identification of appropriate patients and therapeutic guidance for rehabilitation interventions, with the goal of preserving physical function in individuals aging with knee OA. Knowing the specific targets for strength that may prevent or delay disability in women with symptomatic knee OA advances our ability to provide strengthening interventions that are most likely to result in clinically meaningful functional improvement. For those individuals whose strength is below the identified thresholds, strengthening interventions may improve functional performance on relevant tasks such as getting out of a chair and ascending stairs, while for those who are not weak and who are struggling with these functions, other rehabilitation interventions with less focus on strengthening, e.g., weight management, or attention to flexibility, or balance, may be preferred. This stratification of patients echoes personalized medicine approaches and would be consistent with arguments favoring phenotype differences among OA patients (27–29) that would motivate different treatment approaches.

Although our data suggest thresholds of strength in women below which the strength function relationship is steeper, we found no such thresholds in men, who were in general much stronger than women (30,31). This sex difference is likely because a certain threshold of strength is needed to carry out these particular tasks and few men fell below this threshold. Therefore, we lacked statistical power to evaluate this relationship. In secondary analyses using the ratio of strength:weight as the exposure, for the 20-meter-walk test, only 17% of the individuals with the estimated ratio 0.74 Nm:kg were men. Men and stronger women may encounter thresholds with more difficult tasks, which were not included in this study.

The MOST data set has several key strengths for this type of analysis. MOST is a community sample that did not select individuals with OA based on the severity of disease. Participants in this sample were chosen to mirror a typical clinical situation in which patients with knee OA may be referred to physical therapy for strengthening. Although many individuals in this sample had advanced OA by K/L grade (Table 1) in the worst knee, this study should still be broadly representative of OA severity in the community, and therefore be representative of those seeking rehabilitative treatment for OA. Knee extensor strength was measured using an isokinetic dynamometer, measuring knee strength while the leg is in motion, rather than pressing against a static instrument. We included both men and women, and because all participants had some level of OA, they represent a sample that is neither severely disabled, nor extremely healthy, but rather span a range of performance.

A limitation of this analysis is that it is cross-sectional, although one might claim that the effects of strength on function should be immediate. This is not a treatment study, so the effectiveness of strengthening at different levels of advanced disease can not be predicted. In addition, control for other factors that affect function, such as pain, other muscle involvement, medical conditions, or psychosocial factors possibly would diminish the association of quadriceps strength with function, although quadriceps strength is likely to still be a major contributing factor for physical function.

Future studies should evaluate longitudinal associations of changes in strength and changes in function, exploring whether inflection points that appeared for some outcomes in this cross-sectional analysis are present in longer-term analyses, e.g., do weaker women (below the threshold at baseline) have different response over time compared to stronger women, or men? Future analyses should include individuals without OA and explore the role of other factors besides quadriceps strength, which may affect function, such as pain, hip or other muscle strength, other medical conditions, or psychosocial factors such as self-efficacy or depressive symptoms, because consideration of these factors may aid in tailoring rehabilitation programs as well. More difficult or strenuous measures of func-

tion and endurance should be evaluated to investigate other potential thresholds of strength that may occur in stronger individuals.

In summary, the major finding of this study in individuals with knee OA is evidence of strength thresholds in several basic functional tasks, such as rising from a chair, in women but not in men. Although other factors also affect physical function capabilities, these thresholds in strength may provide a means of identifying individuals at the brink of disability, insofar as the contribution of quadriceps strength to basic functional tasks in daily life. These individuals may benefit most from strengthening interventions.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Bacon 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. Bacon, Segal, Øiestad, Brown, Felson.
Acquisition of data. Lewis, Nevitt, Felson.
Analysis and interpretation of data. Bacon, Brown, LaValley, McCulloch, Felson.

REFERENCES

1. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson P, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. Am J Public Health 1994;84:351–8.
2. Segal NA, Torner JC, Felson D, Niu J, Sharma L, Lewis CE, et al. Effect of thigh strength on incident radiographic and symptomatic knee osteoarthritis in a longitudinal cohort. Arthritis Rheum 2009;61:1210–7.
3. Culvenor AG, Ruhdorfer A, Juhl C, Eckstein F, Øiestad BE. Knee extensor strength and risk of structural, symptomatic, and functional decline in knee osteoarthritis: a systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2017;69:649–58.
4. Alnahdi AH, Zeni JA, Snyder-Mackler L. Muscle impairments in patients with knee osteoarthritis. Sports Health 2012;4:284–92.
5. Roos EM, Herzog W, Block JA, Bennell KL. Muscle weakness, afferent sensory dysfunction and exercise in knee osteoarthritis. Nat Rev Rheumatol 2011;7:57–63.
6. Buchner DM, Beresford SA, Larson EB, LaCroix AZ, Wagner EH. Effects of physical activity on health status in older adults II: intervention studies. Annu Rev Public Health 1992;13:469–88.
7. Yoshioka S, Nagano A, Hay DC, Fukashiro S. The minimum required muscle force for a sit-to-stand task. J Biomech 2012;45:699–705.
8. Buchner DM, Larson EB, Wagner EH, Koepsell TD, de Lateur BJ. Evidence for a non-linear relationship between leg strength and gait speed. Age Ageing 1996;25:386–91.
9. Ferrucci L, Guralnik JM, Buchner D, Kasper J, Lamb SE, Simonsick EM, et al. Departures from linearity in the relationship between measures of muscular strength and physical performance of the lower extremities: the Women's Health and Aging Study. J Gerontol A Biol Sci Med Sci 1997;52:M275–85.
10. Buchner DM, Cress ME, de Lateur BJ, Esselman PC, Margherita AJ, Price R, et al. The effect of strength and endurance training on gait, balance, fall risk, and health services use in community-living older adults. J Gerontol A Biol Sci Med Sci 1997;524:M218–24.
11. Lord SR, Murray SM, Chapman K, Munro B, Tiedemann A. Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. J Gerontol A Biol Sci Med Sci 2002;57:M539–43.
12. Kennedy D, Stratford PW, Pagura SM, Walsh M, Woodhouse LJ. Comparison of gender and group differences in self-report and physical performance measures in total hip and knee arthroplasty candidates. J Arthroplasty 2002;17:70–7.
13. Davis AM. Osteoarthritis year 2011 in review: rehabilitation and outcomes. Osteoarthritis Cartilage 2012;20:1–6.
14. Felson DT, Niu J, Guermazi A, Roemer F, Aliabadi P, Clancy M, et al. Correlation of the development of knee pain with enlarging bone marrow lesions on magnetic resonance imaging. Arthritis Rheum 2007;56:2986–92.
15. Segal NA, Nevitt MC, Gross KD, Hietpas J, Glass NA, Lewis CE, et al. The Multicenter Osteoarthritis Study (MOST): opportunities for rehabilitation research. PM R 2013;5:647–54.
16. Kellgren J, Lawrence J. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
17. Bazett-Jones DM, Cobb SC, Joshi MN, Cashin SE, Earl JE. Normalizing hip muscle strength: establishing body-size-independent measurements. Arch Phys Med Rehabil 2011;92:76–82.
18. Jaric S. Muscle strength testing. Sports Med 2002;32:615–31.
19. Skelton DA, Greig CA, Davies JM, Young A. Strength, power and related functional ability of healthy people aged 65-89 years. Age Ageing 1994;23:371–7.
20. Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. J Am Stat Assoc 1988;83:596–610.
21. Ryan SE, Porth LS. A tutorial on the piecewise regression approach applied to bedload transport data. In: Gen. Tech. Rep. RMRS-GTR-189. Fort Collins (CO): U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station; 2007.
22. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. Regression methods in biostatistics. In: Linear, logistic, survival, and repeated measures models. 2nd ed. New York: Springer; 2012.
23. Bennell K, Dobson F, Hinman R. Measures of physical performance assessments: Self-Paced Walk Test (SPWT), Stair Climb Test (SCT), Six-Minute Walk Test (6MWT), Chair Stand Test (CST), Timed Up & Go (TUG), Sock Test, Lift and Carry Test (LCT), and Car Task. Arthritis Care Res (Hoboken) 2011;63 Suppl:S350–70.
24. Grindem H, Snyder-Mackler L, Moksnes H, Engebretsen L, Risberg MA. Simple decision rules can reduce reinjury risk by 84% after ACL reconstruction: the Delaware-Oslo ACL cohort study. Br J Sports Med 2016;50:804–8.
25. Kuenze C, Hertel J, Saliba S, Diduch DR, Weltman A, Hart JM. Clinical thresholds for quadriceps assessment after anterior cruciate ligament reconstruction. J Sport Rehabil 2015;24:36–46.
26. Ploutz-Snyder LL, Manini T, Ploutz-Snyder RJ, Wolf DA. Functionally relevant thresholds of quadriceps femoris strength. J Gerontol A Biol Sci Med Sci 2002;57:B144–52.
27. Knoop J, van der Leeden M, Thorstensson CA, Roorda LD, Lems WF, Knol DL, et al. Identification of phenotypes with different clinical outcomes in knee osteoarthritis: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2011;63:1535–42.
28. Felson DT. Identifying different osteoarthritis phenotypes through epidemiology. Osteoarthritis Cartilage 2010;18:601–4.
29. Dell'Isola A, Allan R, Smith SL, Marreiros SS, Steultjens M. Identification of clinical phenotypes in knee osteoarthritis: a systematic review of the literature. BMC Musculoskelet Disord 2016;17:425.
30. Newman AB, Kupelian V, Visser M, Simonsick EM, Goodpaster BH, Kritchevsky SB, et al. Strength, but not muscle mass, is associated with mortality in the health, aging and body composition study cohort. J Gerontol A Biol Sci Med Sci 2006;61:72–7.
31. Segal NA, Glass NA, Felson DT, Hurley M, Yang M, Nevitt M, et al. Effect of quadriceps strength and proprioception on risk for knee osteoarthritis. Med Sci Sports Exerc 2010;42:2081–8.

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1194–1201
DOI 10.1002/acr.23772

American College of Rheumatology
Empowering Rheumatology Professionals

Factors Associated With Patients' Willingness to Consider Joint Surgery After Completion of a Digital Osteoarthritis Treatment Program: A Prospective Cohort Study

Anna Cronström,[1] Håkan Nero,[2] and Leif E. Dahlberg[3]

Objective. To examine patient willingness and a possible shift in willingness for surgery and to investigate factors associated with this shift, following participation in the digital nonsurgical osteoarthritis (OA) treatment program Joint Academy.

Methods. A total of 458 individuals (mean ± SD age 62 ± 5.6 years, 67% women) with diagnosed hip or knee OA were evaluated after 6 weeks in the Joint Academy program, comprising education and exercise as well as asynchronous chat with a physical therapist. Data describing joint pain, health-related quality of life (the EuroQol 5-domain [EQ-5D] questionnaire in 3 levels), walking difficulties, the 30-second chair stand test, and willingness to consider surgery were collected at baseline and at 6 weeks.

Results. At follow-up, 31% of those participants willing to consider surgery at baseline no longer considered surgery. Of those participants who were unwilling to consider surgery at baseline, 6% reconsidered and decided in favor of surgery at follow-up. Less pain and a higher EQ-5D score at 6 weeks were associated with the change from being willing to unwilling to consider surgery at follow-up (odds ratio [OR] 0.67–1.64; $P < 0.05$). Worse pain, a lower EQ-5D score, and having walking difficulties at 6 weeks, and a lower pain and EQ-5D score improvement were associated with the change from being unwilling to willing to consider surgery at 6 weeks (OR 0.51–4.30; $P < 0.005$).

Conclusion. Evidence-based nonsurgical OA treatment, at least delivered in a digital format, may reduce the need for surgery and should therefore be offered as the first-line treatment option to patients with hip and knee OA. The results also support the idea that such treatment programs have the potential to improve selection of patients for total joint replacement.

INTRODUCTION

Total joint replacement (TJR) of the knee and/or hip is a common treatment for end-stage osteoarthritis (OA) and >1.2 million hip and knee TJRs are performed annually in the US alone, with an estimated total financial burden of 20 billion dollars (1). While the number of TJRs is expected to gradually rise with the increasing aging population (2), some studies propose that this procedure may not be effective for all patients (3), and in some cases TJR will even increase hospitalization and health care costs (4). Previous studies have shown that between 25% and 34% of all hip and knee TJRs may be inappropriate (5,6), and nearly one-fifth of patients undergoing TJR are not satisfied with the outcome (7).

According to international guidelines, first-line treatment in hip and knee OA should be based on education and exercise, as well as weight loss if needed (8). To implement those guidelines, different self-management programs, including education and either optional (9) or compulsory (10,11) exercises aiming at improved strength and neuromuscular control, have been developed in Sweden (Better Management of Patients with OsteoArthritis [BOA]) in 2008 (9) and in Denmark (Good Life with Osteoarthritis in Denmark) (10,11) in 2013, and a similar program, the stepped-care approach, is offered in the Netherlands (12). Findings from these programs show significant improvements in pain level, physical function, and quality of life as well a decrease in medication intake and sick leave in patients with hip and knee OA that may last for up to 2 years after completion of the program (10,13). Most importantly, findings also indicate that education and exercise may delay or reduce the need for hip and knee replacements in these patients (14–17). Despite these

Supported by Vinnova (Ministry of Enterprise and Innovation, National Health Service), and by the Region Skåne.

[1]Anna Cronström, PhD: Lund University, Lund, Sweden; [2]Håkan Nero, PhD: Lund University, Lund, and Joint Academy, Malmö, Sweden; [3]Leif E. Dahlberg, PhD: Lund University and Skåne University Hospital, Lund, Sweden.

Dr. Dahlberg is the cofounder and Chief Medical Officer of Joint Academy by Arthro Therapeutics AB. No other disclosures relevant to this article were reported.

Address correspondence to Anna Cronström, PhD, Lund University, PO Box 117, 221 00 Lund, Sweden. E-mail: anna.cronstrom@med.lu.se.

Submitted for publication May 2, 2018; accepted in revised form September 25, 2018.

SIGNIFICANCE & INNOVATIONS

- Willingness to consider surgery is the most prominent indicator for referral to total joint replacement (TJR) in patients with hip and knee osteoarthritis (OA).
- Willingness for surgery may not necessarily indicate future beneficial outcomes of surgery.
- After 6 weeks in a nonsurgical digital OA treatment program, 31% of those willing to consider surgery at baseline had reconsidered. Of those who were unwilling to consider surgery at baseline, 6% reconsidered and decided in favor of surgery at 6 weeks. The shift in attitude, in either direction, was highly dependent on the success of the treatment program in reducing the OA symptoms.
- A structured and evidence-based nonsurgical OA treatment program may reduce the need for TJR and should be offered as the first-line treatment option to patients with hip and knee OA**.** The patients' willingness for TJR before completing nonsurgical OA treatment may be a poor indicator for surgery.

findings, many patients do not receive adequate information on treatment options, and surgery is often offered before nonsurgical treatments have been adequately used (9,18,19).

Identifying the most appropriate patients as candidates for TJR is not a straightforward process, as the opinions regarding indications for TJR seem to differ among physicians (20,21). In addition, findings from a systematic review highlight the fact that patients' willingness to undergo surgery has been shown to be the most prominent indicator for referral to TJR in individuals with hip and knee OA (22). This fact may be problematic, since the willingness to consider TJR is influenced by factors such as sociodemographic status and expectations of surgery (22) and may not necessarily indicate future beneficial outcomes of surgery. In a study by Hawker et al (23), more severe OA symptoms and impaired walking ability were reported to be associated with the patients' willingness to consider TJR in an elderly population with symptomatic hip or knee OA in Canada. However, whether and how patients' willingness to consider surgery may change after completing structured evidence-based nonsurgical OA treatment is unclear. Such knowledge may further improve the identification of patients eligible for TJR. Thus, the aim of this study was to investigate any possible shift in willingness to consider surgery and to investigate factors associated with this shift, following completion of a digital treatment program for hip and knee OA, including education and exercise as well as asynchronous chat with a physical therapist.

SUBJECTS AND METHODS

Intervention. This study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology guidelines for observational studies (24). The intervention consisted of a digital, nonsurgical OA treatment program (called Joint Academy), detailed in a previous publication (25). Briefly, the first 6 weeks of the program comprise 8 video lectures about OA, physical activity, and self-management in OA as well as different levels of exercises aimed at improving strength and neuromuscular control, based on each individual's progression in the program. The participants are also able to chat asynchronously with a physical therapist throughout the duration of the program. Joint Academy is a digital version of the Swedish evidence-based face-to-face BOA self-management treatment program (9), and Joint Academy has previously been found to reduce pain and improve function and quality of life in patients with hip and knee OA (16,25).

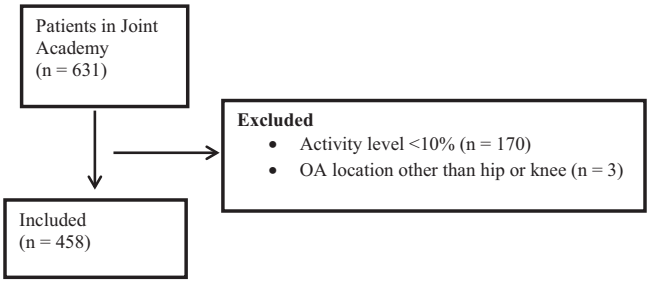

Figure 1. Flow chart of the inclusion process. OA = osteoarthritis.

Table 1. Baseline characteristics of included participants (n = 458)*

Characteristic	Values
Women, %	67.8
Age, years	62 ± 5.6
Body mass index, kg/m^2	26.9 ± 4.9
Working situation, %	
Working	46.6
Retired	45.5
Unemployed	3.3
Sick-leave	4.2
OA medication last 6 months, %	49.9
Previous surgery to other joint, %	13.3
Pain location knee, %	58.2
Pain baseline	5.6 ± 2.2
Walking difficulties at baseline, %	84.5
30CST baseline median (quartiles)†	10 (8–12)
EQ-5D baseline score	0.64 ± 0.2
Fear of physical activity at baseline, %	23.6
Consider surgery at baseline, %	23.2
Compliance level in percentage	78.2 ± 17.3

* Values are the mean ± SD unless indicated otherwise. OA = osteoarthritis; 30CST = 30-second chair stand test; EQ-5D = EuroQol 5-domain questionnaire.
† Non-normally distributed data.

Participants. Of 631 participants who completed the Joint Academy program (16,25), register data from 458 patients (mean ± SD age 62 ± 5.6 years, 67% women) between November 2015 and January 2018 were used in this study. Inclusion criteria were hip or knee OA diagnosed by an orthopedic surgeon and/or a physiotherapist involved in the Joint Academy program; completion of the patients' first 6-weeks in the treatment program for OA; and reporting at least 1 of the following factors at baseline and at 6 weeks: pain, health-related quality of life, and physical function. Exclusion criteria were reporting another joint than hip or knee as the primary OA location and a level of program compliance of <10%. This level of compliance has been used in previous studies on the effect of Joint Academy (16,25) and was defined as the proportion of completed videos, exercises, and questionnaires offered in the program. A flow chart of the inclusion process is shown in Figure 1. Participant characteristics are shown in Table 1. The study was approved by the Regional Ethics Review Board in Lund, Sweden (Dnr 2017/651; Dnr 2017/980), and all patients gave their informed consent at registration.

Data collection. The following demographic data were collected at baseline registration: age, sex, body mass index (BMI), employment status, primary OA location (hip or knee), previous surgery on any joint, and intake of OA medications during the last 6 months. Prior to starting the program (at baseline) and at follow-up (at 6 weeks), the participants were asked whether they had any walking difficulties, whether they had any fear of physical activity, and whether they were willing to consider surgery due to OA-related symptoms (yes/no). In addition, they were asked to fill out a questionnaire on health-related quality of life, the EuroQol 5-domain (EQ-5D) questionnaire in 3 levels. The EQ-5D includes questions about mobility, self-

Table 2. Differences in baseline demographics, pain, and function between those participants who considered surgery at baseline and those who did not (n = 458)*

Factor	Willing	Unwilling	*P*
Sex			
Women	60 (19.4)	250 (80.6)	0.005†
Men	46 (31.1)	102 (68.9)	–
Age, mean ± SD years	61 ± 9.5	62 ± 9.6	0.154‡
Body mass index, mean ± SD kg/m^2	28 ± 5.7	26.5 ± 4.7	0.007§
Working situation (n = 423)			
Working	48 (22.4)	166 (77.6)	0.642¶
Retired	43 (20.6)	166 (79.4)	–
OA medication last 6 months			
Yes	65 (28.5)	173 (71.5)	0.007†
No	41 (17.8)	189 (82.2)	–
Previous surgery			
Yes	22 (36.1)	39 (63.9)	0.005†
No	84 (21.2)	313 (78.9)	–
Pain location			
Hip	43 (22.5)	148 (77.5)	0.787¶
Knee	63 (23.6)	204 (76.4)	–
Pain at baseline, mean ± SD	6.8 ± 1.8	5.2 ± 2.1	<0.001§
EQ-5D score, mean ± SD	0.54 ± 0.19	0.68 ± 0.12	<0.001§
Baseline walking difficulties			
Yes	102 (26.4)	205 (73.6)	<0.001†
No	4 (5.6)	67 (94.4)	–
Baseline 30CST, median (quartiles)#	10.0 (8–11)	10.0 (8–12)	0.02**
Fear of physical activity at baseline	32 (29.4)	77 (70.6)	0.078¶

* Values are the number (%) unless indicated otherwise. OA = osteoarthritis; EQ-5D = EuroQol 5-domain questionnaire; 30CST = 30-second chair stand test.
† Statistically significant by the chi-square test.
‡ Independent *t*-test.
§ Statistically significant by independent *t*-test.
¶ Chi-square test.
# Non-normally distributed data (n = 447).
** Statistically significant by Wilcoxon signed rank test for non-normally distributed data.

care, usual activities, pain/discomfort, and anxiety/depression. A higher EQ-5D score indicates better health-related quality of life (26,27). Participants were also asked to rate their current pain on an 11-point numerical rating scale (NRS; where 0 = no pain and 10 = the worst possible pain) (28). Physical function was assessed using the 30-second chair stand test (30CST) (29), in which the number of repetitions of sitting to standing from a chair during a period of 30 seconds was recorded (self-reported).

Statistical analysis. All statistics were calculated using SPSS software, version 24. Data were explored for normality using visual inspection of histograms and interpretation of skewedness and kurtosis. All data met the assumptions of normality except physical function. To assess the proportion of cross overs from considering surgery at baseline to not considering surgery after completion of the program, and the reverse, cross tabulation and the chi-square test were used. Cross tabulation and the chi-square test were also used when the data were dichotomous, and Student's *t*-test (normally distributed data) and Wilcoxon's signed rank test (non-normally distributed data) were used for continuous data, to assess differences in demographics, pain, and function between those participants who were willing to consider surgery and those participants who were not. At 6 weeks, the cohort was divided into 2 groups, 1 group that had been willing to consider surgery at baseline (n = 104) and 1 group that had been unwilling to consider surgery at baseline (n = 348). Due to multicollinearity between pain, physical function, and walking ability, separate logistic regressions adjusted for age, sex, BMI, and previous surgery were performed to evaluate associations between each independent variable (pain, EQ-5D score, fear of physical activity, walking difficulties, and 30CST) and the dependent variable of willingness to consider surgery, in the 2 groups at 6 weeks. In the group of patients who were willing to consider surgery at baseline, unwillingness to consider surgery at 6 weeks was given the value 1, and in the group of patients who were unwilling to consider surgery at baseline, willingness to consider surgery at 6 weeks was given the value 1 in the regression analyses. For the purpose of regression, the EQ-5D score (0–1) was multiplied by 10. *P* values less than or equal to 0.05 were considered statistically significant. Due to the exploratory design of the study, no adjustments for multiple comparisons were made (30).

Table 3. Differences in pain and function at 6 weeks from baseline willingness to consider surgery*

	Willing at baseline (n = 104)			Unwilling at baseline (n = 348)		
Factor	Willing at 6 weeks (n = 72)	Reconsidered at 6 weeks (n = 32)	*P*	Unwilling at 6 weeks (n = 327)	Reconsidered at 6 weeks (n = 21)	*P*
Pain at 6 weeks	5.8 ± 1.9	4.1 ± 2.4	<0.001†	3.6 ± 2.2	5.5 ± 2.2	0.001†
Change in pain, baseline to 6 weeks	–1.3 ± 1.9	–2 ± 3.6	0.330‡	–1.5 ± 2.2	–0.5 ± 2.0	0.043†
EQ-5D at 6 weeks	0.55 ± 0.18	0.69 ± 0.18	0.001†	0.72 ± 0.12	0.56 ± 0.16	<0.001†
Change in EQ-5D, baseline to 6 weeks	0.05 ± 0.15	0.06 ± 0.12	0.806‡	0.04 ± 0.13	–0.04 ± 0.13	0.011†
Walking difficulties at 6 weeks, %						
Yes	73.3	26.6	0.031§	90.9	9.1	0.009§
No	46.7	53.3	–	97.9	2.1	–
Fear of physical activity at 6 weeks, %						
Yes	80.0	20	0.341¶	95.7	4.3	0.687¶
No	67.8	32.2	–	93.5	6.5	–
30CST at 6 weeks, median (quartiles)#	11 (9–15)	12 (10–15)	0.289**	12 (10–16)	12 (9–15)	0.18**
Change in 30CST, baseline to 6 weeks	1.9 ± 4.5	2.2 ± 3.5	0.755‡	2.2 ± 4.2	2.74 ± 4.9	0.602‡

* Values are the mean ± SD unless indicated otherwise. EQ-5D = EuroQol 5-domain questionnaire; 30CST = 30-second chair stand test.
† Statistically significant by independent *t*-test.
‡ Independent *t*-test.
§ Statistically significant by the chi-square test.
¶ Chi-square test.
# Non-normally distributed data.
** Wilcoxon signed rank test for non-normally distributed data.

RESULTS

After 6 weeks in the nonsurgical digital OA treatment program, 32 of 104 participants (31%) of those who were willing to consider surgery at baseline no longer considered surgery. Of those who were unwilling to consider surgery at baseline, 21 of the 348 participants (6%) reconsidered and decided in favor of surgery after 6 weeks ($P < 0.001$).

Differences between those participants who were willing to consider surgery at baseline and those who were not. Male participants, participants taking any OA-related medication during the last 6 months, those who had had previous surgery, and participants reporting walking difficulties at baseline were more likely to be willing to consider surgery at baseline ($P < 0.05$). Participants who considered surgery at baseline also had a higher BMI, greater pain, a lower EQ-5D score, and worse physical function at baseline compared to those participants who were unwilling to consider surgery ($P < 0.05$). No differences in age, working situation, pain location, or fear of physical activity were observed (Table 2).

Factors associated with the shift from being willing to consider surgery at baseline to no longer considering surgery after completion of the program. Of participants who said they had considered surgery at baseline, those who reconsidered after completion of the program were less likely to have walking difficulties at 6 weeks and had less pain and a higher EQ-5D score at 6 weeks than those who still considered having surgery after completion of the program ($P < 0.005$) (Table 3). After adjusting for age, sex, BMI, and previous surgery, the only variables associated with the shift from willingness to consider surgery to no longer considering surgery at 6 weeks were less pain at 6 weeks (odds ratio [OR] 0.67) and a higher EQ-5D score (OR 1.64). In other words, for every step-increase in NRS pain, the likelihood of having reconsidered after the program decreased by 33%, and for every 0.1 step increase in the EQ-5D score, the likelihood of having reconsidered increased by 64% (Table 4 and Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23772/abstract).

Factors associated with the shift from being unwilling to consider surgery at baseline to willingness to consider surgery after completion of the program. Of participants who said they were unwilling to consider surgery at baseline, those who reconsidered were more likely to have walking difficulties, a lower EQ-5D score at 6 weeks, and greater pain at 6 weeks. They had also experienced smaller improvements in pain and the EQ-5D score compared to those who still did not consider surgery (Table 3). The adjusted regression models showed that worse pain at 6 weeks (OR 1.63), a lower EQ-5D score at 6 weeks (OR 0.51), less pain improvement (OR 1.30), a smaller EQ-5D score improvement (OR 0.63), and having walking difficulties at 6 weeks (OR 4.30) were independently associated with the shift from being unwilling at baseline to being willing to consider surgery at 6 weeks (Table 4 and Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23772/abstract).

DISCUSSION

In this first study to our knowledge investigating factors associated with the shift in willingness to consider surgery after participation in a digital nonsurgical treatment program for OA, nearly one-third of the participants changed their attitude and no longer considered surgery after completion of the program. Less pain and a better health-related quality of life after completion of the program were independently associated with the participants' shift from being willing to unwilling to consider surgery at 6 weeks. Worse pain, health-related quality of life, and walking ability, and less improvement in pain and health-related quality of life after completing the program were independently associated with the participants' shift from being unwilling to willing to consider surgery at 6 weeks.

Table 4. Factors associated with the shift in willingness to consider surgery after completion of the program, adjusted for age, sex, and body mass index*

Independent variable	Adjusted OR (95% CI)	*P*†
Shift from willing to unwilling		
Pain at 6 weeks (n = 104)	0.67 (0.53–0.85)	<0.001
EQ-5D score at 6 weeks (n = 103)	1.64 (1.17–2.30)	0.004
Shift from unwilling to willing		
Pain at 6 weeks (n = 347)	1.63 (1.27–2.08)	<0.001
Pain change (n = 347)	1.30 (1.07–1.48)	0.009
Walking difficulties at 6 weeks (n = 347)	4.30 (1.24–14.94)	0.022
EQ-5D score at 6 weeks (n = 343)	0.51 (0.39–0.67)	<0.001
EQ-5D score change (n = 343)	0.63 (0.44–0.88)	0.007

* The dependent variable is willingness to consider surgery after completion of the program (at 6 weeks). Number of participants who shifted from willing to unwilling: n = 32; number of participants who shifted from unwilling to willing: n = 21. OR = odds ratio; 95% CI = 95% confidence interval. For the purpose of regression the EuroQol 5-domain questionnaire (EQ-5D) score was multiplied by 10.
† All *P* values are statistically significant.

Consistent with previous research, worse OA symptoms were associated with a willingness to consider surgery before entering the treatment program (23), whereas age and work situation seem to have little importance. We also found men to be more willing to consider surgery at baseline than women. This finding is in contrast, however, with a recent review that showed men and women to be equally willing to consider surgery due to OA-related symptoms, but that women were less likely than men to be referred to surgery, despite OA severity (22). Post hoc analyses revealed no differences in OA symptoms between men and women, but a higher proportion of the men (22% versus 9%) in this cohort had undergone a previous TJR in another joint. Given that previous surgery was associated with willingness to have surgery, experiences and expectations of surgery may be one explanation for the fact that a higher proportion of the men compared to women considered surgery in this study.

As shown in previous research (23) and the baseline data in the current study, the severity of OA symptoms may influence willingness to consider surgery. However, whether completing a nonsurgical treatment program aimed at reducing OA symptoms may alter the attitude toward surgery in either direction has not been previously clarified. In the current study, 31% of those participants who considered surgery as a treatment option before entering the online OA treatment program reconsidered and no longer considered surgery as an option after completion of the program. This result is in line with previous studies that showed a reduction in surgery interest of between 24% and 67% after participants were enrolled in structured nonsurgical treatment programs including education and exercise (14,16,17). Furthermore, in another study, only 26% of patients eligible for TJR actually underwent surgery after being enrolled in a nonsurgical treatment option (15). This is the first study to investigate whether completing a nonsurgical treatment program may be associated with crossing over from being unwilling to consider surgery at baseline to being willing to consider surgery after the program. In the current study, approximately 6% of the participants shifted in this direction. The adjusted result from this study indicates that patients who experienced reduced pain and better health-related quality of life after completing the program more often changed their mind and no longer considered surgery. On the other hand, some of the patients who did not consider surgery at baseline and then experienced small improvements in pain level and health-related quality of life, and who still had walking difficulties after the program, also reconsidered and changed their preference in favor for TJR. For example, the improvements in pain in the group who were willing to consider surgery at baseline but reconsidered after completion of the program correspond to a clinically significant change (–2 points on an NRS) (31), whereas the participants who still considered surgery or were unwilling at baseline but reconsidered after completion of the program did not reach clinically significant changes. That is, the individual patient's willingness to consider surgery after the program is highly dependent on the success of the treatment program in reducing their OA symptoms.

In TJR, identifying the patients for whom surgery will be beneficial is a crucial matter. Today, approximately 20% of the patients who undergo TJR for hip or knee OA are not satisfied with the result, which, to some extent, may be attributed to presurgery expectations (7). Studies also showed that the willingness to consider surgery is highly dependent on factors not related to OA symptoms, such as social network, socioeconomic status, and expectations of surgery (32). The result from this study indicates that a significant number of patients will change their attitude toward surgery, in either direction, after completing a treatment program including education and exercise. Thus, offering nonsurgical treatment to patients with hip and knee OA before they make any decision regarding TJR is essential. In this study, approximately one-third of the participants no longer considered surgery after the program. This number also corresponds to the proportion of performed hip and knee TJRs that may be deemed inappropriate each year (6). In other words, in the US alone, unnecessary surgery costing approximately $8.3 billion is performed annually (1). Furthermore, some patients (6%) changed their attitude in the opposite direction. Given this fact, a structured nonsurgical treatment program, when delivered in a digital format online, may reduce the need for TJR and the financial burden of inappropriate surgeries, and in addition assist in selecting those for whom surgery will be beneficial and therefore may also increase the postsurgery satisfaction rate.

Some limitations of this study should be recognized. First, similar to previous studies on the effect of the Joint Academy program (16,25), to increase study power, the lowest level of compliance with the program to be eligible for this study was set at 10%. This setting is a relatively low compliance level, and thus the results in this study may be underestimated, compared to what might have been the case if a higher compliance level in the program had been used. However, since the mean level of compliance in the program was 78%, the compliance level did not likely have an effect on the result. Second, we combined patients with hip and knee OA into 1 group in the analyses. Patients with hip and knee OA are suggested to constitute 2 populations with different expectations of surgery and different surgical outcomes (33), and separate analyses may thus be warranted. However, post hoc analyses revealed no difference in baseline pain and function or willingness to consider surgery between those participants with hip and knee OA (33). Thus the location of OA, i.e., the hip or knee joint, did not likely affect the results in this study.

Third, due to the choice of an observational study design, we do not know whether the results of the digital management program are generalizable to patients receiving no treatment or those undergoing face-to-face programs, such as BOA (9) or Good Life with Osteoarthritis in Denmark (10). Future results from ongoing studies may give further insight into these questions (34). Nevertheless, nonsurgical OA management programs including

education and exercise are evidence-based, and data indicate that this digital program encouraging patients to carry out daily treatment may be at least as effective in reducing OA symptoms as face-to-face treatment (16), and also add long-term effects (35). Therefore, the present results are likely to apply for all types of nonsurgical OA treatments that include education and exercise, regardless of how they are delivered. Fourth, willingness to consider TJR was only evaluated after 6 weeks in the program, and future studies on the long-term willingness for TJR after participation in such OA treatment program are thus warranted. However, Skou et al (15) showed that 75% of patients appointed for knee TJR reconsidered after completion of a nonsurgical treatment program. At follow-up 1 year later, those patients still did not find a knee replacement necessary, indicating long-term effects.

Finally, individual decision-making on important health care aspects such as TJR is complex and cannot solely be explained by the factors investigated in this study. Qualitative studies have highlighted factors such as ability to cope with pain, expectations of surgery, the patient-doctor relationship, and personal views on eligibility criteria for TJR to be important factors when experiencing hip and knee OA and considering TJR (22,33). None of these factors were evaluated in this study. Furthermore, the decision-making process can be divided into 2 stages, the deliberation stage, when the patients consider their options, gather information and review the advantages and disadvantages of these options, and the decision-making stage, where the actual decision is determined (36). In a recent review, Barlow et al (33) discuss the fact that future research is needed to investigate the likelihood of patients to go back to the deliberation stage if their OA symptoms decrease. The result from this study provides evidence that points in that direction. However, studies that include satisfaction after TJR as well as qualitative studies that include patients who have already completed a structured nonsurgical treatment program are warranted, to improve our understanding of the individual factors that are involved in TJR decision-making after nonsurgical treatment, and to further improve the identification of patients who should be referred to TJR.

Structured nonsurgical OA treatment, when delivered in a digital format online, may reduce the number of patients interested in having surgery and can possibly delay or reduce the need for surgical joint replacement. The result showing that one-third of the patients who were willing to consider surgery before entering the online OA treatment program reconsidered after completion supports the idea that exercise and education should be offered as the first-line treatment option for patients with hip and knee OA. A patient's willingness to have TJR before nonsurgical OA treatment may therefore be a poor indicator for surgery. Less improvement in pain, walking ability, and health-related quality of life after completion of the program may cause the patients to change their attitude in favor of surgery. Taken together, these results show that a patient's attitude for and against surgery may shift after program completion. This result suggests that participation in a structured evidence-based nonsurgical OA treatment program has the potential to improve selection of patients for TJR.

ACKNOWLEDGMENT

The authors thank the study participants for contributing essential 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. Cronström 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. Cronström, Nero, Dahlberg.
Acquisition of data. Cronström, Nero, Dahlberg.
Analysis and interpretation of data. Cronström, Nero, Dahlberg.

REFERENCES

1. Lam V, Teutsch S, Fielding J. Hip and knee replacements: a neglected potential savings opportunity. JAMA 2018;319:977–8.
2. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am 2007;89:780–5.
3. Dowsey MM, Gunn J, Choong PF. Selecting those to refer for joint replacement: who will likely benefit and who will not? Best Pract Res Clin Rheumatol 2014;28:157–71.
4. Bozic KJ, Stacey B, Berger A, Sadosky A, Oster G. Resource utilization and costs before and after total joint arthroplasty. BMC Health Serv Res 2012;12:73.
5. Cobos R, Latorre A, Aizpuru F, Guenaga JI, Sarasqueta C, Escobar A, et al. Variability of indication criteria in knee and hip replacement: an observational study. BMC Musculoskelet Disord 2010;11:249.
6. Riddle DL, Jiranek WA, Hayes CW. Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: a multicenter longitudinal cohort study. Arthritis Rheumatol 2014;66:2134–43.
7. Dunbar MJ, Richardson G, Robertsson O. I can't get no satisfaction after my total knee replacement: rhymes and reasons. Bone Joint J 2013;95 Suppl A:148–52.
8. Jevsevar DS. Treatment of osteoarthritis of the knee: evidence-based guideline. J Am Acad Orthop Surg 2013;21:571–6.
9. Thorstensson CA, Garellick G, Rystedt H, Dahlberg LE. Better management of patients with osteoarthritis: development and nationwide implementation of an evidence-based supported osteoarthritis self-management programme. Musculoskeletal Care 2015;13:67–75.
10. Skou ST, Roos EM. Good Life with osteoArthritis in Denmark (GLA:D): evidence-based education and supervised neuromuscular exercise delivered by certified physiotherapists nationwide. BMC Musculoskelet Disord 2017;18:72.
11. Davis AM, Kennedy D, Wong R, Robarts S, Skou ST, McGlasson R, et al. Cross-cultural adaptation and implementation of Good Life with osteoArthritis in Denmark (GLA:D): group education and exercise for hip and knee osteoarthritis is feasible in Canada. Osteoarthritis Cartilage 2018;26:211–9.
12. Smink AJ, Dekker J, Vliet Vlieland TP, Swierstra BA, Kortland JH, Bijlsma JW, et al. Health care use of patients with osteoarthritis of

the hip or knee after implementation of a stepped-care strategy: an observational study. Arthritis Care Res (Hoboken) 2014;66: 817–27.

13. Kroon FP, van der Burg LR, Buchbinder R, Osborne RH, Johnston RV, Pitt V. Self-management education programmes for osteoarthritis. Cochrane Database Syst Rev 2014;1:CD008963.
14. Svege I, Nordsletten L, Fernandes L, Risberg MA. Exercise therapy may postpone total hip replacement surgery in patients with hip osteoarthritis: a long-term follow-up of a randomised trial. Ann Rheum Dis 2015;74:164–9.
15. Skou ST, Roos EM, Laursen MB, Rathleff MS, Arendt-Nielsen L, Simonsen O, et al. A randomized, controlled trial of total knee replacement. N Engl J Med 2015;373:1597–606.
16. Nero H, Dahlberg J, Dahlberg LE. A 6-week web-based osteoarthritis treatment program: observational quasi-experimental study. J Med Internet Res 2017;19:e422.
17. Smittenaar P, Erhart-Hledik JC, Kinsella R, Hunter S, Mecklenburg G, Perez D. Translating comprehensive conservative care for chronic knee pain into a digital care pathway: 12-week and 6-month outcomes for the Hinge Health program. JMIR Rehabil Assist Technol 2017;4:e4.
18. Li LC, Cott C, Jones CA, Badley EM, Davis AM, PEOPLE meeting participants. Improving primary health care in chronic musculoskeletal conditions through digital media: the PEOPLE meeting. JMIR Res Protoc 2013;2:e13.
19. Li LC, Sayre EC, Kopec JA, Esdaile JM, Bar S, Cibere J. Quality of nonpharmacological care in the community for people with knee and hip osteoarthritis. J Rheumatol 2011;38:2230–7.
20. Wright JG, Hawker GA, Hudak PL, Croxford R, Glazier RH, Mahomed NN, et al, for the Toronto Arthroplasty Research Group Writing Committee. Variability in physician opinions about the indications for knee arthroplasty. J Arthroplasty 2011;26:569–75.
21. Coyte PC, Hawker G, Croxford R, Attard C, Wright JG. Variation in rheumatologists' and family physicians' perceptions of the indications for and outcomes of knee replacement surgery. J Rheumatol 1996;23:730–8.
22. Mota RE, Tarricone R, Ciani O, Bridges JF, Drummond M. Determinants of demand for total hip and knee arthroplasty: a systematic literature review. BMC Health Serv Res 2012;12:225.
23. Hawker GA, Wright JG, Badley EM, Coyte PC, for the Toronto Arthroplasty Health Services Research Consortium. Perceptions of, and willingness to consider, total joint arthroplasty in a population-based cohort of individuals with disabling hip and knee arthritis. Arthritis Rheum 2004;51:635–41.
24. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, for the STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008;61:344–9.
25. Dahlberg LE, Grahn D, Dahlberg JE, Thorstensson CA. A web-based platform for patients with osteoarthritis of the hip and knee: a pilot study. JMIR Res Protoc 2016;5:e115.
26. EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.
27. Payakachat N, Ali MM, Tilford JM. Can the EQ-5D detect meaningful change? A systematic review. Pharmacoeconomics 2015;33:1137–54.
28. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs 2005;14:798–804.
29. Dobson F, Hinman RS, Hall M, Marshall CJ, Sayer T, Anderson C, et al. Reliability and measurement error of the Osteoarthritis Research Society International (OARSI) recommended performance-based tests of physical function in people with hip and knee osteoarthritis. Osteoarthritis Cartilage 2017;25:1792–6.
30. Bender R, Lange S. Adjusting for multiple testing: when and how? J Clin Epidemiol 2001;54:343–9.
31. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004;8:283–91.
32. Toye FM, Barlow J, Wright C, Lamb SE. Personal meanings in the construction of need for total knee replacement surgery. Soc Sci Med 2006;63:43–53.
33. Barlow T, Griffin D, Barlow D, Realpe A. Patients' decision making in total knee arthroplasty: a systematic review of qualitative research. Bone Joint Res 2015;4:163–9.
34. National Library of Medicine. ClinicalTrials.gov. Evaluation of a web-pbased platform for osteoarthritis treatment. ClinicalTrial.gov identifier: NCT03328741; 2017. URL: https://www.clinicaltrials.gov/ct2/show/results/NCT03328741?view=results.
35. Nero H, Dahlberg J, Dahlberg LE. Joint academy: a six-week online treatment program for osteoarthritis [abstract]. Osteoarthritis Cartilage 2018;26 Suppl 1:S315.
36. Elwyn G, Miron-Shatz T. Deliberation before determination: the definition and evaluation of good decision making. Health Expect 2010;13:139–47.

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1202–1208
DOI 10.1002/acr.23774

American College of Rheumatology
Empowering Rheumatology Professionals

Relationship Between Hip Morphology and Hip-Related Patient-Reported Outcomes in Young and Middle-Aged Individuals: A Population-Based Study

Jacek A. Kopec,[1] Hong Qian,[2] Jolanda Cibere,[1] Hubert Wong,[3] Linda C. Li,[1] Morgan Barber,[4] Helen M. Prlic,[4] Charlie Zhang,[5] Charles Ratzlaff,[6] Bruce B. Forster,[3] and John M. Esdaile,[1] on behalf of the IMPAKT-HIP Study Team

Objective. Radiographic measurements of the alpha angle and the lateral center edge (LCE) angle in the hip joint are important for the diagnosis of femoroacetabular syndrome, a potential risk factor for hip osteoarthritis. Our objective was to determine whether these measurements are associated with hip-related patient-reported outcomes in young and middle-aged individuals.

Methods. A stratified random sample of white men and women ages 20–49 years, with and without hip pain, was selected using random digit dialing from the population of metro Vancouver, Canada. The alpha and LCE angles were measured bilaterally on radiographs using Dunn and anteroposterior views, respectively. Patient-reported outcomes were measured by the Copenhagen Hip And Groin Outcome Score (HAGOS), which has scales for symptoms, pain, daily activities, sports, physical activity, and quality of life (QoL). We performed descriptive analyses and a regression analysis with restricted cubic splines, adjusted for age and sex and weighted for the sampling design.

Results. Data were obtained for 500 subjects. The alpha angle distribution was strongly skewed, with a mean of 54°. The LCE angle distribution was symmetric, with a mean of 34°. In the restricted cubic splines analysis, the relationship between the alpha angle and HAGOS scores was nonlinear, with higher alpha angles generally associated with worse HAGOS scores for alpha >60°. The associations were statistically significant for symptoms, sports, and QoL. No association was found between the LCE angle and HAGOS scales.

Conclusion. In a general population sample ages 20–49 years, we have found an association between the alpha angle and hip-related patient-reported outcomes.

INTRODUCTION

Femoroacetabular impingement has been proposed as an important risk factor for hip pain and hip osteoarthritis (OA) (1,2). A recent consensus statement defined femoroacetabular impingement syndrome (FAIS) as a combination of symptoms in the hip joint associated with activity, physical signs (mainly range of motion limitations due to pain), and radiographic evidence of cam or pincer morphology (CPM) (3). Cam morphology is a bony prominence at the femoral head-neck junction, whereas pincer morphology is an excessive coverage of the femoral head by the acetabulum (1–3). Common radiographic measurements used to determine CPM are the alpha angle for cam and the lateral center edge (LCE) angle and crossover sign for pincer (4).

Over the past decade, epidemiologic studies have shown a correlation between advanced hip OA and cam morphology (5–9). Nonsurgical treatment and surgical correction of CPM are increasingly offered to patients with FAIS (10). However, the concept of FAIS is relatively new, and important questions surrounding the epidemiology of this condition remain to be elucidated (3). A recent systematic review of 30 studies showed that the current data are insufficient to estimate the population prevalence of cam morphology or to determine its relationship with pain (11).

Supported by the Canadian Institutes of Health Research (grant PAF-107513).

[1]Jacek A. Kopec, PhD, Jolanda Cibere, MD, PhD, Linda C. Li, PhD, John M. Esdaile, MD, MPH: University of British Columbia, Vancouver, and Arthritis Research Canada, Richmond, British Columbia, Canada; [2]Hong Qian, MSc: St. Paul's Hospital, Vancouver, British Columbia, Canada; [3]Hubert Wong, PhD, Bruce B. Forster, MD: University of British Columbia, Vancouver, British Columbia, Canada; [4]Morgan Barber, MHA, Helen M. Prlic, MA: Arthritis Research Canada, Richmond, British Columbia, Canada; [5]Charlie Zhang, MD: University of Calgary, Calgary, Alberta, Canada; [6]Charles Ratzlaff, PhD: University of Arizona, Tucson.

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

Address correspondence to Jacek A. Kopec, PhD, School of Population and Public Health, University of British Columbia, Senior Scientist, Arthritis Research Canada, 5591 No. 3 Road, Richmond, British Columbia, V6X 2C7 Canada. E-mail: jkopec@arthritisresearch.ca.

Submitted for publication April 17, 2018; accepted in revised form October 2, 2018.

SIGNIFICANCE & INNOVATIONS

- Hip morphology is considered a risk factor for the development of hip pain and osteoarthritis, but reports on the association between hip morphology and hip-related patient-reported outcomes have been inconsistent.
- This is the first study to our knowledge to demonstrate a significant effect of cam morphology (and no effect of pincer morphology) on patient-reported outcomes, such as hip symptoms, limitations in sports activities, and quality of life, in young and middle-aged individuals in the general population.
- The study also shows that this relationship is nonlinear and limited to an alpha angle above 60°, which supports previous recommendations to use this cutoff for the diagnosis of femoroacetabular impingement syndrome.

Studies of the association between CPM and pain have produced inconsistent results. While several authors reported no significant differences in hip morphology between those individuals with and without hip pain (12–16), some studies showed cam morphology to be predictive of pain and other symptoms in athletes and other selected groups (17–19). In a recent population-based study, CPM (defined as alpha >55°, LCE >40°, or crossover sign) was found in 49% of individuals with hip symptoms and 44% of asymptomatic controls, but the difference was nonsignificant in a multivariate analysis (20).

The lack of a significant association between CPM and symptoms in several published studies may be due to a number of reasons. First, hip pain in individuals with CPM may depend on the level of physical activity (20). Second, the relationship between CPM and pain may be nonlinear (5), so that showing an association would be more difficult. Third, the association may be limited to certain forms of CPM (e.g., cam morphology) or severity levels (e.g., a high alpha angle) (18). Finally, in some of the previous studies, the sample size may have been too small and/or the measurements of CPM (e.g., use of anteroposterior view) and pain (e.g., yes/no) too imprecise and insufficiently sensitive to detect an association. The purpose of the current study was to determine whether radiographic measurements of alpha angle and LCE angle, treated as continuous variables, are associated with hip-related patient-reported outcomes in young and middle-aged individuals in the general population.

SUBJECTS AND METHODS

Study design and subjects. A random sample of white men and women ages 20–49 years was selected using random digit dialing from the population of metro Vancouver, Canada. The sample was stratified according to hip pain, as assessed in a telephone interview. Pregnant women and individuals with bilateral hip replacement were excluded. Subjects who agreed to participate obtained radiographs of both hips and filled out a self-administered questionnaire. All subjects provided informed consent and the study was approved by the University of British Columbia Clinical Research Ethics Board.

Radiographic measurements. Standardized radiographs of the pelvis (anteroposterior) and Dunn views of both hips were obtained, as described in detail in a previous study (21). For the anteroposterior pelvis view, the subject was in a weight-bearing position, with legs internally rotated 15°. For the bilateral Dunn view, the subject was supine and the hip was positioned in 45° flexion and 20° abduction while maintaining neutral rotation. The alpha angle was defined in the Dunn view as the angle formed by the axis of the femoral neck and a line connecting the center of the femoral head to the point where the contour begins to stray from a spherical radius (4). The LCE angle was defined in the anteroposterior view as the angle between a line through the center of the femoral head, perpendicular to the transverse axis, and a line through the center of the femoral head, passing through the most superolateral point of the sclerotic weight-bearing zone of the acetabulum (4). All radiographic measurements were performed by a single, trained reader. In a reliability study in 49 subjects with the same reader, the intrarater intraclass correlation coefficient (ICC) was 0.97 for the alpha angle and 0.87 for the LCE angle (21).

Assessment of patient-reported outcomes. To measure patient-reported outcomes we used the Copenhagen Hip And Groin Outcome Score (HAGOS) (22). HAGOS is a validated, multidimensional 37-item instrument developed specifically for use in young to middle-aged patients with chronic hip and/or groin pain and recommended by Griffin et al (3) for assessing outcomes in FAIS. It consists of 6 subscales: symptoms (7 items), pain (10 items), physical function in daily living (activities of daily living [ADL], 5 items), physical function in sport and recreation (sports, 8 items), participation in physical activities (2 items), and hip and/or groin-related quality of life (QoL, 5 items). Each scale is scored on a scale of 0–100, with a higher score indicating better health. Internal consistency and reliability in the validation study were high, with Cronbach's alpha ranging from 0.79 (symptoms) to 0.91 (pain) and test–retest ICC from 0.82 (physical activities) to 0.91 (ADL) (22). Construct validity and responsiveness of the HAGOS scales were also assessed and found adequate (22,23).

Statistical analysis. We calculated weighted percentages for demographic variables and weighted means, medians, and frequency distributions for CPM measurements and HAGOS scores. To assess the relationship between radiographic measurements and HAGOS scores, we used weighted linear regression adjusted for age and sex. In the model, the potential nonlinear relations were identified using restricted cubic splines with 3 knots placed at

Table 1. Descriptive data for the study sample (n = 500)*

Variable	Values	Weighted % (95% CI)
Sex		
Male	181 (36.2)	48.9 (41.6–56.2)
Female	319 (63.8)	51.1 (43.8–58.4)
Age, years		
20–29	50 (10.0)	32.2 (23.8–40.6)
30–39	109 (21.8)	31.4 (25.1–37.7)
40–49	341 (68.2)	36.4 (30.5–42.3)
Education		
High school or less (0–13 grade)	101 (20.2)	27.3 (20.0–34.6)
Vocational or some college	178 (35.6)	27.8 (21.8–33.7)
College or university	221 (44.2)	44.9 (37.7–52.1)
Body mass index, kg/m^2		
<25	241 (48.2)	58.8 (51.9–65.6)
25–29.9	155 (31.0)	26.1 (20.3–31.9)
≥30	104 (20.8)	15.1 (10.6–19.6)
Hip injury		
Yes	26 (5.2)	2.0 (0.8–3.1)
No	474 (94.8)	98.0 (96.9–99.2)

* Values are the number (%) unless indicated otherwise. 95% CI = 95% confidence interval.

the 5th, 50th, and 95th percentiles. Splines are smooth functions that can assume virtually any shape, and the most useful type of spline is generally a cubic spline function, which is restricted to be smooth at the junction of each cubic polynomial (24). *P* values were obtained for tests of overall association and nonlinearity. The analysis was done by selecting for each subject the worst hip, defined as the hip with a largest alpha or LCE angle. All descriptive statistics and analyses were conducted using Proc Survey procedures in SAS software, version 9.4, to account for the sampling design of the study. We performed a sensitivity analysis of the associations between alpha angle and HAGOS scales restricted to individuals reporting any hip pain.

RESULTS

A sample of 858 potential subjects was generated by the random digit dialing screening and we were able to contact 754 (87.9%). Of those, 254 (33.7%) did not provide data: 41 were ineligible, 66 not interested, 84 not available, 53 declined for other/unknown reasons, and 10 were excluded due to incomplete data. Thus, data were obtained for 500 subjects, of whom 64% (unweighted percentages) were female, 68% were ages 40–49 years, 44% had a college education, 21% had a body mass index (BMI) ≥30, and 5% reported a hip injury in the past (Table 1).

The weighted mean alpha angle was 54.7° (95% confidence interval [95% CI] 53.7–55.8, median 53) on the left side and 54.1° (95% CI 53.1–55.0, median 52) on the right side. The mean LCE angle was 34.3° (95% CI 33.7–34.9, median 34) on the left side and 34.6° (95% CI 34.0–35.2, median 34) on the right side (Table 2). Figure 1 shows the weighted distributions of the 2 angles. For the alpha angle, the distribution was strongly skewed and possibly bimodal. The LCE angle distribution was symmetric. Weighted mean scores for HAGOS scales ranged from 80.1 (95% CI 77.2–83.0, median 78.7) for physical activity to 93.6 (95% CI 92.3–94.9, median 96.0) for ADL. HAGOS scores did not differ significantly between groups classified according to alpha angle ≤60° versus >60° and LCE angle ≤40° versus >40° (Table 3).

In the restricted cubic spline analysis, the relationship between the alpha angle and HAGOS scales was nonlinear. The shapes of the curves were similar for all scales (see Supplementary Figure 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23774/abstract). Among those with an alpha angle >60°, the graphs showed HAGOS scores to be worse for higher alpha values. Statistically significant nonlinear associations ($P < 0.05$) were identified for the symptoms, sports, and QoL scales (Figure 2). No association was found between the LCE angle and any of the HAGOS scales (data not shown).

Sensitivity analyses. In an analysis among individuals reporting hip pain, the alpha angle in the worst hip was statistically significantly associated with the symptoms, pain, sports, and QoL scales (see Supplementary Figure 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23774/abstract). A nonlinear association was identified for symptoms only, though the slopes of the

Table 2. Weighted radiographic measurements and HAGOS scores in the study population*

Variable	Mean (95% CI)	Median (95% CI)
Radiographic, degrees		
Alpha angle left hip	54.7 (53.7–55.8)	53 (52–53)
Alpha angle right hip	54.1 (53.1–55.0)	52 (52–53)
Alpha angle worst hip	55.4 (54.4–56.4)	53 (53–54)
LCE angle left hip	34.3 (33.7–34.9)	34 (34–35)
LCE angle right hip	34.6 (34.0–35.2)	34 (33–35)
LCE angle worst hip	35.3 (34.7–35.9)	35 (34–36)
HAGOS scale scores		
Symptoms	86.8 (85.2–88.5)	89.9 (87.5–91.7)
Pain	91.5 (89.9–93.1)	97.7 (96.4–98.0)
ADL	93.6 (92.3–94.9)	96.0 (95.5–96.6)
Sports	91.3 (89.8–92.9)	97.0 (95.6–97.5)
PA	80.1 (77.2–83.0)	78.7 (73.4–85.0)
QoL	88.4 (86.4–90.4)	95.3 (92.5–96.0)

* HAGOS = Copenhagen Hip And Groin Outcome Score; 95% CI = 95% confidence interval; LCE = lateral center edge; ADL = activities of daily living; Sports = physical function in sport and recreation; PA = participation in physical activities; QoL= hip and/or groin-related quality of life.

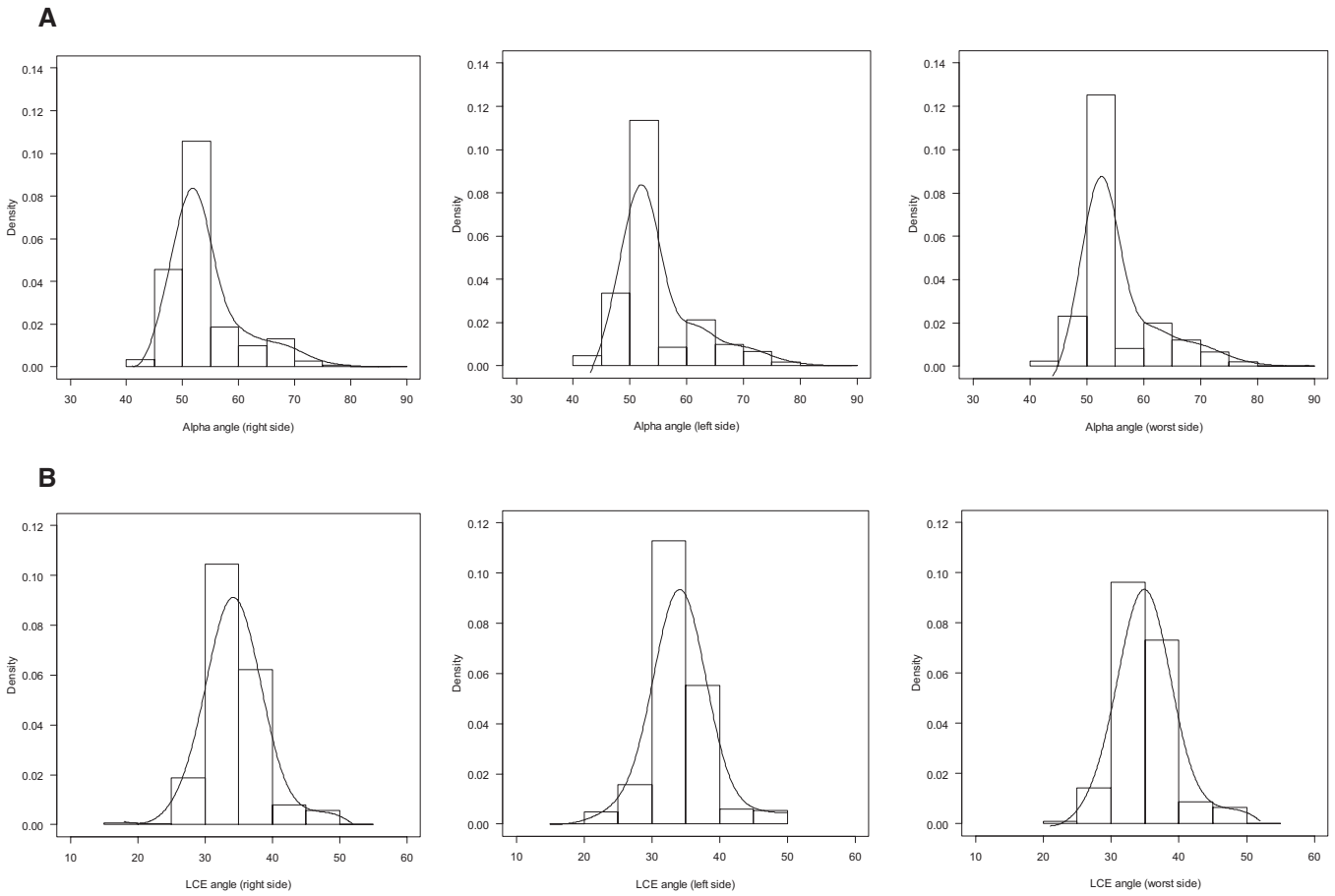

Figure 1. Weighted distributions of **A**, alpha angle and **B**, lateral center edge (LCE) angle in the study population for the left hip, right hip, and worst hip.

curves for alpha >60° appeared steeper than those observed in the weighted pain and nonpain combined analyses.

DISCUSSION

The purpose of this study was to assess the association between radiographic measurements of hip morphology (cam and pincer) and patient-reported hip-related outcomes in individuals ages 20–49 years, selected from the general population. The mean alpha and LCE angles in our study were similar to those reported by Frank et al for asymptomatic hips (25). In a systematic review of cam and pincer morphology prevalence studies, these authors found mean values of 54.1° and 31.2° for alpha and LCE angles, respectively.

In an analysis using restricted cubic splines, we have found a relationship between HAGOS scores and the alpha angle, but not the LCE angle. The general shape of the relationship was similar for all 6 HAGOS scales and consistently

Table 3. Weighted mean HAGOS scores according to alpha angle and LCE angle (worst hip)*

HAGOS scale	Alpha ≤60° (n = 413)	Alpha >60° (n = 87)	LCE ≤40° (n = 457)	LCE >40° (n = 43)
Symptoms	86.5 (84.8–88.3)	88.0 (83.4–92.6)	86.8 (85.1–88.5)	87.1 (79.1–95.1)
Pain	91.6 (89.9–93.2)	91.3 (86.9–95.8)	91.5 (89.9–93.2)	91.4 (83.7–99.0)
ADL	93.5 (92.2–94.8)	94.0 (90.4–97.7)	93.7 (92.5–95.0)	92.2 (84.4–100.0)
Sports	91.3 (89.7–92.9)	91.4 (87.2–95.5)	91.4 (89.8–92.9)	90.4 (82.2–98.6)
PA	80.5 (77.4–83.6)	78.7 (70.9–86.4)	79.8 (76.8–82.9)	83.8 (73.3–94.4)
QoL	88.0 (85.8–90.2)	89.9 (85.1–94.6)	88.2 (86.2–90.3)	89.9 (82.1–97.7)

* Values are the Copenhagen Hip And Groin Outcome Score (HAGOS) (95% confidence interval). LCE = lateral center edge; ADL = activities of daily living; Sports = physical function in sport and recreation; PA = participation in physical activities; QoL = hip and/or groin-related quality of life.

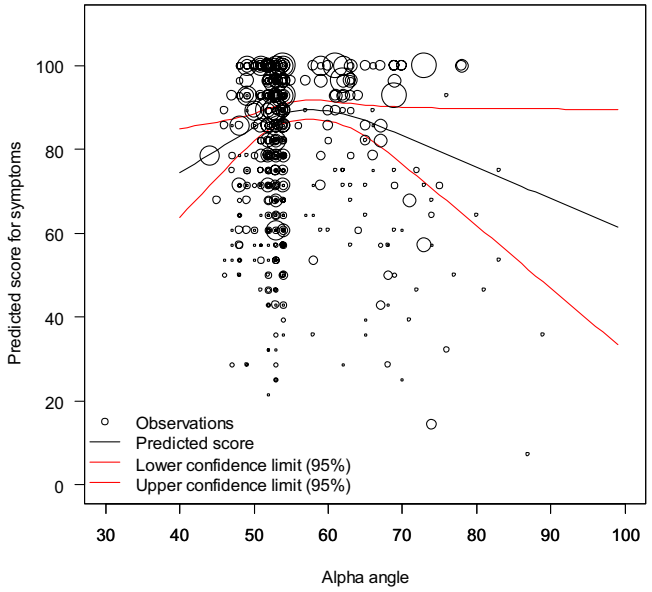

Figure 2. Relationship between the alpha angle and the Copenhagen Hip And Groin Outcome Score symptoms scale. The figure shows the predicted symptoms score and its 95% confidence interval for a man age 42 years (worst hip) from a restricted cubic spline regression analysis, using 3 knots at the 5th, 50th, and 95th percentiles. The size of the observations reflects their corresponding sample weight.

indicated a negative correlation (a higher alpha correlated with a worse score) for alpha >60°. The slope of the curve was significantly different from the null for 3 scales: symptoms, sports, and QoL. This finding may be due to discrepancies in how different outcomes are measured and related to CPM. For example, the pain questions in HAGOS ask about the frequency of hip/groin pain, pain on walking on various surfaces, pain on walking up/down stairs, standing, sitting, and lying, as well as bending and straightening the hip. These normal activities may be less likely to cause pain as a result of hip morphology. Furthermore, the physical activity scale has only 2 items and is the least reliable of the HAGOS scales (22). In contrast, other HAGOS scales ask about difficulty in performing more specific movements, for example, "stretching your legs far out to the side" (symptoms), more demanding activities, such as "running as fast as you can" (sports), and their impact on QoL. Such questions may be more sensitive to the effect of cam morphology, and if so, our results are plausible and consistent with the current concept of FAIS.

The graphic representation of the nonlinear relationship between alpha angle and HAGOS scores can be used to assess the clinical importance of the effect observed. For example, the predicted symptoms score for a man age 42 years with an alpha angle of approximately 80° would be close to 75 (of 100), compared with a score close to 90 for alpha 55–60°. This difference would be considered large and clinically important. On the other hand, as shown in Table 3, the difference in symptoms scores between those with alpha scores >60° versus ≤60° is only 1.5, i.e., very small and clinically insignificant.

Our sensitivity analysis generally confirmed the results observed in the main analysis. The relationship appeared stronger when the analysis was restricted to individuals reporting any hip pain. While this result is not generalizable to the population at large, it is plausible. We would expect the association of alpha angle with hip function and pain to be stronger and easier to detect in this group, compared to a general population sample in which most subjects report no hip pain.

Both in clinical settings and epidemiologic studies, various cutoff values for the alpha angle have been proposed to determine whether cam morphology is present and to diagnose FAIS (3). The cutoff values for the alpha angle in published studies varied from 50° to 83° and none of the published studies was truly population-based (11). Owing to differences in populations, definitions, and methods of assessment, the prevalence of cam morphology has been difficult to determine. In our study, the population (weighted) proportions of individuals ages 20–49 years with cam morphology (worst side) ranged from 87.3% for alpha >50° to 24.6% for alpha >55°, 20.4% for alpha >60°, and 10.4% for alpha >65° (Figure 1). In their analysis of data from the Chingford and the Cohort Hip and Cohort Knee studies, Agricola et al (7) suggested alpha >60° as the best cutoff point to define cam morphology, based on the bimodal distribution they observed. In our data, a greater alpha score was associated with lower patient-reported outcomes scores when the alpha was above 60°, which would support this cutoff.

Our data may also show a negative impact on hip outcomes at lower alpha values (alpha <50°). This possibility should be treated with caution, because our data in this range were sparse, and we are not aware of other studies showing a similar association (although a nonlinear relationship between the alpha angle and radiographic OA risk was reported by Thomas et al [5]).

We have found no association between HAGOS scores and the LCE angle. This result is unlikely to be due to sample size or other methodologic aspects of the study. The LCE distribution in our study was symmetric, with a mean similar to that found in other studies (25). In a previous study, we reported unweighted prevalence of pincer morphology (LCE angle >40°) to be 8% in subjects with pain and 9% in asymptomatic controls (20). To our knowledge, no study has shown a significant association between hip symptoms, function, or OA and isolated pincer morphology or a high LCE angle. On the other hand, a low LCE angle, indicative of hip dysplasia, may be associated with OA (26). It is possible that the LCE angle is not an optimal measure of pincer morphology; however, another common measure, the crossover sign, has been criticized for low specificity in assessing retroversion of the acetabulum (27) and therefore was not used in the current analysis. In a recent review of the criteria for the surgical treatment of FAIS, Peters et al (28) reported that the LCE angle was used in approximately half of the studies.

Our data suggest that the methods for determining pincer morphology and its relationship with hip symptoms and OA require further study.

Several limitations of the study need to be acknowledged. First, the association between alpha angle and HAGOS scores that we found in a cross-sectional observational study does not necessarily imply causation. In theory, the results may be due to confounding by unmeasured risk factors for hip-related outcomes that are correlated with alpha angle. BMI was not related to the alpha or LCE angle, and adjusting for BMI did not change the results. We did not adjust for hip OA because OA can be a mediator of the association under study. Second, the possibility of other biases, and specifically, measurement, selection, or reverse causality bias, also needs to be recognized. Despite our use of valid and reliable measures of the key variables, some degree of error in measuring alpha angle and self-reported outcomes is inevitable. Such errors would be unlikely to be differential, because the subjects were unaware of their radiograph findings and our radiograph readers were unaware of the questionnaire data. Nondifferential errors would dilute the correlations and make them less statistically significant. Selection bias could occur if participation in the study was related to both hip morphology and outcomes. This possibility seems unlikely, because subjects were unaware of their radiograph findings at the time of recruitment. Third, the confidence bands for the restricted cubic splines curves were relatively wide. Thus the lack of statistical significance for some HAGOS scales does not imply that a relationship does not exist. Since the general shape of the relationship was similar across the scales, the weaker associations could become significant in a larger study.

Our study had some methodologic strengths that are worth noting. The study was carried out in a stratified random population sample, and the data were properly weighted to be representative of the general white population of metro Vancouver. As a result, generalizability of the findings is high. We used the Dunn view for assessing the alpha angle; this method is considered more precise than the anteroposterior view employed in most published studies (3). The ICC for alpha angle was 0.97, indicating almost perfect interrater reliability. Rather than using an arbitrary cut point, we analyzed the alpha angle and LCE angle as continuous variables using modern statistical methods (restricted cubic splines). For measuring hip outcomes, we employed the best measure currently available, the HAGOS questionnaire. This instrument has been recommended for research on FAIS (3). Finally, where comparable data were available, our findings were consistent with the literature.

In conclusion, we have for the first time demonstrated the association of cam morphology with poor patient-reported outcomes, such as hip symptoms, limitations in sports activities, and QoL, in a general population sample. We have also shown that this relationship is limited to an alpha angle above 60°, which supports previous recommendations to use this cutoff for the diagnosis of FAIS.

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. Kopec 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. Kopec, Cibere, Li, Ratzlaff, Forster, Esdaile.

Acquisition of data. Cibere, Li, Barber, Prlic, Zhang, Ratzlaff, Forster.

Analysis and interpretation of data. Qian, Wong.

REFERENCES

1. Ganz R, Parvizi J, Beck M, Leunig M, Notzli H, Siebenrock KA. Femoroacetabular impingement: a cause for osteoarthritis of the hip. Clin Orthop Relat Res 2003;417:112–20.
2. Leunig M, Ganz R. Femoroacetabular impingement: a common cause of hip complaints leading to arthrosis. Unfallchirurg 2005;108: 9–17. In German.
3. Griffin DR, Dickenson EJ, O'Donnell J, Agricola R, Awan T, Beck M, et al. The Warwick Agreement on femoroacetabular impingement syndrome (FAI syndrome): an international consensus statement. Br J Sports Med 2016;50:1169–76.
4. Tannast M, Siebenrock KA, Anderson SE. Femoroacetabular impingement: radiographic diagnosis. What the radiologist should know. AJR Am J Roentgenol 2007;188:1540–52.
5. Thomas GE, Palmar AJ, Batra RN, Kiran A, Hart D, Spector T, et al. Subclinical deformities of the hip are significant predictors of radiographic osteoarthritis and joint replacement in women: a 20 year longitudinal cohort study. Osteoarthritis Cartilage 2014;22:1504–10.
6. Nicholls AS, Kiran A, Pollard TC, Hart DJ, Arden CP, Spector T, et al. The association between hip morphology parameters and nineteen-year risk of end-stage osteoarthritis of the hip: a nested case–control study. Arthritis Rheum 2011;63:3392–400.
7. Agricola R, Heijboer MP, Bierma-Zeinstra SM, Verhaar JA, Weinans H, Waarsing JH. Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study (CHECK). Ann Rheum Dis 2013;72:918–23.
8. Doherty M, Courtney P, Doherty S, Jenkins W, Maciewicz RA, Muir K, et al. Nonspherical femoral head shape (pistol grip deformity), neck shaft angle, and risk of hip osteoarthritis: a case–control study. Arthritis Rheum 2008;58:3172–82.
9. Bardakos NV, Villar RN. Predictors of progression of osteoarthritis in femoroacetabular impingement: a radiological study with a minimum of ten years follow-up. J Bone Joint Surg Br 2009;91:162–9.
10. Harris JD, Erickson BJ, Bush-Joseph CA, Nho SJ. Treatment of femoroacetabular impingement: a systematic review. Curr Rev Musculoskelet Med 2013;6:207–18.
11. Dickenson E, Wall PD, Robinson B, Fernandez M, Parsons H, Buchbinder R, et al. Prevalence of cam hip shape morphology: a systematic review. Osteoarthritis Cartilage 2016;24:949–61.
12. Gosvig KK, Jacobsen S, Sonne-Home S, Palm H, Troelsen A. Prevalence of malformations of the hip joint and their relationship to sex, groin pain, and risk of osteoarthritis: a population-based survey. J Bone Joint Surg Am 2010;92:1162–9.
13. Kapron AL, Peters CL, Aoki SK, Beckmann JT, Erickson JA, Anderson MB, et al. The prevalence of radiographic findings of structural hip deformities in female collegiate athletes. Am J Sports Med 2015;43:1324–30.

14. Anderson LA, Anderson MB, Kapron A, Aoki SK, Erickson JA, Chrastil J, et al. The 2015 Frank Stinchfield Award: radiographic abnormalities common in senior athletes with well-functioning hips but not associated with osteoarthritis. Clin Orthop Relat Res 2016;474:342–52.
15. Nardo L, Parimi N, Liu F, Lee S, Jungmann PM, Nevitt MC, et al. Osteoporotic Fractures in Men (MrOS) Research Group. Femoroacetabular impingement: prevalent and often asymptomatic in older men. The Osteoporotic Fractures in Men Study. Clin Orthop Relat Res 2015;473:2578–86.
16. Tsitskaris K, Sharif K, Meacock LM, Bansal M, Ayis S, Li PL, et al. The prevalence of cam-type femoroacetabular morphology in young adults and its effect on functional hip scores. Hip Int 2012;22:68–74.
17. Allen D, Beaulé PE, Ramadan O, Doucette S. Prevalence of associated deformities and hip pain in patients with cam-type femoroacetabular impingement. J Bone Joint Surg Br 2009;91:589–94.
18. Mariconda M, Cozzolino A, Di Pietto F, Ribas M, Bellotti V, Soldati A. Radiographic findings of femoroacetabular impingement in capoeira players. Knee Surg Sports Traumatol Arthrosc 2014;22:874–81.
19. Larson CM, Sikka RS, Sardelli MC, Byrd JW, Kelly BT, Jain RK, et al. Increasing alpha angle is predictive of athletic-related "hip" and "groin" pain in collegiate National Football League prospects. Arthroscopy 2013;29:405–10.
20. Kopec JA, Cibere J, Li LC, Zhang C, Barber M, Qian H, et al. IMPAKT-HIP Study Team. Relationship between physical activity and hip pain in persons with and without cam or pincer morphology: a population-based case-control study. Osteoarthritis Cartilage 2017;25:1055–61.
21. Ratzlaff C, Zhang C, Korzan J, Josey L, Wong H, Cibere J, et al. The validity of a non-radiologist reader in identifying cam and pincer femoroacetabular impingement (FAI) using plain radiography. Rheumatol Int 2016;36:371–6.
22. Thorborg K, Hölmich P, Christensen R, Petersen J, Roos EM. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. Br J Sports Med 2011;45:478–91.
23. Kemp JL, Collins NJ, Roos EM, Crossley KM. Psychometric properties of patient-reported outcome measures for hip arthroscopic surgery. Am J Sports Med 2013;41:2065–73.
24. Harrell FE Jr, Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. J Natl Cancer Inst 1998;80:1198–202.
25. Frank JM, Harris JD, Erickson BJ, Slikker W III, Bush-Joseph CA, Salata MJ, et al. Prevalence of femoroacetabular impingement imaging findings in asymptomatic volunteers: a systematic review. Arthroscopy 2015;31:1199–204.
26. Wyles CC, Heidenreich MJ, Jeng J, Larson DR, Trousdale RT, Sierra RJ. The John Charnley Award: redefining the natural history of osteoarthritis in patients with hip dysplasia and impingement. Clin Orthop Relat Res 2017;475:336–50.
27. Hashemi SA, Dehghani J, Vosoughi AR. Can the crossover sign be a reliable marker of global retroversion of the acetabulum? Skeletal Radiol 2017;46:17–21.
28. Peters S, Laing A, Emerson C, Mutchler K, Joyce T, Thorborg K, et al. Surgical criteria for femoroacetabular impingement syndrome: a scoping review. Br J Sports Med 2017;51:1605–10.

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1209–1215
DOI 10.1002/acr.23874

American College of Rheumatology
Empowering Rheumatology Professionals

Prevalence of Arthritis and Rheumatoid Arthritis in Coal Mining Counties of the United States

Gabriela Schmajuk,[1] Laura Trupin,[2] Edward Yelin,[2] and Paul D. Blanc[1]

Objective. Exposure to inhaled mineral dust, in particular silica, is associated with increased odds of rheumatoid arthritis (RA) and other autoimmune diseases. We studied the association of RA with work-related coal and silica exposure in the Appalachian region of the US.

Methods. We carried out a random-digit dialed telephone survey in selected counties in Appalachia that had elevated coal workers' pneumoconiosis mortality. Our study cohort included men ages ≥50 with any employment history, and we assessed exposure to coal mining employment, other work-related dust, and ergonomic factors. We ascertained self-reported physician diagnosis of any arthritis and of RA with glucocorticoid treatment. We used multivariable logistic regression analysis to estimate the odds ratios (ORs) and associated population attributable fraction (PAF) estimates.

Results. Among the 973 men who met study entry criteria (mean ± SD ages 66 ± 10 years; 54% ever smokers), 266 (27%) reported coal mining work and 189 (19%) reported other work-related silica exposure. There were 517 men (53%), who reported any arthritis and 112 (12%) whose disease met the study definition of RA. Adjusting for covariates, coal mining was associated with elevated odds of RA (OR 3.6 [95% confidence interval (95% CI) 2.1–6.2]), which accounted for a PAF of 33% (95% CI 26–40%) of the men studied. For any arthritis, the coal mining–associated OR was 2.3 (95% CI 1.6–3.2), with an associated PAF of 20% (95% CI 14–25%).

Conclusion. In this population of older males living in a coal mining region, we estimated that 20% of arthritis and 33% of RA may be attributable to coal mining work.

INTRODUCTION

Rates of arthritis are elevated in states in the US that have large numbers of coal miners. Based on 2015 data, West Virginia has the highest prevalence of arthritis among adult males (32.8%, age adjusted) and one of the narrowest gender gaps (only 1.4% lower than the prevalence in women) of any state in the US (1). Further, Tennessee and Kentucky have the third and fourth leading state rates for arthritis among adult males (26.3% and 26.2%, respectively), while Pennsylvania is ranked 14 and Ohio is 16. In comparison, the age-adjusted prevalence rate of arthritis in adult men in California, a state without many coal miners, is 15.7%. The cause of this geographic clustering is unknown, either for degenerative arthritis (the dominant form of the condition) or for inflammatory/autoimmune arthritis.

Multiple independent studies have found that occupational exposure to mineral dust is strongly associated with rheumatologic disease risk (2–4). Rheumatoid arthritis (RA) has been the condition most strongly implicated in mineral dust inhalation (5–7). The role that coal and silica dust inhalation may play in the colocation of US regions in which coal mining is concentrated and where there is a high prevalence of arthritis in males is not clear. The aim of this study was to examine whether coal mining is the nexus between the prevalence of arthritis, especially rheumatoid disease, and being a man in Appalachia.

To address the question of whether a job in coal mining explains, at least in part, the elevated prevalence of arthritis in West Virginia and surrounding areas, we conducted a population-based survey of men ages ≥50, living in coal mining areas in the Appalachian region. The survey included items to identify

Supported by the Russell/Engleman Rheumatology Research Center for Arthritis.

Dr. Schmajuk's work was supported by the NIH (grant K23-AR-063770 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases). Dr. Blanc's work was supported by the Alpha Foundation (grant AFC618-5).

[1]Gabriela Schmajuk, MD, MSc, Paul D. Blanc, MD, MSPH: University of California, San Francisco and the San Francisco Veterans Affairs Health Care System, San Francisco, California; [2]Laura Trupin, MPH, Edward Yelin, PhD: University of California, San Francisco.

Address correspondence to Paul D. Blanc, MD, MSPH, Box 0924, University of California San Francisco, San Francisco, CA 94143-0924. E-mail: Paul.Blanc@ucsf.edu.

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

Submitted for publication October 13, 2018; accepted in revised form March 5, 2019.

SIGNIFICANCE & INNOVATIONS

- Silica has a well-established association with rheumatoid arthritis (RA), but coal mining, especially in the US, has not been well-studied.
- To our knowledge, this is the first modern US population-based study of degenerative arthritis and RA showing a strong association with coal mining.
- The potential occupational risk of arthritis in a current or retired coal miner should be considered by clinicians treating such patients.

occupational exposure to coal dust and to silica exposure, self-reported diagnoses of arthritis (including autoimmune disease subtypes), smoking (given its association with RA), and ergonomic factors that characterize the industrial conditions that miners and other silica-exposed workers face (given their likely association with osteoarthritis [OA]).

MATERIALS AND METHODS

Data source. Data for this study derive from a random-digit dial (using both landline and cellular phone sampling) population-based telephone survey of men ages ≥50 with a history of labor force participation who reside in coal mining areas. We targeted persons living in Appalachia (selected counties in Kentucky, Ohio, Pennsylvania, Tennessee, Virginia, and West Virginia) with historically high mortality rates from coal workers' pneumoconiosis based on data from the National Institute for Occupational Safety and Health (8). The study was approved by the University of California, San Francisco Committee on Research, and all participants provided verbal consent to proceed with the interview.

Study sample. From 30,448 call attempts, we made 7,710 contacts with potential participants, 3,704 were excluded for age, sex, or language (non-English) or because they resided outside of the catchment area. There were 3,003 eligible individuals who refused to participate and an additional 30 who reported no work history, leaving a final study sample of 973 (24% of eligible contacts).

Survey instrument. The brief interview (average time 10 minutes) addressed employment, smoking history, sociodemographics, and arthritis and related diagnoses. Where appropriate, questions were adapted from standard survey items, most importantly in assessing self-report of a health care provider diagnosis of a health condition, the approach used by the US National Health Interview Survey. Duration and type of coal mining experience were ascertained, as well as type and duration of exposure to inhaled dusts (employment for ≥1 year that involved "regular exposure to breathing dusty air"). The employment section also included a 13-item list of physical work hazards (e.g., lifting, bending, using power tools), experienced regularly on any job held for at least 1 year. The health section ascertained whether the respondent had ever received a diagnosis from a health professional of arthritis of any kind, with follow-up items to specify RA, psoriatic arthritis (PsA), or gout. Other autoimmune conditions, including systemic lupus erythematosus (SLE), PsA, and systemic sclerosis (SSc), were also queried. All respondents were also asked about joint swelling, stiffness, or pain, and those who responded affirmatively were asked if they had ever been treated with oral glucocorticoids ("prednisone or steroid pills") for these symptoms.

Disease classification. Arthritis classification was based on an affirmative response to the primary question regarding receipt of a health care provider's diagnosis of arthritis. RA was defined based on the follow-up question about the type of arthritis, restricted to individuals who also reported receiving glucocorticoids for joint symptoms. A non-RA arthritis category was also defined and included all those who responded positively to the initial arthritis question, but did not meet the study definition of RA. This category is likely to include predominately degenerative arthritis, but includes those with RA who were not being treated with corticosteroids as well as persons with other autoimmune arthritis. Our rationale for these definitions was to increase specificity of the RA classification, recognizing that, as a result, the non-RA category may be less precise.

Exposure classification. We categorized coal and silica dust exposure based on questionnaire responses. Coal mining was based on either occupational history of coal mining employment or self-report of coal dust exposure. Other silica dust exposure (among occupations other than coal mining) was based on affirmative responses to any of a list of 7 categories of exposure, including silica, sand, or concrete dust, sandblasting, rock drilling or roof bolting, rock crushing or quarry work, foundry work, concrete finishing, cutting, or drilling, or masonry work or tip-pointing (items that did not specifically elicit employment or history of employment in selected other, less frequent silica trades in the region such as glassmaking or pottery works). We assessed lifetime employment without regard to longest-held job, but did elicit total years of employment in jobs with dust exposure. On an empiric basis, the ergonomics score based on the checklist that we developed for this study was dichotomized at the top quartile (11–13 points versus <11 points).

Statistical analysis. We used multivariable logistic regression analysis to separately model the risk of all arthritis, RA, and non-RA arthritis associated with coal mining employment and other silica exposure. We evaluated either of these exposures in

additional models. All models controlled for age, race/ethnicity (Hispanic or nonwhite versus white non-Hispanic), smoking status (current, former, never), and for high levels of ergonomic exposure. We also calculated the population attributable fraction (PAF) of prevalence to estimate the proportion of prevalent cases (in males) that could be attributed to coal and/or silica exposure, following the method originally proposed in a study by Greenland and Drescher (9) that uses maximum likelihood estimates from multivariable logistic regression models. We tested interaction terms between coal/silica exposure and smoking status for the odds of disease. In order to examine further potential interaction between ergonomic factors and coal/silica, we carried out an analysis stratified by level of ergonomic factors and conducted a formal test of interaction between high ergonomic score and coal/silica exposure. Statistical analyses were carried out using SAS, version 9.4 and Stata, version 15.

RESULTS

Of the 973 respondents meeting study entry criteria, 888 (91%) were white and the mean age was 66 years (Table 1). A total of 852 respondents (87%) were ever smokers, among whom the mean pack-years was 29.8 (median 22). More than half of the respondents reported having received a diagnosis of arthritis from a health care provider. A diagnosis of RA was reported by 188 respondents (19%), but restricting these reports to a more conservative case definition of RA with glucocorticoid treatment at any point yielded a disease prevalence of 12%. Three percent of those surveyed (n = 30) reported at least 1 non-RA autoimmune condition. The prevalence of these autoimmune conditions included 7 respondents with SLE, 5 with SSc, and 20 with PsA. There was overlap among diagnoses, including 11 respondents who also were in the RA with glucocorticoids group.

Table 1. Subject demographics, smoking status, and arthritis diagnoses*

Characteristics	Frequency
Age, mean ± SD years	66.0 ± 9.6
Race/ethnicity	
White, non-Hispanic	88 (91)
Black	31 (3)
Hispanic	16 (2)
Asian/other	38 (4)
Smoking status	
Never smoker	452 (46)
Former smoker	400 (41)
Current smoker	121 (12)
Pack years (among ever smokers), mean ± SD/median (25th–75th percentile)	29.8 ± 28.8/ 22 (9–43)
Reported health care provider arthritis diagnosis	
Any arthritis diagnosis reported	517 (53)
Arthritis, excluding RA	329 (34)
Any RA reported	188 (19)
RA, without ever prednisone use	76 (8)
RA, with ever prednisone use	112 (12)
Any other autoimmune arthritis (not mutually exclusive)	30 (3)
SLE	7 (1)
SSc	5 (1)
PsA	20 (2)

* Values are the number (%) unless indicated otherwise. RA = rheumatoid arthritis; SLE = systemic lupus erythematosus; SSc = systemic sclerosis; PsA = psoriatic arthritis.

Table 2. Exposure status for 973 survey respondents with any work history

Employment status and exposures	No. (%)
Ever employed	973 (100)
Currently employed	407 (42)
Any coal mining employment	266 (27)
Underground coal mining	133 (14)
Any dust exposure	524 (54)
Non-silica dust exposure	124 (13)
Silica exposure, any	400 (41)
Silica exposure, non-coal*	189 (19)
Coal mining and/or silica exposure)	455 (47)

* Coal and silica are not wholly overlapping; 55 respondents with coal employment did not also report any of a checklist of 7 sources of silica exposure, including silica, sand, or concrete dust, sandblasting, rock drilling or roof bolting, rock crushing or quarry work, foundry work, concrete finishing, cutting, or drilling, masonry work or tip-pointing.

More than 1 in 4 respondents (n = 266; 27%) reported coal mining employment, 50% of whom reported work underground, which confers a higher exposure to coal dust (Table 2). The mean ± SD duration of coal mining employment was 21 ± 13 years (60 miners had worked ≥30 years). Independent of coal mining experience, more than half (54%) of those miners with any current or past employment reported regular exposure to dusty air, with a mean ± SD duration of 22 ± 14 years (data on length of exposure not shown). Of the 973 participants surveyed, 400 (41%) responded positively to at least 1 of the 7-item silica exposure checklist. Among the 133 underground coal miners, 78 (59%) reported rock drilling or roof bolting work, a coal mining task recognized to confer high silica exposure (10). A total of 455 respondents (47%) reported either coal mining or silica dust exposure.

Table 3 shows the frequency of 13 work-related ergonomic factors experienced on any job for ≥1 year "on a daily or almost daily basis." The exposure prevalence differed significantly in 3-way comparisons among coal mining exposure, other silica exposure, and all other types of exposure, with a substantially lower prevalence in the latter group. Those reporting ≥11 ergonomic factors comprised 30% of the entire group, but made up more than half

Table 3. Ergonomic factors associated with coal and other silica exposure*

Ergonomic factors	All respondents (n = 973)	Coal mining (n = 266)	Other silica (n = 189)	Neither (n = 518)
Lifting/carrying >30 lbs.	69.8	83.1	93.7	54.2
Arms overhead	55.3	64.3	77.8	42.5
Knee bend/squat/kneel	69.3	82.7	91.0	54.4
Back bend/twist	68.1	81.6	88.4	53.9
Hand grip/wrist bend	71.8	85.0	89.9	58.5
Shaking/vibrating equipment	36.3	57.1	58.7	17.4
Hammer/chisel/saw/drill	50.5	71.1	80.4	29.0
Stoop over	71.1	82.7	91.0	57.9
Pneumatic tools	32.7	51.5	56.1	14.5
Pedal/treadle	37.3	53.0	46.0	26.1
Push/pull >50 lbs.	55.5	71.1	80.4	38.4
Neck twist/bend	51.7	73.3	66.7	35.1
Stand >8 hrs/day	65.6	73.3	83.6	55.0
High exposure (≥11 factors)	30.4	50.8	53.4	11.6

* Values are the percent of patients. All measures differed at $P < 0.001$ across the 3 occupational groups.

of the coal mining and other silica exposure groups, while making up only 12% of the all other types of exposure group.

The estimated odds for all arthritis, RA, and arthritis without RA are shown in Table 4. We estimated the odds for RA excluding the 407 respondents who reported having non-RA arthritis or other rheumatic autoimmune diseases (SLE, SSc, or PsA) without concomitant RA, and the models for non-RA arthritis excluded those who reported having RA (n = 112). Coal mining was associated with more than 2 times the odds of having arthritis (odds ratio [OR] 2.3 [95% confidence interval (95% CI) 1.6–3.2]). The estimated OR of RA associated with coal mining was 3.6 (95% CI 2.1–6.2) and for non-RA arthritis it was 2.0 (95% CI 1.4–2.9). Silica exposure, exclusive of coal mining, was also associated with increased odds of any arthritis (OR 1.8 [95% CI 1.2–2.6]), RA (OR 2.1 [95% CI 1.1–3.9]) and non-RA arthritis (OR 1.7 [95% CI 1.2–2.6]). Exposure to 11 to 13 ergonomic factors was associated with statistically significant and increased odds of any arthritis (OR 1.5 [95% CI 1.1–2.0]) and non-RA arthritis (OR 1.4 [95% CI 1.01–2.0]), while the odds of RA were slightly higher, but the CI did not exclude 1.0 (OR 1.6 [95% CI 0.97–2.8]). Current smoking was associated with 2 times the odds of RA (OR 2.0 [95% CI 1.1–3.9]) but was not associated with either any arthritis or non-RA arthritis.

Table 4. Multivariate analysis: arthritis and RA associated with coal and silica exposure adjusted for smoking, ergonomic factors, age, and race/ethnicity*

Associated factors	All arthritis model (n = 973)	RA model (n = 566)†	Non-RA arthritis model (n = 861)‡
Coal and silica exposure			
Coal mining work	2.3 (1.6–3.2)	3.6 (2.1–6.2)	2.0 (1.4–2.9)
Silica, no coal exposure	1.8 (1.2–2.6)	2.1 (1.1–3.9)	1.7 (1.2–2.5)
Smoking			
Current	1.2 (0.8–1.9)	2.0 (1.1–3.9)	1.1 (0.7–1.7)
Former	1.1 (0.8–1.5)	1.2 (0.7–1.9)	1.1 (0.8–1.5)
Ergonomic exposure			
11–13 factors	1.5 (1.1–2.0)	1.6 (0.97–2.8)	1.4 (1.01–2.0)
Age (per year)	1.03 (1.01–1.04)	1.04 (1.01–1.06)	1.03 (1.01–1.04)
Hispanic ethnicity or nonwhite race	1.4 (0.8–2.4)	1.3 (0.6–3.0)	1.5 (0.9–2.5)

* Values are the odds ratio (OR) 95% confidence interval (95% CI). For coal and silica, referent category = neither exposure; for smoking, referent = never smoker; for ergonomic exposure, referent category = 0 to 10 factors; for race/ethnicity, referent category = White, non-Hispanic; RA = rheumatoid arthritis.
† Excludes 407 reporting non-RA arthritis or selected autoimmune diseases without concomitant RA.
‡ Excludes 112 participants reporting RA and glucocorticoid treatment.

Table 5. All arthritis and RA population attributable fraction associated with coal and silica exposure*

Exposure	All arthritis	RA
Coal and silica exposure		
Either exposure	29 (21–37)	44 (31–54)
Coal mining work	20 (14–25)	33 (26–40)
Other occupational silica exposure	10 (5–14)	10 (4–16)

* Values are the percent of population attributable risk (PAF) 95% confidence interval (95% CI). RA = rheumatoid arthritis. For coal or silica exposure, for all arthritis OR = 2.1 (95% CI 1.5–2.8); for RA, OR = 2.9 (95% CI 1.8–4.9). For coal mining work and other silica in model adjusting for each, see ORs in Table 4. All estimates derived from multivariable models adjusted for age, race/ethnicity, smoking, and ergonomic exposures.

Former smoking had no statistically significant associations with disease status. There was no statistical evidence supporting an interaction effect for smoking and coal or silica exposure in association with disease in any of the models used in the present study.

In order to examine the potential interaction between ergonomic factors and coal/silica exposure for the odds of arthritis, we carried out an analysis stratified by level of ergonomic factors. Among those respondents who reported 11–13 ergonomic factors (n = 296), there was no statistically significant association between combined coal or silica exposure and the odds of all arthritis (OR 1.5 [95% CI 0.8–2.7]). Among the stratum of respondents with a lower ergonomic burden (n = 677), the OR was higher and statistically significant (OR 2.2 [95% CI 1.6–3.1]). A formal test of the interaction term between coal or silica exposure and ergonomic factors, however, was not statistically significant ($P = 0.24$). To assess the coal work and silica exposure burden for all arthritis and RA, we estimated the PAF for either coal or silica exposure and for coal and silica separately (Table 5). For coal or silica exposure, the PAF for all arthritis was 29% (95% CI 21–37%) and was 44% (95% CI 31–54%) for RA. The major contributor was coal mining: the PAF for all arthritis was 20% and was 33% for RA.

DISCUSSION

In this population-based study of arthritis in Appalachia, 1 in 2 men older than the age of 50 reported having arthritis; more than 1 in 10 met our case definition of RA. Just over one-quarter of men reported coal mining work and 47% altogether either had been coal miners or had been otherwise occupationally exposed to silica. Because this exposure was common and, because the odds of RA were substantially increased in association with such exposure, we estimated that fully a third of the RA cases in the men in the present study were attributable to coal work and, combining that with other silica exposures, the PAF was 42%.

Our findings are consistent with previous studies of coal and silica exposure. In the early 1950s, there were nearly simultaneous observations that both silica dust (nearly pure silica "flour") and coal mining work were associated with RA. These observations were revealed in a study by Colinet in Belgium (11) and Caplan and colleagues in the UK (12,13). By the 1990s, researchers identified mineral dust as a factor in a range of autoimmune diseases (2–4,14,15). Although much of the biomedical literature has focused on silica, there is emerging recognition that coal dust (with likely silica co-exposure, much of it of a particle size in the respirable range) represents an important factor in what has come to be recognized more broadly as "coal mine dust lung disease" (16,17).

Despite this, there have been relatively few studies of RA among US coal miners. Nearly 50 years ago, a 1969 community-based study that included 560 miners ages 20–69 years in West Virginia observed that radiographic OA of the hands was present in 40.2% of the miners (18). A 1973 serologic study of 207 underground coal miners in Pennsylvania and West Virginia (all with radiographic disease) found that 6% of miners were positive for rheumatoid factor and 34% were positive for antinuclear antibodies (19). Contemporaneous clinical studies of coal miners from the same region suggested that exposure-related RA was more common than appreciated (20,21). Only 1 other study of rheumatologic disease in Appalachian coal miners has appeared since that time (1981), which found that, among 353 miners (130 without radiographic lung disease), 69 (19.5%) were RA positive (22). A recent extensive review of occupational RA included only a brief mention of coal mining and did not contain any recent citations on that subject (23).

The differences that we observed in coal mining associated odds of RA juxtaposed with odds for all other arthritis suggest interesting insights. The estimated coal mining odds for RA (OR 3.5) were greater than for all other arthritis (OR 2.0), consistent with a strong relationship to autoimmune arthritis in particular. The odds for RA associated with other silica, however, was only slightly higher than that for other arthritis, although most of those with silica exposure had concomitant coal mining experience.

We recognized, a priori, that ergonomic exposures in coal mining might also be associated with increased arthritis risks. This was, in fact, the case (OR 1.5 for all arthritis, 1.7 for RA, and 1.4 for arthritis excluding RA). Ergonomic factors with established relationships to degenerative arthritis included kneeling, bending, squatting, crawling, whole-body vibration, lifting heavy loads, and repetitive motion (24–29). In degenerative arthritis of the knee, coal mining–specific data show a strong link to disease (30–36). It is reasonable to assume that this can be generalized to other body parts as well. Thus, the arthritis that we observed is consistent with the pattern of ergonomic factors reported by the coal- and silica-exposed participants in our study. Nonetheless, ergonomics alone does not account for all the associated odds, given that our multivariable modeling of the coal mining OR for arthritis and RA took into account ergonomic factors. Indeed, the association of a high ergonomic load with RA was of a similar

magnitude as that for arthritis overall. Further, we did not identify a significant interaction between ergonomic exposures and coal or silica exposure for arthritis risk, although study power was limited, with fewer than 300 respondents in the stratum with a heavy ergonomic load. We also cannot exclude unmeasured confounding that might explain the associations we observed, for example, body habitus or socioeconomic factors beyond work itself.

The findings revealed in our study have potential limitations. The diagnosis of RA that we used was based on respondent report of a health care provider's diagnosis. This is the approach used in many questionnaire-based studies, most notably the US National Health Interview Survey. Nonetheless, self-reported disease can be subject to random misclassification or reporting bias. Random misclassification of disease should have resulted in a reduced association of RA with coal or silica exposures. In contrast, systematic reporting bias could lead persons with coal mining histories to be more likely to report disease, thus leading to a false association. Our telephone survey length was constrained such that we could not ascertain duration or dosage of reported corticosteroids nor obtain a detailed history of other disease modifying antirheumatic drugs. Further, we did not have access to medical records, serologic data (e.g., for rheumatoid factor), or physical examinations, which are all sources of confirmatory data that mitigate against reporting bias. Our finding of a 53% prevalence of arthritis overall, although high, is consistent with estimated rates in Appalachia (1). Nonetheless, it is probable that some persons with degenerative arthritis but not autoimmune disease misreported their condition as RA. In particular, the term "rheumatism," as it is commonly used, may manifest geographic regional differences that magnify this problem in the counties from which we recruited (37). We attempted to address this, in part, by using a conservative definition of disease that also included reported glucocorticoid use. In addition, the higher OR associated with coal mining exposure that we observed for RA as compared to other arthritis, noted previously, argues against selective overreporting of arthritis as RA among coal miners. Further, the association with current smoking for RA (OR 2.0), but not all arthritis (OR 1.2), is consistent with previous observations specific to RA generally but also in silica exposure (6). Our observed overall response rate (24%), although similar to that reported for nonfederal telephone surveys, is less than that of the Behavioral Risk Factor Surveillance System (38). Because we do not have information on nonparticipants we cannot assess unmeasured selection effects that this response rate may represent. Finally, the high prevalence of disease in the study population could lead to overestimation of the association with coal and silica exposure when relying on prevalence ORs.

In summary, our findings of increased ORs for arthritis and RA among coal miners in Appalachia are robust, unlikely explained by biased reporting or confounding, and are consistent with other studies that focused primarily on silica exposure outside of coal mining. This association, at the regional level of Appalachia, is relevant to the delivery of health care services and to individual case attribution and compensation. Tertiary prevention of disease progression and disability is especially noteworthy. Given that treatment guidelines for RA indicate that a disease-modifying agent should be initiated soon after onset of disease, earlier disease detection could be achieved through targeted surveillance among current and former coal miners. However, the results reported in the present study suggest that primary prevention of arthritis through workplace protections against dust inhalation may reduce the prevalence of arthritis in general and RA in particular.

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. Blanc 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. Schmajuk, Trupin, Yelin, Blanc.
Acquisition of data. Schmajuk, Trupin, Yelin, Blanc.
Analysis and interpretation of data. Schmajuk, Trupin, Yelin, Blanc.

REFERENCES

1. Centers for Disease Control. State-Specific 2015 BRFSS arthritis prevalence estimates. URL: https://www.cdc.gov/arthritis/data_statistics/state-text.htm.
2. Steenland K, Goldsmith DF. Silica exposure and autoimmune diseases. Am J Ind Med 1995;28:603–8.
3. Parks CG, Conrad K, Cooper GS. Occupational exposure to crystalline silica and autoimmune disease. Environ Health Perspect 1999;107 Suppl 5:793–802.
4. Lee S, Hayashi, Mastuzaki H, Kumagai-Takei N, Otsuki T. Silicosis and autoimmunity. Curr Opin Allergy Clin Immunol 2017;17:78–84.
5. Stolt P, Källberg H, Lundberg I, Sjögren B, Klareskog L, Alfredsson L, et al. Silica exposure is associated with increased risk of developing rheumatoid arthritis: results from the Swedish EIRA study. Ann Rheum Dis 2005;64:582–6.
6. Stolt P, Yahya A, Bengtsson C, Källberg H, Rönnelid J, Lundberg I, et al. Silica exposure among male current smokers is associated with a high risk of developing ACPA-positive rheumatoid arthritis. Ann Rheum Dis 2010;69:1072–6.
7. Blanc PD, Järvholm B, Torén K. Prospective risk of rheumatologic disease associated with occupational exposure in a cohort of male construction workers. Am J Med 2015;128:1094–101.
8. Centers for Disease Control. URL: Coal workers' pneumoconiosis: top 50 counties with highest age-adjusted death rates (per million population), U.S. residents age 15 and over, 2000–2009. https://wwwn.cdc.gov/eworld/Data/Coal_Workers_Pneumoconiosis_Top_50_counties_with_highest_age-adjusted_death_rates_per_million_population_US_residents_age_15_and_over_20002009/767.
9. Greenland S, Drescher K. Maximum likelihood estimation of the attributable fraction from logistic models. Biometrics 1993;49:865–72.
10. Grové T, Van Dyk T, Franken A, Du Plessis J. The evaluation and quantification of respirable coal and silica dust concentrations: a task-based approach. J Occup Environ Hyg 2014;11:406–14.
11. Colinet E. Polyarthrite chronique evolutive et silcose pulmonaire. Acta Physiother Rheumatol Belg 1953;8:37–41.
12. Caplan A. Certain unusual radiological appearances in the chest of coal-miners suffering from rheumatoid arthritis. Thorax 1953;8:29–37.

13. Miall WE, Caplan A, Cochrane AL, Kilpatrick GS, Oldham PD. An epidemiological study of rheumatoid arthritis associated with characteristic chest x-ray appearances in coal-workers. Br Med J 1953;2:1231–6.
14. Sanchez-Roman J, Wichmann I, Salaberri J, Varela JM, Nuriez-Roldan A. Multiple clinical and biological autoimmune manifestations in 50 workers after occupational exposure to silica. Ann Rheum Dis 1993;52:534–8.
15. Koeger AC, Lang T, Alcaix D, Milleron B, Rozenberg S, Chaibi et al. Silica-associated connective tissue disease. A study of 24 cases. Medicine (Baltimore) 1995;74:221–37.
16. Cohen RA, Petsonk EL, Rose C, Young B, Regier M, Najmuddin A, et al. Lung pathology in U.S. coal workers with rapidly progressive pneumoconiosis implicates silica and silicates. Am J Respir Crit Care Med 2016;193:673–80.
17. Cohen RA, Go LH, Green FH. Coal mine dust lung disease. In: Newman, Taylor A, Cullinan P, Blanc P, Pickering A, editors. Parkes' occupational lung disorders. 4th ed. Boca Raton: CRC Press; 2017;207–24.
18. Lockshin MD, Higgins IT, Higgins MW, Dodge HJ, Canale N. Rheumatism in mining communities in Marion County, West Virginia. Am J Epidemiol 1969;90:17–29.
19. Lippmann M, Eckert HL, Hahon N, Morgan WK. Circulating antinuclear and rheumatoid factors in coal miners: a prevalence study in Pennsylvania and West Virginia. Ann Intern Med 1973;79:807–11.
20. Benedek TG. Rheumatoid pneumoconiosis: documentation of onset and pathogenic considerations. Am J Med 1973;55:515–24.
21. Benedek TG, Zawadzki ZA, Medsger TA Jr. Serum immunoglobulins, rheumatoid factor, and pneumoconiosis in coal miners with rheumatoid arthritis. Arthritis Rheum 1976;19:731–6.
22. Pearson DJ, Mentnech MS, Elliott JA, Price CD, Taylor G, Major PC. Serologic changes in pneumoconiosis and progressive massive fibrosis of coal workers. Am Rev Respir Dis 1981;124:696–9.
23. Murphy D, Hutchinson S. Is male rheumatoid arthritis an occupational disease? A review. Open Rheum J 2017;11:88–105.
24. 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.
25. Allen KD, Golightly YM. State of the evidence. Curr Opin Rheumatol 2015;27:276–83.
26. Palmer KT. Occupational activities and osteoarthritis of the knee. Br Med Bull 2012;102:147–70.
27. Yucesoy B, Charles LE, Baker B, Burchfiel CM. Occupational and genetic risk factors for osteoarthritis: a review. Work 2015;50:261–73.
28. Richmond SA, Fukuchi RK, Ezzat A, Schneider K, Schneider G, Emery CA. Are joint injury, sport activity, physical activity, obesity, or occupational activities predictors for osteoarthritis? A systematic review. Orthop Sports Phys Ther 2013;43:515–19.
29. Ezzat AM, Li LC. Occupational physical loading tasks and knee osteoarthritis: a review of the evidence. Physiother Can 2014;66:91–107.
30. Lawrence JS. Aitken-Swan J. Rheumatism in miners. 1: Rheumatic complaints. Br J Ind Med 1952;9:1–18.
31. Kellgren JH, Lawrence JS. Rheumatism in miners. II. X-ray study. Br J Ind Med 1952;9:197–207.
32. Lawrence JS. Rheumatism in miners. Part III occupational factors. Br J Ind Med 1955;12:249–61.
33. Sharrard WJW. Pressure effects on the knee in kneeling miners. Ann R Coll Surg Engl 1965;36:309–24.
34. O'Reilly SC, Muir KR, Doherty M. Occupation and knee pain: a community study. Osteoarthritis Cartilage 2000;8:78–81.
35. McMillan G, Nichols L. Osteoarthritis and meniscus disorders of the knee as occupational diseases of miners. Occup Environ Med 2005;62:567–75.
36. Bhattacherjee A, Bertrand JP, Meyer JP, Benamghar L, Otero Sierra C, Michaely JP, et al. Relationships of physical job tasks and living conditions with occupational injuries in coal miners. Ind Health 2007;45:352–8.
37. Cavender A. Folk medicines in Southern Appalachia. Chapel Hill (NC): UNC Press; 2003. p. 113.
38. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System. 2016 Summary Data Quality Report June 29, 2017. URL: https://www.cdc.gov/brfss/annual_data/2016/pdf/2016-sdqr.pdf.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1216–1223
DOI 10.1002/acr.23779

Effect of Physical State on Pain Mediated Through Emotional Health in Rheumatoid Arthritis

Yukako Nakagami,[1] Genichi Sugihara,[1] Nori Takei,[2] Takao Fujii,[3] Motomu Hashimoto,[1] Kosaku Murakami,[1] Moritoshi Furu,[1] Hiromu Ito,[1] Miyabi Uda,[1] Mie Torii,[1] Kazuko Nin,[1] Toshiya Murai,[1] and Tsuneyo Mimori[1]

Objective. Pain is one of the main symptoms of patients with rheumatoid arthritis (RA). Pain in RA is caused by specific physical changes, such as joint destruction, and is therefore used as a disease activity marker. Although pain can also be influenced by emotional factors, neither the effect of emotional health nor the indirect effect of the physical state mediated by emotional health on pain has been quantified.

Methods. A total of 548 patients with RA participated. Emotional health was assessed using the Hospital Anxiety and Depression Scale (HADS). Measures routinely used in practice were used to evaluate the physical state and pain. To quantify the effects of the physical state on emotional health, and the effects of both physical and emotional health on pain, we used structural equation modeling, with emotional health, physical state, and pain as latent variables.

Results. The prevalence of anxiety and depression (HADS score ≥8 for each) among patients with RA was 18.7% and 29.4%, respectively. Emotional health was significantly influenced by the physical state (β = 0.21). Pain was affected by physical (β = 0.54) and emotional health (β = 0.29). The effect of the physical state on pain was mediated by emotional health, with this mediation effect (β = 0.06) accounting for 10.2% of the total effect.

Conclusion. The magnitude of pain in RA is determined by the mediation effect of emotional health as well as the direct physical state. Our findings suggest that emotional factors should be taken into account when assessing RA disease activity.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic inflammatory disease with a prevalence of 0.5–1% in northern Europe and North America (1,2). RA is characterized by persistent synovitis, which leads to joint destruction and pain (1). Patients with RA experience various problems, which have been classified as physical dimensions (mobility level, walking and bending, hand and finger function, arm function, self-care tasks, and household tasks), social interaction (social activity and support from family and friends), symptoms (pain), role (work), and affect (tension and mood) (3). Approximately 70% of patients with RA report that pain is among the top 3 priorities for improvement (3). Furthermore, when patients evaluate their own RA disease activity, pain is considered the most clinically important component (4). Consistent with this patient perspective, physicians regard pain as an important aspect of RA disease activity measurement and a pivotal treatment target.

Pain experienced by patients with RA is a subjective measure of disease activity that may be affected by emotional factors. For example, pain sometimes persists despite RA remission, especially when patients have emotional problems such as anxiety (5,6). Interestingly, a recent report showed a link between anxiety

Supported by a Grant-in-Aid for Scientific Research B (15H04771), Grant-in-Aid for Scientific Research C (15K09825), Grant-in-Aid for Scientific Research on Innovative Areas (16H06402), Grant-in-Aid for Challenging Exploratory Research (16K15558), and Grants-in-Aid for JSPS Research Fellow (17J00953).

[1]Yukako Nakagami, MD, Genichi Sugihara, MD, PhD, Motomu Hashimoto, MD, PhD, Kosaku Murakami, MD, PhD, Moritoshi Furu, MD, PhD, Hiromu Ito, MD, PhD, Miyabi Uda, RN, MSN, Mie Torii, PhD, Kazuko Nin, RN, PhD, Toshiya Murai, MD, PhD, Tsuneyo Mimori, MD, PhD: Kyoto University Graduate School of Medicine, Kyoto, Japan; [2]Nori Takei, MD, PhD, MSc: Hamamatsu University School of Medicine, Hamamatsu, Japan, and King's College, London, UK; [3]Takao Fujii, MD, PhD: Wakayama Medical University, Wakayama, Japan.

Dr. Fujii has received research grants and/or speaking fees from Pfizer Japan, Ono, Daiichi Sankyo, Mitsubishi-Tanabe, Eisai, AbbVie GK, and Astellas (more than $10,000 each). Dr. Hashimoto has received grants from the Astellas Kyoto University project. Dr. Murakami has received speaking and/or consulting fees from Eisai, Chugai, Pfizer Japan, Bristol-Myers Squibb, Mitsubishi-Tanabe, UCB, Daiichi Sankyo, and Astellas (less than $10,000 each). Drs. Hashimoto, Furu, and Ito belong to a department that receives financial support from Mitsubishi-Tanabe, Chugai, Bristol-Meyers Squibb, AbbVie, and Eisai. Dr. Ito has received research grants from Astellas and Bristol-Myers Squibb. Dr. Mimori has received grants from Astellas, Daiichi Sankyo, Mitsubishi-Tanabe, Ayumi, Esai, and Chugai, and speaking fees from Bristol-Meyers Squibb, Chugai, and Mitsubishi-Tanabe (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Nori Takei, MD, PhD, MSc, Head and Professor, Research Center for Child Mental Development, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu City 431-3192, Japan, or Department of Psychosis Studies, Institute of Psychiatry, Psychosis and Neuroscience, King's College London, London SE5 8AF, UK. E-mail: noritakei-psy@umin.ac.jp.

Submitted for publication April 11, 2018; accepted in revised form September 25, 2018.

SIGNIFICANCE & INNOVATIONS

- Patients with rheumatoid arthritis (RA) have a high incidence of anxiety and depression.
- Pain among patients with RA is influenced by both physical state and emotional health, and in addition, the effect of the physical state on pain is mediated through emotional health.
- Improvement of emotional health in patients with RA may partially alleviate the level of pain.
- RA disease activity in patients with anxiety and/or depressive symptoms may be overestimated, because disease activity measures include pain components.

and/or depression in patients with RA and increased pain 1 year later (7). Based on this finding, a concept of noninflammatory pain in RA has been proposed (7). This proposal suggests that emotional factors have a substantial influence on pain in RA.

Evidence shows that the physical state affects emotional health (8–13). Several hypotheses concerning the mechanisms whereby inflammation potentially influences emotional health have been put forward, including dysregulation of the hypothalamic pituitary adrenal axis, changes in glial function and glutamate release in the brain, and activation of the tryptophan metabolizing enzyme (8,10,12,14). Of particular interest are findings of clinical studies showing that antiinflammatory drugs are effective in treating patients with depression (15). In addition, a meta-analysis showed a prevalence of major depressive disorder among patients with RA of approximately 17%, a rate much higher than in the general population (16). This finding suggests that having an inflammatory condition such as RA confers risks for emotional health.

Considering the information above, pain in RA may be related to both physical state and emotional health, and emotional health may be affected by the physical state. Therefore, in addition to the direct effect of physical condition (RA) on pain, the effect of RA on pain may be mediated by emotional health. This idea led us to construct a model of relationships among physical state, emotional health, and pain, as shown in Figure 1. In this study, we evaluated this model using structural equation modeling (SEM), by quantifying the direct (physical state) and indirect effects (physical state as it affects emotional health) on pain in patients with RA.

In order to evaluate the effect of the physical state on pain through its effect on emotional health, we focused on variables that reflect purely physical and objective components to define physical state. To our knowledge, this is the first study to investigate the effects of the physical state in a comprehensive and quantitative manner. Previous studies have evaluated the physical state using disease activity composite indices (7,17) or a single item such as the erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) level, or tender joint count (TJC) (9,17–19). These evaluations may be problematic, because disease activity composite indices generally include subjective components and therefore may be influenced by emotional factors, and single items may be insufficient to assess multifaceted aspects of the physical state.

PATIENTS AND METHODS

Patients. This study included 548 participants from the 2014 Kyoto University Rheumatoid Arthritis Management Alliance (KURAMA) cohort (20). The KURAMA cohort study was started in 2011 at the Center for Rheumatic Diseases, Kyoto University Hospital, and aimed to investigate RA in clinical practice. Approximately 500 consecutive patients with RA were enrolled annually. Because we included evaluation of anxiety and depression in 2014, we used data assembled in 2014 to investigate the effect of mental and physical states on pain. All participants met the American College of Rheumatology (ACR) revised criteria for RA 1987, or the ACR and European League Against Rheumatism classification criteria for RA 2010 (1,21,22). The study was conducted according to the Declaration of Helsinki. Ethics approval was obtained from the Medical Ethics Committee of Kyoto University Graduate School and Faculty of Medicine. All participants were age >18 years and provided written informed consent.

Measures. To ascertain the direct effect of the physical state and the mediating effect of emotional health on pain, we developed a conceptual model representing relationships among physical state, emotional health, and pain (Figure 1). Physical state, emotional health, and pain were evaluated with separate sets of indicators (symptoms and laboratory findings); that is, each concept was regarded as having a unique latent trait measured by several observed variables. Observed indicators for the 3 latent traits are described in the following paragraph. In addition, demographic (age, sex, employment status) and medical information were included in the analysis. Medical information included current psychiatric treatment, physical complications, and RA treatments such as methotrexate (MTX), disease-modifying antirheumatic drugs (DMARDs) other than MTX, and biologics.

Physical state. The activity of RA cannot be examined by a single clinical or laboratory test, and composite indices using several clinical and laboratory variables have been proposed. Three widely used composite indices are the Disease Activity

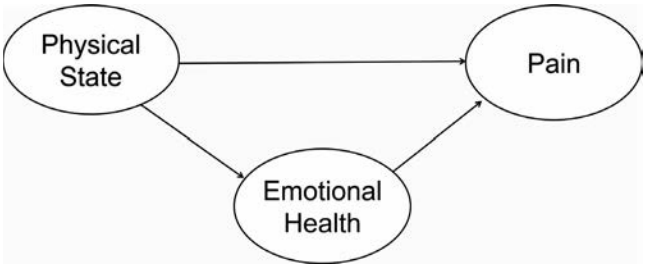

Figure 1. Hypothesized model of the effects of physical and emotional health on pain.

Score in 28 joints (DAS28), the Simplified Disease Activity Index (SDAI), and the Clinical Disease Activity Index (CDAI), all of which are considered valid tools for evaluating RA disease activity (1,20). However, these composite indices include subjective items, which may obscure accurate measurement of disease activity due to modification by emotional factors (4). In this study, we aimed to extract purely physical aspects of RA and evaluate the physical state independently of emotional factors and pain. Therefore, we eliminated subjective components and items associated with pain from the latent structure for RA physical state. We selected observed variables for RA physical state through a 3-step process. All 7 components of the 3 validated RA disease activity measurements (DAS28, SDAI, and CDAI) were extracted: ESR 1-hour level, serum CRP level, swollen joint count (SJC), TJC, patient's global assessment (PtGA) of disease activity, evaluator's global assessment (EGA) of disease activity, and the patient's assessment of general health. Next, we excluded TJC, PtGA, EGA, and general health, because TJC is a pain indicator (see Pain section below), PtGA and general health are both subjective patient self-evaluations and influenced by emotional health (e.g., depression) (4,7), and EGA is a subjective evaluator assessment and is not sensitive to objective RA disease outcome or radiographic progression (4). Last, we added matrix metalloproteinase 3 (MMP-3) to the remaining 3 components (i.e., ESR, CRP, and SJC) to form a physical state set, because MMP-3 has been reported to directly reflect joint inflammation and destruction (23–25). In effect, in our data, the MMP-3 level was correlated with the yearly radiologic progression (change in modified total Sharp/van der Heijde score per year) (Spearman's $r = 0.12$, $P < 0.05$ [n = 344]). We used MMP-3 in our analysis because of the relatively limited data available on radiologic progression. Finally, the physical state comprised 4 variables: ESR, CRP level, MMP-3, and SJC. For SJC, we incorporated 68 joints rather than the 28 joints used in the DAS28, SDAI, and CDAI to make our evaluation as thorough as possible.

Pain. We evaluated pain using 2 aspects: distribution and intensity. We used TJC in 68 joints as an indicator of pain distribution, which the ACR recommends to measure disease activity in clinical trials (26). Pain intensity was ascertained using a visual analog scale (VAS; range 0–100, where 0 = no pain and 100 = the worst pain) as a continuous variable. Thus, we used 2 indicators to structure the latent variable of pain, TJC, and pain VAS.

Emotional health. We used the Hospital Anxiety and Depression Scale (HADS) to assess anxiety and depression (2,7,16,27). The latent trait for emotional health comprised the HADS anxiety and depression subscales. The HADS comprises 14 items (7 items each for anxiety and depression), scored on a 4-point Likert scale (range 0–3), with total subscale scores ranging from 0 to 21. We chose the most commonly used cutoffs. A total subscale score of 0–7 indicated no anxiety/depression, 8–10 indicated possible anxiety/depression, and 11–21 indicated probable anxiety/depression (2,16). The validity of the HADS has been established with a sensitivity and specificity of approximately 0.80, and internal consistency >0.85 (2).

Statistical analysis. We examined the conceptual model (Figure 1) using SEM, a powerful tool that is widely used in sociology, psychology, and other social sciences to analyze complex modeling with latent factors. SEM techniques have also been used in research with patients with RA. For example, SEM analysis revealed that disease activity, mood disturbance, and sleep quality contributed to fatigue in patients with RA (28). Another SEM analysis involving patients with RA showed that self-efficacy for pain partially mediated the relationship between disease activity and pain (29). However, no previous studies have investigated the mediation effect of emotional health on the relationship between the physical state and pain in RA, where the physical state is defined by purely physical and objective components.

Before the main SEM analysis, we examined correlations of indicators constituting the 3 structures (i.e., physical state, emotional health, and pain) using Spearman's correlation coefficients because of their non-normal distribution (indicated by the Shapiro-Wilk test). We used robust maximum likelihood estimation in the SEM analysis to allow for non-normality distributions. Age and sex were included as covariates. Unless otherwise indicated, we used the full information maximum likelihood procedure to manage missing data. Fit between the conceptual model and the data was examined using several fit indices: root mean square error of approximation (RMSEA), comparative fit index, Tucker-Lewis fit index, and standardized root mean square residual (SRMR). Along with model fit procedures, we scrutinized modification indices to reach the optimal final model. To verify the indirect effect of physical state on pain through emotional health, we carried out the Sobel test; the 95% confidence interval (95% CI) of the effect was obtained from the bias-corrected bootstrap estimation, using 10,000 samples (30,31).

The Wald test (32) was used to evaluate any difference of moderated mediation effect in binary factors, including employment status, presence/absence of current psychiatric treatment, any physical complications (extraarticular manifestations, connective tissue diseases, and respiratory diseases), and RA treatments. SEM analyses were performed using Mplus software, version 7.31 for Mac. Other statistical analyses were performed using STATA software, version 14.

RESULTS

Demographic characteristics. As shown in Table 1, 456 (83.5%) of the 548 participants were women, which is consistent with previous studies showing approximately 70–90% of individuals with RA are female (6,8,17). The participants' mean ± SD age was 62.6 ± 13.1 years, and the mean ± SD disease duration was 14.8 ± 12.0 years. Classification by Steinbrocker stage showed that 118 participants (21.6%)

Table 1. Participants' demographic characteristics*

Characteristic	Values	No.
Age, mean ± SD years	62.6 ± 13.1	546
Women	456 (83.5)	546
Employed (including housework)	410 (74.8)	548
Current psychiatric treatment	25 (4.6)	540
RA treatments		548
MTX	370 (67.5)	
DMARDs other than MTX	213 (38.9)	
Biologic drugs	216 (39.4)	
Stage		546
I	118 (21.6)	
II	135 (24.7)	
III	101 (18.4)	
IV	192 (35.1)	
Disease duration from the onset of symptoms, mean ± SD years	14.8 ± 12	443
Physical complications		547
Extraarticular manifestations	110 (20.1)	
Connective tissue diseases	17 (3.1)	
Respiratory diseases	149 (27.2)	
EQ-5D, mean ± SD	7.5 ± 2	548
HADS anxiety, mean ± SD†	4.8 ± 3.7	539
Possible anxiety	56 (10.4)	
Probable anxiety	45 (8.3)	
HADS depression, mean ± SD‡	5.8 ± 3.7	537
Possible depression	93 (17.3)	
Probable depression	65 (12.1)	
ESR 1h, mm/hour		527
Mean ± SD	22.5 ± 18.7	
Median (IQR)	16 (9–29)	
CRP, mg/dl		515
Mean ± SD	0.38 ± 0.73	
Median (IQR)	0.1 (0–0.3)	
MMP-3, ng/ml		528
Mean ± SD	103.5 ± 110.6	
Median (IQR)	68.05 (44.5–120.9)	
SJC		529
Mean ± SD	1.03 ± 1.87	
Median (IQR)	0 (0–1)	
TJC		529
Mean ± SD	1.18 ± 2.08	
Median (IQR)	0 (0–2)	
Pain VAS		547
Mean ± SD	27.8 ± 25.8	
Median (IQR)	18 (5–49)	

* Values are the number (%) unless indicated otherwise. RA = rheumatoid arthritis; MTX = methotrexate; DMARDs = disease-modifying antirheumatic drugs; EQ-5D = EuroQol 5 domain instrument; HADS = Hospital Anxiety and Depression Scale; ESR 1h = erythrocyte sedimentation rate 1-hour level; IQR = interquartile range; CRP = C-reactive protein; MMP-3 = matrix metalloproteinase 3; SJC = swollen joint count; TJC = tender joint count; VAS = visual analog scale.

† Possible anxiety: score of 8–10 on the anxiety subscale of HADS; probable anxiety: score of 11–21 on the anxiety subscale of HADS.

‡ Possible depression, score of 8–10 on the depression subscale of HADS; probable depression, score of 11–21 on the depression subscale of HADS.

were stage I, 135 (24.7%) were stage II, 101 (18.4%) were stage III, and 192 (35.1%) were stage IV. The distribution of Steinbrocker stages was similar to that of a survey of prevalent RA cases in Akita Prefecture, a northern area of Japan (33), implying that our sample was representative. The representativeness of the sample was further supported by the fact that complications with RA in the current study were comparable with those in other studies (1,8). In our study, 27.2% of participants (n = 149) had respiratory diseases, 20.1% (n = 110) had extraarticular manifestation, and 3.1% (n = 17) had connective tissue diseases.

Variables related to the physical state and pain are summarized in Table 1. For emotional factors, the mean ± SD HADS anxiety score was 4.8 ± 3.7, median 4, and interquartile range (IQR) 2–7; and the mean ± SD HADS depression score was 5.8 ± 3.7, median 5, and IQR 3–8. HADS scores indicated that 56 patients (10.4%) had possible anxiety and 45 (8.3%) had probable anxiety, and that 93 (17.3%) had possible depression and 65 (12.1%) had probable depression.

Correlation analysis. Spearman's correlation analysis showed significant correlations with a medium effect size ($r > 0.3$) (34) among observed variables representing the same latent factor, except 1 relationship between SJC and ESR 1hour ($r = 0.272$) (Table 2).

SEM. Following scrutiny of modification indices during the model fit procedures, the final model included covariance between ESR and CRP, CRP and MMP-3, and SJC and TJC (Figure 2). Inclusion of these covariances was clinically justified for 3 reasons. First, ESR and CRP are markers of inflammation and have a strong correlation among RA patients (1,23). Second, CRP and MMP-3 reflect radiologic damage and are related to each other in the course of RA with/without progression of radiologic damage (23). Third, SJC and TJC represent symptomatic (i.e., swollen and tender) joint counts, and it is reasonable to assume that they are closely related. Figure 2 shows factor loadings for the 3 latent factors. The fit between the hypothesized model and the data was satisfactory (32): RMSEA = 0.069, comparative fit index = 0.917, Tucker-Lewis fit index = 0.847, and SRMR = 0.040.

The path coefficient in Figure 2 showed a significant direct effect of the physical state on pain ($\beta = 0.54$, $P < 0.001$). Path analyses showed that the physical state exerted a significant effect on emotional health ($\beta = 0.21$, $P < 0.05$). In turn, emotional health had a significant influence on pain ($\beta = 0.29$, $P < 0.01$). These results afforded a premise for proceeding to the next step of mediation analysis, as suggested by Baron and Kenny (35).

The indirect effect was calculated in the SEM framework (Table 3). A significant indirect effect led from the physical state through emotional health to pain (delta: β indirect = 0.06, $P < 0.05$ [bias-corrected bootstrap (30) 95% CI 0.021–0.374]). This finding indicated that the relationship between the physical state and pain was partially but significantly mediated by emotional health (model 1), with the indirect effect accounting for 10.2% of the total effect of the physical state on pain. The model yielded an R^2 value of 0.428 for the dependent variable (pain), indicating that direct and indirect paths accounted for 43% of the variance of pain. However, this value could not be partitioned. Therefore, we computed R^2 using estimated parameters for the relevant paths. The results indicated that the physical state accounted for 28.9% of the variance of pain, 8.6% of which was attributable to emotional health.

The Wald test (32) was found to be nonsignificant, indicating that there was no significant difference in the moderated indirect effect between groups created by the binary variables (employment status, current psychiatric treatment, any physical complications, or RA treatments such as MTX, DMARDs other than MTX, and biologics) (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23779/abstract).

Table 2. Correlations between observed indicator variables*

	Physical state				Pain		Emotional health	
	ESR 1h	CRP	MMP-3	SJC	TJC	VAS	Anxiety	Depression
ESR 1h	1	–	–	–	–	–	–	–
CRP	0.603†	1	–	–	–	–	–	–
MMP-3	0.345†	0.408†	1	–	–	–	–	–
SJC	0.272†	0.330†	0.325†	1	–	–	–	–
TJC	0.223†	0.291†	0.272†	0.552†	1	–	–	–
Pain VAS	0.207†	0.244†	0.290†	0.371†	0.443†	1	–	–
Anxiety	0.03	0.061	0.073	0.044	0.112‡	0.267†	1	–
Depression	0.097‡	0.102‡	0.166†	0.061	0.130§	0.254†	0.526†	1

* ESR 1h = erythrocyte sedimentation rate 1-hour level; CRP = C-reactive protein; MMP-3 = matrix metalloproteinase 3; SJC = swollen joint count; TJC = tender joint count; VAS = visual analog scale.
† $P < 0.001$.
‡ $P < 0.05$.
§ $P < 0.01$.

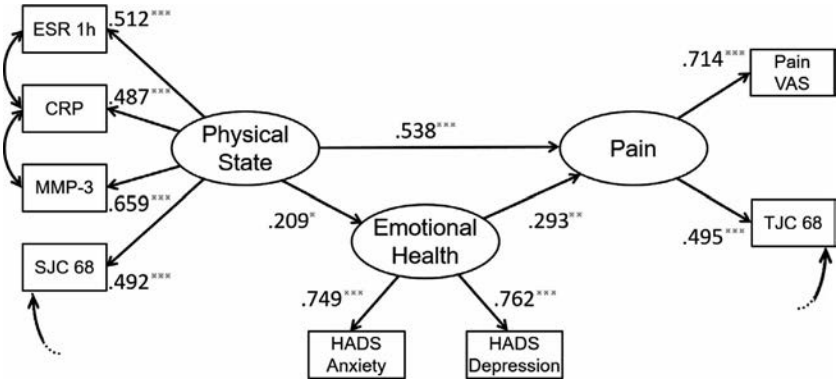

Figure 2. Estimated factor loadings and path coefficients for structural equation modeling, adjusted for age and sex as covariates in the model. All paths displayed in the figure are statistically significant. Curved arrows represent covariances. Curved arrows with a dashed line denote omission in the covariance between swollen joint count (SJC 68) and tender joint count (TJC 68). The fit between the hypothesized model and the data was fairly good: root mean square error of approximation = 0.069 (90% confidence intervals 0.054–0.085), comparative fit index = 0.917, Tucker–Lewis fit index = 0.847, and standardized root mean square residual = 0.040. ESR 1h = erythrocyte sedimentation rate 1-hour level (mm/hour); CRP = C-reactive protein level (mg/dl); MMP-3 = matrix metalloproteinase 3 (ng/ml); pain VAS = visual analog scale for pain; HADS Anxiety = anxiety subscale of the Hospital Anxiety and Depression Scale; HADS Depression = HADS depression subscale. * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$.

We used full information maximum likelihood throughout the SEM analyses, meaning that all available information was used to maximize the estimation of model parameters. This approach is appropriate when missing data occur at random or completely at random. However, there is a possibility of data not missing at random (NMAR). In explicating NMAR, we re-examined the data with a listwise deletion method. This approach with complete cases (n = 481) revealed virtually identical path coefficients, indicating that missing data occurred at random, which refuted any involvement of NMAR: physical state to emotional health ($\beta = 0.25$, $P < 0.05$), emotional health to pain ($\beta = 0.26$, $P < 0.01$), and physical state to pain ($\beta_{direct} = 0.55$, $P < 0.01$; $\beta_{indirect} = 0.07$, $P < 0.05$). Good model fit was retained: RMSEA = 0.066, comparative fit index = 0.923, Tucker-Lewis fit index = 0.858, and SRMR = 0.038.

DISCUSSION

The SEM analyses of patients with RA showed that the physical state significantly affected emotional health. Furthermore, both the physical state and emotional health had significant effects on pain. We found that 10.4% of patients with RA had possible anxiety and 8.3% had probable anxiety. The reported prevalence of anxiety varies across studies. One study using UK and Australian data showed a prevalence of 16.7% possible and 18.6% probable anxiety (2), while a Japanese study showed 7.0% probable anxiety among patients with RA (27). In this study, 17.3% of patients with RA had possible depression, and 12.1% had probable depression, which is similar to the results of a recent meta-analysis that used identical cutoffs (19.4% with possible depression and 14.8% with probable depression) (16).

Table 3. Direct and indirect effects obtained from structural equation modeling*

	Std. est.	SE	*P*
Direct effect			
Physical state → emotional health	0.209	0.097	0.032
Emotional health → pain	0.293	0.088	0.001
Physical state → pain	0.538	0.098	<0.001
Indirect effect			
Physical state → emotional health → pain	0.061	0.028	0.028

* Std. est. = standardized estimate.

The significant effect of emotional health on pain found in this study is consistent with previous reports (7,9). Longitudinal studies have demonstrated a significant effect of anxiety and depression on TJC (an indicator of pain) (7). Another recent study showed that depression affects pain severity (9). These studies suggest that emotional factors may exacerbate pain in patients with RA.

Our finding of a significant influence of emotional health on pain may have clinical implications for current RA treatment. According to the ACR guideline (36), treatment choices depend on RA disease activity. For example, for an individual with RA with low disease activity, DMARD monotherapy is strongly recommended if the patient is naive to DMARDs. For patients with moderate or high disease activity receiving DMARD monotherapy, combination therapy or biologic treatment is recommended. Therefore, disease activity plays a key role in RA treatment decision-making. In current clinical settings, RA disease activity is evaluated by composite indices such as DAS28, SDAI, and CDAI (1,20). However, these composite indices include TJC (an indicator of pain), and thus are potentially influenced by emotional health, as this study illus-

trated. Therefore, assessment of disease activity using these measures may be overestimated if an individual has emotional problems. If some degree of high disease activity is attributable to emotional factors, overtreatment may occur. Consideration of emotional factors in evaluating disease activity may avoid undue RA treatment, which may, in part, lead to a reduction of inflated health care expenditure ascribed to the increased use of biologic drugs (1). Assessing true disease activity targeting for immunosuppressive therapy is strongly required (5,6).

Our findings suggest that RA treatments targeting emotional factors can ameliorate pain. A meta-analysis of randomized controlled trials showed that psychological interventions for RA have positive effects on pain as well as on anxiety and depressive symptoms (37). Furthermore, antidepressants have been shown to relieve pain in individuals with both RA and anxiety/depression (38). These studies and our findings suggest that interventions for emotional health in individuals with RA may substantively alleviate pain, which in turn results in improvement of quality of life (39). Therefore, emphasis should be placed on detailed assessment of emotional health among individuals with RA.

This study has several limitations. First, there might be problems related to the construction of hypothetical models in SEM. For example, we did not incorporate the effect of pain on emotional health in our theory-driven model, because we focused on the effects of the physical state and emotional health on pain, as well as the effect of the physical state on emotional health. Further research using longitudinal data is required to investigate causal relationships between the 3 factors (i.e., physical state, emotional health, and pain). Second, the selection of variables in the hypothetical models had limitations. The estimated effects in this study might be influenced by unknown variables, such as social factors (e.g., socioeconomic factors, lifestyle, degree of social activities, health behaviors, and lack of support from others), although basic characteristics such as age and sex were incorporated in the model as covariates. Third, all participants were from the KURAMA cohort (20), which comprised outpatients at Kyoto University Hospital (located in one of the largest cities in Japan). This cohort includes patients with RA in clinical remission, which might have led to underestimation of the relationship between inflammation and pain. Thus, caution is needed in generalizing our findings to a broader sample of patients with RA. However, the clinical characteristics of participants in this study were similar to those described in other international reports (1,5,7,9,16,33). Finally, psychiatric evaluation was based on self-reported questionnaire responses, which may not objectively reflect the psychological state.

In summary, our study indicates that emotional health partially but significantly mediates the relationship between the physical state and pain in RA, and emotional distress exacerbates pain. Although pain is currently included in RA disease activity measurement, physicians need to be aware of the possibility of overestimated disease activity due to the inclusion of pain components. When ascertaining RA disease activity, emotional factors should be considered to circumvent overestimation of disease activity and unnecessary, excessive treatment.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Takei 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. Nakagami, Sugihara, Takei, Fujii, Hashimoto, Murai, Mimori.

Acquisition of data. Nakagami, Sugihara, Fujii, Hashimoto, Murakami, Furu, Ito, Uda, Torii, Nin, Murai, Mimori.

Analysis and interpretation of data. Nakagami, Sugihara, Takei, Fujii, Hashimoto, Murakami, Murai, Mimori.

REFERENCES

1. Isaacs JD, Moreland LW. Fast facts: rheumatoid arthritis. Oxford: Health Press Limited; 2011.
2. Covic T, Cumming SR, Pallant JF, Manolios N, Emery P, Conaghan PG, et al. Depression and anxiety in patients with rheumatoid arthritis: prevalence rates based on a comparison of the Depression, Anxiety and Stress Scale (DASS) and the Hospital, Anxiety and Depression Scale (HADS). BMC Psychiatry 2012;12:1–10.
3. Heiberg T, Kvien TK. Preferences for improved health examined in 1,024 patients with rheumatoid arthritis: pain has highest priority. Arthritis Rheum 2002;47:391–7.
4. Kaneko Y, Kuwana M, Kondo H, Takeuchi T. Discordance in global assessments between patient and estimator in patients with newly diagnosed rheumatoid arthritis: associations with progressive joint destruction and functional impairment. J Rheumatol 2014;41:1061–6.
5. Ferreira RJ, Duarte C, Ndosi M, de Wit M, Gossec L, da Silva JA. Suppressing inflammation in rheumatoid arthritis: does patient global assessment blur the target? A practice-based call for a paradigm change. Arthritis Care Res (Hoboken) 2018;70:369–78.
6. Lee YC, Cui J, Lu B, Frits ML, Iannaccone CK, Shadick NA, et al. Pain persists in DAS28 rheumatoid arthritis remission but not in ACR/EULAR remission: a longitudinal observational study. Arthritis Res Ther 2011;13:R83.
7. Matcham F, Ali S, Irving K, Hotopf M, Chalder T. Are depression and anxiety associated with disease activity in rheumatoid arthritis? A prospective study. BMC Musculoskelet Disord 2016;17:155.
8. Choy E. Understanding the dynamics: pathways involved in the pathogenesis of rheumatoid arthritis. Rheumatology (Oxford) 2012;51:v3–11.
9. Kojima M, Kojima T, Suzuki S, Oguchi T, Oba M, Tsuchiya H, et al. Depression, inflammation, and pain in patients with rheumatoid arthritis. Arthritis Rheum 2009;61:1018–24.
10. Reichenberg A, Yirmiya R, Schuld A, Kraus T, Haack M, Morag A, et al. Cytokine-associated emotional and cognitive disturbances in humans. Arch Gen Psychiatry 2001;58:445–52.
11. Cooper CM, Godlewska B, Sharpley AL, Barnes E, Cowen PJ, Harmer CJ. Interferon-α induces negative biases in emotional processing in patients with hepatitis C virus infection: a preliminary study. Psychol Med 2018;48:998–1007.
12. Khandaker GM, Dantzer R, Jones PB. Immunopsychiatry: important facts. Psychol Med 2017;47:2229–37.
13. Koyanagi A, Oh H, Stubbs B, Haro JM, DeVylder JE. Epidemiology of depression with psychotic experiences and its association with

chronic physical conditions in 47 low- and middle-income countries. Psychol Med 2016;47:531–42.

14. Cowen PJ. Backing into the future: pharmacological approaches to the management of resistant depression. Psychol Med 2017;47:2569–77.
15. Kopschina Feltes P, Doorduin J, Klein HC, Juarez-Orozco LE, Dierckx RA, Moriguchi-Jeckel CM, et al. Anti-inflammatory treatment for major depressive disorder: implications for patients with an elevated immune profile and non-responders to standard antidepressant therapy. J Psychopharmacol 2017;31:1149–65.
16. Matcham F, Rayner L, Steer S, Hotopf M. The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis. Rheumatology (Oxford) 2013;52:2136–48.
17. Matcham F, Davies R, Hotopf M, Hyrich KL, Norton S, Steer S, et al. The relationship between depression and biologic treatment response in rheumatoid arthritis: an analysis of the British Society for Rheumatology Biologics Register. Rheumatology (Oxford) 2018;57:835–43.
18. Smedstad LM, Vaglum P, Moum T, Kvien TK. The relationship between psychological distress and traditional clinical variables: a 2 year prospective study of 216 patients with early rheumatoid arthritis. Br J Rheumatol 1997;36:1304–11.
19. Schieir O, Thombs BD, Hudson M, Taillefer S, Steele R, Berkson L, et al. Symptoms of depression predict the trajectory of pain among patients with early inflammatory arthritis: a path analysis approach to assessing change. J Rheumatol 2009;36:231–9.
20. Terao C, Hashimoto M, Yamamoto K, Murakami K, Ohmura K, Nakashima R, et al. Three groups in the 28 joints for rheumatoid arthritis synovitis: analysis using more than 17,000 assessments in the KURAMA database. PLoS One 2013;8:e59341.
21. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
22. 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.
23. Posthumus MD, Limburg PC, Westra J, van Leeuwen MA, van Rijswijk MH. Serum matrix metalloproteinase 3 levels in comparison to C-reactive protein in periods with and without progression of radiological damage in patients with early rheumatoid arthritis. Clin Exp Rheumatol 2003;21:465–72.
24. Ribbens C, Andre B, Jaspar JM, Kaye O, Kaiser MJ, De Groote D, et al. Matrix metalloproteinase-3 serum levels are correlated with disease activity and predict clinical response in rheumatoid arthritis. J Rheumatol 2000;27:888–93.
25. Houseman M, Potter C, Marshall N, Lakey R, Cawston T, Griffiths I, et al. Baseline serum MMP-3 levels in patients with rheumatoid arthritis are still independently predictive of radiographic progression in a longitudinal observational cohort at 8 years follow up. Arthritis Res Ther 2012;14:R30.
26. 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.
27. Katayama M, Miyamura T, Suenaga Y, Suematsu E, Urata Y, Matsui T, et al. AB0365 prevalence and factors associated with depression and anxiety in patients with rheumatoid arthritis using data from a large Japanese cohort database in 2013 (Ninja 2013 Database) [abstract]. Ann Rheum Dis 2015;74:1015.
28. Nicassio PM, Ormseth SR, Custodio MK, Irwin MR, Olmstead R, Weisman MH. A multidimensional model of fatigue in patients with rheumatoid arthritis. J Rheumatol 2012;39:1807–13.
29. Somers TJ, Shelby RA, Keefe FJ, Godiwala N, Lumley MA, Mosley-Williams A, et al. Disease severity and domain-specific arthritis self-efficacy: relationships to pain and functioning in patients with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2010;62:848–56.
30. MacKinnon DP, Lockwood CM, Williams J. Confidence limits for the indirect effect: distribution of the product and resampling methods. Multivariate Behav Res 2004;39:99.
31. Sobel ME. Asymptotic confidence intervals for indirect effects in structural equation models. Sociol Methodol 1982;13:290–312.
32. Hoyle RH. Handbook of structural equation modeling. New York: Guilford Publications; 2012.
33. Miyamoto S, Sugimura Y, Kasiwagura T, Urayama M, Kobayashi M, Sakuraba T, et al. Evaluation for treatment of rheumatoid arthritis patients in AORA cohort. Clin Rheumatol 2015;27:135–45. In Japanese.
34. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale (NJ): L. Erlbaum Associates; 1988.
35. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986;51:1173–82.
36. Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheum 2016;68:1–26.
37. Knittle K, Maes S, de Gucht V. Psychological interventions for rheumatoid arthritis: examining the role of self-regulation with a systematic review and meta-analysis of randomized controlled trials. Arthritis Care Res (Hoboken) 2010;62:1460–72.
38. Ash G, Dickens CM, Creed FH, Jayson MI, Tomenson B. The effects of dothiepin on subjects with rheumatoid arthritis and depression. Rheumatology (Oxford) 1999;38:959–67.
39. Garip Y, Eser F, Bodur H. Health-related quality of life in rheumatoid arthritis: comparison of RAQoL with other scales in terms of disease activity, severity of pain, and functional status. Rheumatol Int 2011;31:769–72.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1224–1233
DOI 10.1002/acr.23843

Timing of Abatacept Before Elective Arthroplasty and Risk of Postoperative Outcomes

Michael D. George,[1] Joshua F. Baker,[2] Kevin Winthrop,[3] Evo Alemao,[4] Lang Chen,[5] Sean Connolly,[6] Jesse Y. Hsu,[1] Teresa A. Simon,[4] Qufei Wu,[1] Fenglong Xie,[5] Shuo Yang,[5] and Jeffrey R. Curtis[5]

Objective. Guidelines recommend withholding biologic therapies before hip and knee arthroplasty, yet evidence to inform optimal timing is limited. The aim of this study was to determine whether withholding abatacept infusions is associated with lower risk of adverse postoperative outcomes.

Methods. This retrospective cohort study, which used US Medicare and Truven MarketScan administrative data from January 2006 to September 2015, evaluated adults with rheumatoid arthritis who received intravenous abatacept (precisely dated in claims data) within 6 months of elective primary or revision hip or knee arthroplasty. Propensity weighted analyses using inverse probability weights compared the risk of 30-day hospitalized infection and 1-year prosthetic joint infection (PJI) between patients with different abatacept stop timing (time between last infusion and surgery). Secondary analyses evaluated nonurinary hospitalized infections and 30-day readmissions.

Results. After 1,939 surgeries among 1,780 patients, there were 175 hospitalized infections (9.0%), 115 nonurinary hospitalized infections (5.9%), 39 PJIs (2.4/100 person-years), and 114/1,815 30-day readmissions (6.3%). There were no significant differences in outcomes with abatacept stop timing <4 weeks (1 dosing interval) versus 4–8 weeks (hospitalized infection odds ratio [OR] 0.93 [95% confidence interval (95% CI) 0.65–1.34]; nonurinary hospitalized infection OR 0.93 [95% CI 0.60–1.44]; PJI hazard ratio 1.29 [95% CI 0.62–2.69]; 30-day readmission OR 1.00 [95% CI 0.65–1.54]). Similarly, there were no significant differences in outcomes with abatacept stop timing <4 weeks versus ≥8 weeks. Glucocorticoid use >7.5 mg/day was associated with greater risk of hospitalized infection (OR 2.19 [95% CI 1.28–3.77]) and nonurinary hospitalized infection (OR 2.38 [95% CI 1.22–4.64]).

Conclusion. Compared to continuing intravenous abatacept, withholding abatacept for ≥4 weeks (one dosing interval) before surgery was not associated with a lower risk of hospitalized infection, nonurinary hospitalized infection, PJI, or 30-day readmission.

INTRODUCTION

Orthopedic surgery, especially hip and knee arthroplasty, remains common in patients with rheumatoid arthritis (RA) (1). These surgeries can be complicated by both immediate postoperative infections (e.g., pneumonia, soft tissue infections) and by prosthetic joint infections occurring later after surgery (2). Patients with RA are at increased risk of postoperative infection, with potential contributors including immunosuppression, disease activity, and comorbidities (3,4). The optimal approach to managing immunosuppression before surgery, a modifiable risk factor, has been of particular interest.

Supported by Bristol-Myers Squibb.

Dr. George's work was supported by the Rheumatology Research Foundation Scientist Development Award and the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant 1K23-AR-073931-01). Dr. Curtis' work was supported by the Patient Centered Outcomes Research Institute (grant PPRND-1507-32163).

[1]Michael D. George, MD, MSCE, Jesse Y. Hsu, PhD, Qufei Wu, MS: University of Pennsylvania Perelman School of Medicine, Philadelphia; [2]Joshua F. Baker, MD, MSCE: Philadelphia Veterans Affairs Medical Center, Philadelphia, Pennyslvania; [3]Kevin Winthrop, MD, MPH: Oregon Health & Science University, Portland; [4]Evo Alemao, RPh, PhD, Teresa A. Simon, MPH: Bristol-Myers Squibb, Princeton, New Jersey; [5]Lang Chen, PhD, Fenglong Xie, MS, Shuo Yang, MS, Jeffrey R. Curtis, MD, MS, MPH: University of Alabama at Birmingham; [6]Sean Connolly, PhD: Bristol-Myers Squibb, Plainsboro, New Jersey.

Dr. George has received consulting fees from AbbVie (less than $10,000) and research support from Bristol-Myers Squibb. Dr. Baker has received consulting fees from Bristol-Myers Squibb and Corrona (less than $10,000 each). Dr. Winthrop has received consulting fees from Galapagos, Roche, and Gilead (less than $10,000 each) and from Bristol-Myers Squibb, UCB, Eli Lilly, GlaxoSmithKline, AbbVie, and Pfizer (more than $10,000 each). Dr. Alemao, Dr. Connolly, and Ms Simon own stock or stock options in Bristol-Myers Squibb. Dr. Curtis has received consulting fees from AbbVie, Bristol-Myers Squibb, Eli Lilly, Myriad, Roche/Genentech, and UCB (less than $10,000 each) and from Amgen, Janssen, Pfizer, and Corrona (more than $10,000 each) and research support from Bristol-Myers Squibb.

No other disclosures relevant to this article were reported.

Address correspondence to Jeffrey R. Curtis, MD, MS, MPH, 510 20th Street South Birmingham, AL 35294. E-mail: jrcurtis@uabmc.edu.

Submitted for publication August 13, 2018; accepted in revised form February 5, 2019.

SIGNIFICANCE & INNOVATIONS

- Receiving intravenous abatacept within 4 weeks (one dosing interval) before elective hip or knee replacement was not associated with a greater risk of postoperative infection.
- Glucocorticoids were associated with increased risk of postoperative infection.
- Withholding abatacept may not improve outcomes, particularly if glucocorticoids must be increased to treat disease flares.

Recent guidelines from the American College of Rheumatology (ACR) and American Association of Hip and Knee Surgeons recommend continuing conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), such as methotrexate, throughout the perioperative period (5). Given minimal data to guide optimal timing of biologic DMARDS (bDMARDs), however, these guidelines recommended withholding bDMARDs for one dosing interval before surgery (e.g., 4 weeks for intravenous abatacept). It remains unknown whether withholding therapy prior to surgery reduces risks and, if so, what the optimal timing should be. Our previous observational study suggested that withholding infliximab before surgery was not associated with a reduced risk of postoperative infection (6), but it is unclear whether these results are applicable to other biologic therapies with different mechanisms of action, dosing intervals, and pharmacodynamics.

In the current study, we combined 2 large administrative data sets to evaluate whether the timing of abatacept infusions before elective hip or knee arthroplasty was associated with the risk of postoperative infection (i.e., determining whether withholding intravenous abatacept reduces the risk of postoperative infection). We specifically evaluated abatacept infusions, as these can be precisely dated in claims data.

PATIENTS AND METHODS

This retrospective cohort study evaluated patients with RA undergoing hip or knee arthroplasty using US Medicare claims and Truven MarketScan databases from January 1, 2006 to September 30, 2015. Medicare is a public health plan covering more than 90% of US adults ages ≥65 (7). Younger individuals with certain disabilities (e.g., RA) may also be covered. MarketScan is a US database including inpatient, outpatient, and pharmacy data contributed by large employers, health plans, and government and public organizations for more than 143 million individuals (8).

Cohort identification. Included patients were ages ≥18 years with RA, based on 2 physician office or inpatient International Classification of Diseases, Ninth Revision (ICD-9) codes (714.xx) at least 7 days apart and use of a DMARD (9), who underwent inpatient elective primary or revision hip or knee arthroplasty (requiring ICD-9 and Current Procedural Terminology [CPT] codes for primary surgeries and CPT codes for revisions according to validated algorithms [see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract]) (10–13). In order to identify patients receiving stable long-term therapy, only those patients who had received intravenous abatacept within 6 months of surgery and ≥3 infusions in the past year were included (see Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract). We evaluated abatacept infusions and not subcutaneous abatacept, because infusions are coded as procedures (Healthcare Common Procedure Coding System codes C9230 or J0129) and can be precisely dated in claims data. We required continuous enrollment in MarketScan or in Medicare Part A, B, and D without enrollment in a Medicare Advantage Plan for ≥1 year prior to surgery ("baseline") to allow uniform assessment of covariates.

In order to accurately assess risk of postoperative infection, we excluded patients with evidence of possible preexisting infection or nonelective surgery, which included those who had a diagnosis or treatment for native or prosthetic joint infection (PJI) in the year prior to surgery, diagnoses of femur fracture, bone or metastatic cancer from the index hospitalization, admission through the emergency department or transfer from another acute care hospital, surgery after hospital day 3, or major surgery in the previous 6 months (see Supplementary Table 1, available on the Arthritis Care & Research web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract). Patients with admission status "emergent" were excluded in Medicare (not available in MarketScan). Patients with inflammatory bowel disease, psoriatic arthritis, ankylosing spondylitis, HIV, or active malignancy in the past year were excluded. Patients could contribute multiple surgeries if >6 months apart. In order to avoid including any possible duplicate observations, we excluded patients in MarketScan with derived birth dates within 31 days of any patient in Medicare with the same admission date.

Exposure. The exposure of interest was the abatacept stop timing, which was the time between the last abatacept infusion and surgery (see Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract). Stop timing was categorized in 4-week intervals (<4 weeks, 4 to <8 weeks, ≥8 weeks) based on the usual dosing interval of intravenous abatacept. Stop timing <4 weeks was the reference group, and we evaluated whether stopping abatacept for at least 1 dosing interval (4–8 weeks), or 2 dosing intervals (≥8 weeks) was associated with a lower risk of adverse postoperative outcomes. In a sensitivity analysis, we also evaluated abatacept stop timing in 2-week intervals, with stop timing of 2–4 weeks as the reference group in this

analysis because of the small number of patients with stop timing <2 weeks.

Outcomes. The 2 prespecified primary outcomes were hospitalized (serious) infection within 30 days and PJI within 1 year after surgery. Hospitalized infections were identified based on ICD-9 diagnosis codes from any position of the discharge diagnoses as previously validated with positive predictive value >80% (14,15), including the index hospitalization and any subsequent acute care hospitalizations with admission date within 30 days of surgery (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract). PJI was identified based on inpatient or outpatient physician diagnosis (ICD-9 code 996.66), excluding diagnoses from the index hospitalization (an exclusion criteria because these diagnoses may represent preexisting infections) (16,17). Sensitivity analyses assessed more stringent definitions of PJI, including the requirement of an inpatient diagnosis or an accompanying procedure code within 30 days of PJI diagnosis (i.e., arthrotomy, prosthesis removal, central venous catheter insertion, spacer, or revision surgery), similar to previously used definitions (11) (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract).

We evaluated 2 secondary outcomes. Because urinary tract infections were expected to represent a high proportion of postoperative infections and may at times represent more minor or incidentally discovered infections (e.g., related to a urinary catheter), we evaluated an alternate 30-day hospitalized infection outcome that ignored urinary tract infections. Additionally, we assessed 30-day readmission (a surrogate for complications and an important health services outcome) among patients with a discharge disposition to home, home health care, acute rehabilitation, or a skilled nursing facility (not including admissions within 1 day of discharge or with primary diagnosis indicating rehabilitation) (18).

Exploratory outcomes included prolonged length of stay (a surrogate for postoperative complications) and time to revision surgery among patients undergoing primary knee or hip arthroplasty. Prolonged length of stay was defined as length of stay >90th percentile (empirically derived in the data as >5 days for revision surgeries and >4 days for primary surgeries) (19,20). We also examined wound complications (21) and rates of specific postoperative infections, including pneumonia, septicemia/bacteremia, and urinary infection (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract).

Covariates. Covariates measured during the 365-day baseline period included demographics, comorbid conditions, an adaptation of the Charlson comorbidity index (22), health care utilization (outpatient visits, emergency department visits, hospitalizations), and hospitalized infections. The number of previous biologics was assessed with all available data. We also evaluated the prescriptions filled for nonsteroidal antiinflammatory drugs, opioids, csDMARDs, and antibiotics in the 90 days before surgery. Average glucocorticoid dose in the 90 days prior to surgery was calculated based on oral prescriptions for prednisone, prednisolone, and methylprednisolone, using prescribed dose in prednisone equivalents and days supply to determine each daily dose and truncating prescriptions if a new prescription was filled before the prescription end date. Stress dose or other intravenous glucocorticoids were not included. Disability status, skilled nursing facility residence, and quintiles of median household income based on zip code from the American Community Survey 2009–2013 were available only in Medicare (23). Surgeon and hospital volume (Medicare only) were estimated using a larger cohort of 55,812 hip or knee arthroplasty procedures among patients with any bDMARD or methotrexate use within 6 months of surgery.

Statistical analysis. Associations between abatacept stop timing and postoperative outcomes were assessed using logistic regression for binary outcomes and cause-specific proportional hazards regression for PJI, censoring at the soonest of 1 year, end or interruption of enrollment, subsequent hip or knee arthroplasty, or September 30, 2015 and treating death as censoring (<1% of censoring events) (24). Propensity score–derived inverse probability weights were used to balance confounders across treatment groups (as described below). Potential heterogeneity in associations between exposure and outcome in Medicare versus MarketScan was evaluated using data set/exposure interaction terms. Log-log plots revealed no violation of proportional hazards.

Propensity scores based on the probability of being in each stop timing group were generated using multinomial logistic regression models. In the primary analyses, propensity score models included a data set variable (Medicare versus MarketScan), covariates of interest common to both data sets, and a squared term for age to account for nonlinearity (6,25,26) (see Supplementary Table 3, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract). In a sensitivity analysis restricted only to Medicare data, propensity scores were recalculated with additional variables, available only in the Medicare data, and analysis was repeated. Propensity scores were used to create stabilized inverse probability treatment weights (26–28) truncated at the 1st and 99th percentile (29). The balance of covariates across treatment categories was assessed using standardized mean differences (SMD) compared to the reference group (stop timing <4 weeks) with SMD ≤0.1 indicating good balance (balance before and after weighting for primary analyses are shown in Supplementary Table 3 and Supplementary Figure 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract). In cases of inadequate balance (only present in sensitivity analyses evaluating abatacept stop timing in 2-week intervals), unbalanced variables were added as covariates to the weighted outcome models.

Table 1. Select patient characteristics by abatacept stop timing*

	Abatacept timing					Patient total
	<4 weeks (n = 732)	4 to <8 weeks (n = 862)	SMD vs. <4 weeks†	≥8 weeks (n = 345)	SMD vs. <4 weeks†	(n = 1,939)
Medicare cohort	594 (81.1)	670 (77.7)	−0.085	273 (79.1)	−0.051	1,537 (79.3)
Female	628 (85.8)	741 (86.0)	0.005	303 (87.8)	0.060	1,672 (86.2)
Age, mean ± SD years	67.5 ± 10.0	66.5 ± 10.7	−0.100	67.3 ± 9.7	−0.024	67.0 ± 10.3
Year	2011 [2009–2013]	2012 [2010–2013]	0.130	2012 [2010–2013]	0.112	2012 [2010–2013]
Surgery type						
Primary knee	489 (66.8)	547 (63.5)	−0.07	208 (60.3)	−0.135‡	1,244 (64.2)
Primary hip	180 (24.6)	226 (26.2)	0.037	98 (28.4)	0.086	504 (26.0)
Revision knee	32 (4.4)	50 (5.8)	0.065	26 (7.5)	0.134‡	108 (5.6)
Revision hip	31 (4.2)	39 (4.5)	0.014	13 (3.8)	−0.024	83 (4.3)
Average glucocorticoid dose						
None	363 (49.6)	445 (51.6)	0.041	163 (47.2)	−0.047	971 (50.1)
≤5 mg	218 (29.8)	241 (28.0)	−0.040	111 (32.2)	0.052	570 (29.4)
5–10 mg	114 (15.6)	139 (16.1)	0.015	53 (15.4)	−0.006	306 (15.8)
>10 mg	37 (5.0)	37 (4.3)	−0.036	18 (5.2)	0.007	92 (4.7)
Prior biologics						
0	338 (46.2)	375 (43.5)	−0.054	135 (39.1)	−0.143	848 (43.7)
1	287 (39.2)	365 (42.3)	0.064	151 (43.8)	0.093	803 (41.4)
2	107 (14.6)	122 (14.2)	−0.013	59 (17.1)	0.068	288 (14.9)
MTX	338 (46.2)	375 (43.5)	0.003	135 (39.1)	−0.187	848 (43.7)
HCQ, SSZ, or LEF	201 (27.5)	222 (25.8)	−0.039	94 (27.2)	−0.005	517 (26.7)
NSAID past 90 days	256 (35.0)	279 (32.4)	−0.055	112 (32.5)	−0.053	647 (33.4)
Opioid past 90 days	442 (60.4)	511 (59.3)	−0.022	253 (73.3)	0.277	1,206 (62.2)
Charlson morbidity score, median (IQR)	1 (0,3)	1 (0,3)	−0.040	1 (0,3)	0.120	1 (0,3)
Diabetes mellitus	140 (19.1)	135 (15.7)	−0.091	69 (20.0)	0.022	344 (17.7)
Hypertension	400 (54.6)	459 (53.2)	−0.028	209 (60.6)	0.120	1,068 (55.1)
Congestive heart failure	49 (6.7)	52 (6.0)	−0.027	24 (7.0)	0.010	125 (6.4)
COPD/asthma	111 (15.2)	113 (13.1)	−0.059	50 (14.5)	−0.019	274 (14.1)
Chronic kidney disease	40 (5.5)	49 (5.7)	0.010	29 (8.4)	0.116	118 (6.1)
Obesity	87 (11.9)	92 (10.7)	−0.038	38 (11.0)	−0.027	217 (11.2)
Hospitalizations past year						
0	552 (75.4)	641 (74.4)	−0.024	223 (64.6)	−0.237	1,416 (73.0)
1–2	124 (16.9)	151 (17.5)	0.015	78 (22.6)	0.143	353 (18.2)
≥3	56 (7.7)	70 (8.1)	0.017	44 (12.8)	0.169	170 (8.8)
Hospitalized infection past year	46 (6.3)	63 (7.3)	0.041	30 (8.7)	0.092	139 (7.2)
ED visits past year						
0	492 (67.2)	565 (65.5)	−0.035	211 (61.2)	−0.126	1,268 (65.4)
1	144 (19.7)	203 (23.5)	0.094	77 (22.3)	0.065	424 (21.9)
2–3	73 (10.0)	74 (8.6)	−0.048	42 (12.2)	0.07	189 (9.7)
>3	23 (3.1)	20 (2.3)	−0.05	15 (4.3)	0.063	58 (3.0)
Antibiotic use past 3 months	374 (51.1)	470 (54.5)	0.069	223 (64.6)	0.277	1,067 (55.0)

* Values are the number (%) unless indicated otherwise. Standardized mean difference (SMD) before inverse probability weighting with values >0.1 considered relevant, indicating imbalance between groups. SMD before and after weighting for full set of covariates available in Supplementary Table 3. MTX = methotrexate; HCQ = hydroxychloroquine; SSZ = sulfasalazine; LEF = leflunomide; NSAID = nonsteroidal antiinflammatory drug; IQR = interquartile range; COPD = chronic obstructive pulmonary disease; ED = emergency department.

Table 2. Unadjusted frequencies of primary and secondary outcomes by abatacept stop timing*

Abatacept stop timing	No.	PYs (for PJI)	Hospitalized infection	PJI, no. (incidence/ 100 PYs)	Nonurinary hospitalized infections†	30-day readmission
<4 weeks	732	616	67 (9.1)	13 (2.1)	45 (6.2)	46/697 (6.6)
4 to <8 weeks	862	715	67 (7.8)	18 (2.5)	45 (5.2)	49/805 (6.1)
≥8 weeks	345	283	41 (11.9)	8 (2.8)	25 (7.2)	19/313 (6.1)
Total	1,939	1,614	175 (9.0)	39 (2.4)	115 (5.9)	114/1,815 (6.3)

* Values are the number (%) unless indicated otherwise. Person-years (PYS) shown are for assessment of prosthetic joint infection (PJI) within 1 year of surgery. Thirty-day readmission analyses restricted to patients discharged to home, home health, skilled nursing facility, or inpatient rehabilitation facility (denominators indicate number of eligible surgeries).
† Hospitalized infections excluding urinary tract infections.

In order to assess our ability to detect important differences in our outcomes (i.e., a "positive control"), we performed similar inverse probability weighted analyses to evaluate associations between glucocorticoid dose (none, ≤7.5 mg, >7.5 mg) and outcomes in the same cohort, with weights derived from separate multinomial logistic regression propensity score models based on the probability of being in each glucocorticoid treatment group, including the same covariates as well as abatacept stop timing.

Sensitivity analyses. We repeated analyses with Medicare-only data and also evaluated 2 alternative definitions of PJI as noted above. Additionally, we repeated analyses restricted to primary joint replacement surgeries only. Finally, we repeated analyses with abatacept stop timing categorized in 2-week intervals.

The data set was created with SAS, version 9.4 and analysis was performed using STATA, version 13.1. The protocol was approved by the University of Alabama at Birmingham Institutional Review Board. Use of the data was governed by data use agreements from the Centers for Medicare and Medicaid Services and Truven MarketScan.

RESULTS

Cohort identification and characteristics. We identified 3,984 hip or knee arthroplasties in Medicare and 901 in MarketScan among patients with RA with an abatacept infusion within 6 months before surgery. After applying exclusion criteria, 1,537 Medicare (38.6%) and 421 MarketScan (46.7%) surgeries remained (see Supplementary Table 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract). We excluded 19 surgeries from the MarketScan cohort that were possible duplicates in the 2 data sets, leaving 1,939 surgeries among 1,780 patients for final analysis.

Cohort characteristics are shown in Table 1 and Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract. Medicare patients comprised 79.3% of the cohort, with a mean age of 67.0 years; 86.2% of the patients were women and 90.1% of surgeries were primary hip or knee arthroplasty. Abatacept stop timing was <4 weeks (within 1 dosing interval) in 732 patient surgeries (37.8%), 4 to <8 weeks in 862 (44.5%), and ≥8 weeks in 345 (17.8%). While patients with stop timing <4 weeks versus those with 4–8 weeks were similar, patients with stop timing ≥8 weeks were less likely to have received prior biologic therapy or be treated with methotrexate, more likely to be treated with opioids or antibiotics in the past 3 months, and more likely to have been hospitalized or have an emergency department visit in the past year. The interval between the 2 most recent abatacept infusions was ≤28 days for the majority of patients in all groups, and was ≥35 days in 15.3% of patients with stop timing <4 weeks, 20.5% in those with stop timing of 4–8 weeks, and 27.0% in those with stop timing of ≥8 weeks.

Primary and secondary outcomes. We identified 175 hospitalized infections (9.0%) within 30 days of surgery, with 101 infections from the index hospitalization (57.7%) (Table 2). The most common infections were urinary (103

Table 3. Frequency of specific postoperative outcomes*

Abatacept stop timing	No.	Urinary infection	Pneumonia	Septicemia/ bacteremia	Wound complication
<4 weeks	732	38 (5.2)	5 (0.7)	4 (0.6)	9 (1.2)
4 to <8 weeks	862	43 (5.0)	12 (1.4)	2 (0.2)	7 (0.8)
≥8 weeks	345	22 (6.4)	5 (1.4)	1 (0.3)	4 (1.2)
Total	1,939	103 (5.3)	22 (1.1)	7 (0.4)	20 (1.0)

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

Table 4. Association between abatacept stop timing and primary and secondary outcomes from unadjusted and inverse probability weighted analyses (shown graphically in Figure 1)*

	Hospitalized infection	PJI, HR (95% CI)	Nonurinary hospitalized infection	30-day readmission
Stop timing, unadj. results				
<4 weeks	1	1	1	1
4 to <8 weeks	0.84 (0.59–1.19)	1.18 (0.58–2.41)	0.84 (0.55–1.29)	0.92 (0.61–1.39)
≥8 weeks	1.34 (0.89–2.02)	1.32 (0.55–3.19)	1.19 (0.72–1.98)	0.91 (0.53–1.59)
Stop timing, adj. results†				
<4 weeks	1	1	1	1
4 to <8 weeks	0.93 (0.65–1.34)	1.29 (0.62–2.69)	0.93 (0.60–1.44)	1.00 (0.65–1.54)
≥8 weeks	1.13 (0.73–1.77)	1.20 (0.48–3.05)	0.97 (0.57–1.66)	0.82 (0.46–1.47)

* Values are the odds ratio (95% confidence interval [95% CI]) unless indicated otherwise. Odds ratio and hazard ratio (HR) from logistic regression and cause-specific hazards regression. PJI = prosthetic joint infection; unadj. = unadjusted; adj. = adjusted.
† Adjusted results from inverse probability weighted analyses.

patients [5.3%]), skin and soft tissue (22 patients [1.1%]), and pneumonia (22 patients [1.1%]) (Table 3). Rates of hospitalized infection by abatacept stop timing ranged from 7.8% (4–8 weeks) to 11.9% (≥8 weeks). In propensity weighted analyses that controlled for between-group differences, there were no significant differences in hospitalized infection for longer abatacept stop timing compared to <4 weeks (4–8 weeks OR 0.93 [95% CI 0.65–1.34]; ≥8 weeks OR 1.13 [95% CI 0.73–1.77]) (Table 4, Figure 1). Results were similar in sensitivity analyses only evaluating the Medicare cohort with additional variables included in the propensity score (see Supplementary Tables 3 and 4, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract) and in analyses only including primary hip or knee arthroplasty (see Supplementary Tables 5–6, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract).

There were 39 PJIs within 1 year of surgery (2.4/100 person-years) with 25 infections (64%) occurring within 90

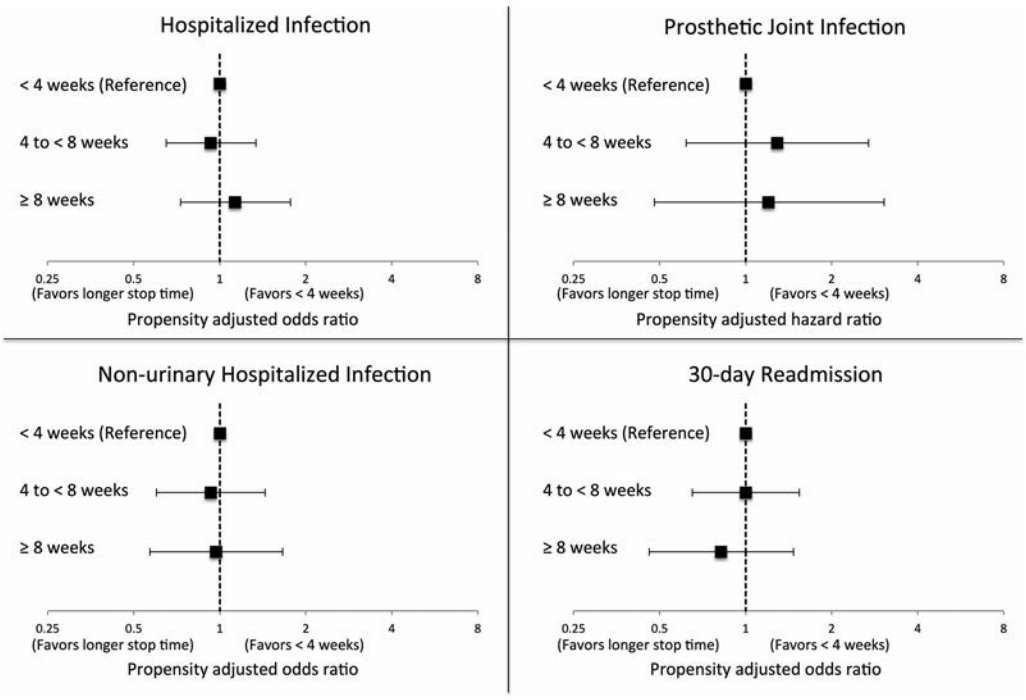

Figure 1. Association between abatacept stop timing and postoperative outcomes from inverse probability weighted analyses. Odds ratios and hazard ratios from logistic and cause-specific hazards regression analyses weighted with inverse probability weights to account for confounders.

days of surgery (Table 2). Crude incidence of PJI ranged from 2.1 (<4 weeks) to 2.8/100 person-years (≥8 weeks), with no significant difference in PJI in propensity adjusted analyses for longer stop timing compared to <4 weeks (4–8 weeks hazard ratio [HR] 1.29 (95% CI 0.62–2.69), ≥8 weeks HR 1.20 (95% CI 0.48–3.05)] (Table 4, Figure 1). Results were similar using 2 alternative definitions of PJI and in sensitivity analyses restricted to the Medicare cohort or evaluating only primary surgeries (see Supplementary Tables 5–7, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract).

Nonurinary hospitalized infection occurred after 115 surgeries (5.9%), and 30-day readmission occurred after 114/1,815 surgeries (6.3%). There were no significant differences in either of these outcomes by abatacept stop timing (Table 4, Figure 1). Results were similar in analyses limited to the Medicare cohort or limited to primary surgeries (see Supplementary Tables 5–6, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract).

Exploratory outcomes. The risk of prolonged length of stay was significantly greater in patients with abatacept stop timing of 4–8 weeks (OR 1.74 [95% CI 1.17–2.58]) or ≥8 weeks (OR 2.26 [95% CI 1.41–3.62]) versus <4 weeks in propensity-weighted analyses (Table 5). Among 1,746 primary hip or knee arthroplasties with a mean follow-up time of 2.5 years, we identified 54 revision surgeries (1.2/100 person-years), with no significant difference in rates of revision by abatacept stop timing (Table 5).

While patients with different abatacept stop timing had similar rates of oral glucocorticoid use prior to surgery (50.4%, 48.4%, and 52.8% for abatacept stop timing <4 weeks, 4–8 weeks, and ≥8 weeks, respectively), patients with stop timing ≥8 weeks were the most likely to fill a glucocorticoid prescription in the 90 days after surgery (42.5%, 42.0%, 49.6% for abatacept stop timing <4 weeks, 4–8 weeks, and ≥8 weeks, respectively; $P = 0.01$ for <8 weeks versus ≥8 weeks). Patients with abatacept stop timing ≥8 weeks were more likely to receive opioids before surgery (60.4%, 59.3%, 73.3% for stop timing <4 weeks, 4–8 weeks, and ≥8 weeks, respectively; $P < 0.001$). These patients were also less likely to restart abatacept after surgery (92.9%, 89.6%, 55.9% for stop timing <4 weeks, 4–8 weeks, and ≥8 weeks, respectively; $P < 0.001$).

Abatacept stop timing in 2-week intervals. Precision was more limited in sensitivity analyses categorizing abatacept stop timing in 2-week intervals. We found that patients with stop timing of 2–4 weeks had similar rates of hospitalized infection, PJI, nonurinary hospitalized infection, and 30-day readmission versus patients with longer stop-timing in propensity-weighted analyses (see Supplementary Table 8 and Supplemental Figure 3, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract). Only 177 patients (9.1% overall) received abatacept within 2 weeks of surgery. Although risks of the outcomes were not significantly different with stop timing <2 weeks versus 2–4 weeks in propensity-weighted analyses, patients with abatacept stop timing <2 weeks did have numerically greater risk of hospitalized infection (OR 1.63 [95% CI 0.91–2.91]), PJI (HR 1.48 [95% CI 0.45–4.93]), nonurinary hospitalized infection (OR 1.45 [95% CI 0.70–2.99]), and 30-day readmission (OR 1.23 [95% CI 0.60–2.51]) versus stop timing of 2–4 weeks (see Supplementary Table 8 and Supplementary Figure 3, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract).

Associations between glucocorticoids and outcomes. In propensity-weighted analyses, we found that an average glucocorticoid dosage of >7.5 mg/day in the 3 months prior to surgery versus no glucocorticoids was associated with a greater risk of hospitalized infection (OR 2.19 [95% CI 1.28–3.77]), nonurinary hospitalized infection (OR 2.38 [95% CI 1.22–4.64]), and a numerically greater risk of PJI (HR 2.13 [95% CI 0.78–5.79]), 30-day readmission (OR 1.52 [95% CI 0.78–2.98]), and prolonged length of stay (OR 1.70 [95% CI 0.99–2.90]) (see Supplementary Tables 9 and 10 and

Table 5. Frequency and results of inverse probability weighted analyses for prolonged length of stay and time to revision surgery*

Abatacept stop timing	No.	Prolonged length of stay	Prolonged length of stay aOR (95% CI)†	PYs	Revision, no. (incidence/100 PYs)‡	Revision, aHR (95% CI)†
<4 weeks	732	43 (5.9)	Ref.	1,753	18 (1.0)	Ref.
4 to <8 weeks	862	79 (9.2)	1.74 (1.17–2.58)	1,870	27 (1.4)	1.30 (0.70–2.41)
≥8 weeks	345	45 (13.0)	2.26 (1.41–3.62)	760	9 (1.2)	0.83 (0.35–1.95)
Total	1,939	167 (8.6)	N/A	4,382	54 (1.2)	N/A

* Values are the number (%) unless indicated otherwise. aOR = adjusted odds ratio; 95% CI = 95% confidence interval; PYs = person-years; aHR = adjusted hazard ratio; Ref. = reference; N/A = not applicable.
† From logistic regression and cause-specific hazards regression with inverse probability weighting.
‡ Shown for revision analyses, restricted to patients with primary hip or knee arthroplasty and censored if a subsequent hip or knee arthroplasty is performed.

Supplementary Figure 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23843/abstract).

DISCUSSION

In this study evaluating a cohort of patients with RA treated with intravenous abatacept before elective hip or knee arthroplasty, we found that patients who received abatacept within 1 dosing interval (4 weeks) of surgery had similar rates of postoperative hospitalized infection, prosthetic joint infection, and 30-day readmission compared to patients with longer stop timing. These results suggest that withholding abatacept for at least 1 dosing interval prior to surgery (i.e., >4 weeks or >8 weeks), while exposing patients to a risk of disease flares, may not reduce the risk of short-term hospitalized infection or subsequent PJI.

Little data exists to inform how biologic therapies should be managed in the perioperative period. Current guidelines from the ACR recommend continuing nonbiologic DMARDs (such as methotrexate) throughout the perioperative period but withholding biologics for 1 dosing interval (i.e., 4 weeks for intravenous abatacept), although these recommendations are based on limited and indirect data from studies of biologics in the nonsurgical setting (5). After the publication of these guidelines, we evaluated the timing of infliximab before hip or knee arthroplasty in a large Medicare cohort and found that rates of postoperative infection were similar in patients who had received infliximab 8–12 weeks prior to surgery (at least 1 dosing interval) compared to those with shorter stop timing of 4–8 weeks (adjusted OR 0.95 [95% CI 0.62–1.36]) or <4 weeks (adjusted OR 0.90 [95% CI 0.60–1.34]) (6). At that time, however, there was not enough data to evaluate other infusion biologic therapies.

The current study provides new data on associations between the timing of abatacept before hip and knee arthroplasty and postoperative outcomes. In this study, we specifically evaluated patients receiving intravenous abatacept because, unlike with prescription drug data from pharmacy claims, the date a patient last received an infusion medication can be precisely determined based on the infusion date. Importantly, we did not compare those being treated with abatacept to those who are not being treated with abatacept because these patients may be quite different. Rather, we compared patients with different abatacept stop timing before surgery. In support of this approach, patients who received abatacept within 1 dosing interval (4 weeks) of surgery were quite similar to those who had abatacept held for at least 1 dosing interval (4–8 weeks), even before propensity score adjustment.

We found that patients who received an infusion within 4 weeks of surgery had similar rates of hospitalized infection, PJI, nonurinary hospitalized infection, and 30-day readmission compared to those who held therapy for 1 dosing interval (4–8 weeks) or more than 2 dosing intervals (≥8 weeks). These results suggest that withholding therapy is not likely to substantially improve postoperative outcomes and that any differences in infection risk, if they exist, are likely to be small. Prolonged withholding of therapy, however, increases the risk of disease flares (30), which could potentially interfere with rehabilitation or lead to increases in glucocorticoid exposure.

We could not directly assess clinical disease flares in these data and were limited in our ability to precisely assess short-term changes in glucocorticoid doses, but we did find that patients with stop timing ≥8 weeks were more likely to fill a glucocorticoid prescription in the 3 months after surgery. Importantly, previous studies have consistently demonstrated an increased risk of adverse postoperative outcomes with glucocorticoids (6,31–34). In this study as well, an average glucocorticoid dosage of >7.5 mg/day (based on oral prescriptions in the 3 months before surgery) was associated with greater risk. Higher disease activity in glucocorticoid-treated patients may contribute to this risk, although the patients in this study were healthy enough for elective surgery.

Interestingly, patients with longer abatacept stop timing of 4–8 weeks or ≥8 weeks were more likely to have a prolonged hospitalization (defined as >90th percentile). While it is tempting to think that disease flares may have contributed to slower recoveries and longer hospitalizations, this observation could also be the result of residual differences in hospital practice or patient characteristics (e.g., abatacept might be held in patients at higher risk of complications), which have been shown to strongly influence length of stay (20,35).

Our ability to evaluate abatacept stop timing in smaller intervals (i.e., 2-week instead of 4-week intervals) was more limited, but it should be noted that adverse outcomes, especially hospitalized infection, were somewhat more common in the small group of patients who received abatacept within 2 weeks before surgery (24% of those who received abatacept within 4 weeks). Confidence intervals in this analysis are wide and results may be due to chance, but we cannot rule out the possibility that receiving an abatacept infusion within 2 weeks before surgery is associated with a clinically important increase in infection risk.

When applying these results to clinical practice, it may be reasonable preoperatively to time intravenous abatacept infusions to occur 2–4 weeks prior to surgery. Postoperatively, if the surgical site is healing well and treatment is resumed 14 days after surgery (as suggested in ACR guidelines) (5), then there will be little to no interruption of abatacept.

Several limitations of this study should be considered. Outcomes were based on claims algorithms, although hospitalized infection has previously been validated in other settings, rates of infection were similar to expected rates (4,36), and we were able to demonstrate differences in the risk of our outcomes with glucocorticoid exposure. Results of PJI analyses were less

precise, yet we found no evidence of greater risk with shorter abatacept stop timing. Residual confounding by indication is possible, although patient characteristics were quite similar even before applying propensity score weighting (especially in the <4 week and 4–8 week categories). Some patients with longer stop timing could have been tapering treatment or less adherent to treatment, but the majority of patients in all groups were receiving abatacept every 4 weeks. Although we did not directly measure disease activity, all patients were well enough to undergo elective surgery and were stable on biologic therapy. We also included measures of health care utilization as a proxy for disease severity. Patients were intentionally evaluated on chronic stable therapy with abatacept, and results might not apply to patients who have recently started abatacept, particularly because infection risk may be higher soon after biologic initiation (37–39). Additionally, it is unclear how the results of this study should be applied to subcutaneous abatacept, which was excluded because of inability to precisely determine a stop date.

Subcutaneous abatacept has a similar half-life (mean of 14.3 days) to intravenous abatacept (mean of 13.1 days) (40), but weekly instead of monthly dosing. Current ACR guidelines recommend withholding subcutaneous abatacept for 1 week before surgery (5).

In conclusion, withholding intravenous abatacept for ≥4 weeks (1 dosing interval) is not associated with a lower risk of postoperative hospitalized infection, PJI, or readmission. The questionable benefits of withholding therapy should be balanced against the known risks of disease flares, especially because glucocorticoids are associated with adverse postoperative outcomes.

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. George 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. George, Baker, Winthrop, Curtis.
Acquisition of data. George, Chen, Wu, Xie, Yang, Curtis.
Analysis and interpretation of data. George, Baker, Winthrop, Alemao, Connolly, Hsu, Simon, Curtis.

ROLE OF THE STUDY SPONSOR

Bristol-Myers Squibb contributed to the study design and reviewed and approved the manuscript prior to submission. The authors independently collected the data, interpreted the results, and had the final decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Bristol-Myers Squibb.

REFERENCES

1. Shourt CA, Crowson CS, Gabriel SE, Matteson EL. Orthopedic surgery among patients with rheumatoid arthritis 1980–2007: a population-based study focused on surgery rates, sex, and mortality. J Rheumatol 2012;39:481–5.
2. Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG. CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. Infect Control Hosp Epidemiol 1992;13:606–8.
3. Ravi B, Croxford R, Hollands S, Paterson JM, Bogoch E, Kreder H, et al. Increased risk of complications following total joint arthroplasty in patients with rheumatoid arthritis. Arthritis Rheumatol 2014;66:254–63.
4. Stundner O, Chiu YL, Sun X, Goodman SM, Russell LA, Calloway JJ, et al. Perioperative outcomes in patients with rheumatoid versus osteoarthritis for total hip arthroplasty: a population-based study. Clin Exp Rheumatol 2013;31:889–95.
5. Goodman SM, Springer B, Guyatt G, Abdel MP, Dasa V, George M, et al. 2017 American College of Rheumatology/American Association of Hip and Knee Surgeons guideline for the perioperative management of antirheumatic medication in patients with rheumatic diseases undergoing elective total hip or total knee arthroplasty. Arthritis Care Res (Hoboken) 2017;69:1111–24.
6. George MD, Baker JF, Hsu JY, Wu Q, Xie F, Chen L, et al. Perioperative timing of infliximab and the risk of serious infection after elective hip and knee arthroplasty. Arthritis Care Res (Hoboken) 2017;69:1845–54.
7. Klees BS, Wolfe CJ, Curtis CA. Brief Summaries of Medicare and Medicaid: Title XVIII and Title XIX of the Social Security Act. 2015. URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MedicareProgramRatesStats/Downloads/MedicareMedicaidSummaries2015.pdf.
8. IBM. IBM MarketScan Research Databases. URL: http://truvenhealth.com/Markets/Life-Sciences/Products/Data-Tools/MarketScan-Databases.
9. Kim SY, Servi A, Polinski JM, Mogun H, Weinblatt ME, Katz JN, et al. Validation of rheumatoid arthritis diagnoses in health care utilization data. Arthritis Res Ther 2011;13:R32.
10. Mahomed NN, Barrett JA, Katz JN, Phillips CB, Losina E, Lew RA, et al. Rates and outcomes of primary and revision total hip replacement in the United States Medicare population. J Bone Joint Surg Am 2003;85:27–32.
11. Katz JN, Barrett J, Mahomed NN, Baron JA, Wright RJ, Losina E. Association between hospital and surgeon procedure volume and the outcomes of total knee replacement. J Bone Joint Surg Am 2004;86:1909–16.
12. Mahomed NN, Barrett J, Katz JN, Baron JA, Wright J, Losina E. Epidemiology of total knee replacement in the United States Medicare population. J Bone Joint Surg Am 2005;87:1222–8.
13. Katz JN, Losina E, Barrett J, Phillips CB, Mahomed NN, Lew RA, et al. Association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States Medicare population. J Bone Joint Surg Am 2001;83:1622–9.
14. Curtis JR, Patkar N, Xie A, Martin C, Allison JJ, Saag M, et al. Risk of serious bacterial infections among rheumatoid arthritis patients exposed to tumor necrosis factor α antagonists. Arthritis Rheum 2007;56:1125–33.
15. Schneeweiss S, Setoguchi S, Weinblatt ME, Katz JN, Avorn J, Sax PE, et al. Anti-tumor necrosis factor α therapy and the risk of serious bacterial infections in elderly patients with rheumatoid arthritis. Arthritis Rheum 2007;56:1754–64.
16. Kurtz SM, Ong KL, Lau E, Bozic KJ, Berry D, Parvizi J. Prosthetic joint infection risk after TKA in the Medicare population. Clin Orthop Relat Res 2010;468:52–6.
17. Bozic KJ, Lau E, Kurtz S, Ong K, Rubash H, Vail TP, et al. Patient-related risk factors for periprosthetic joint infection and postoperative mortality following total hip arthroplasty in Medicare patients. J Bone Joint Surg Am 2012;94:749–800.
18. Centers for Medicare & Medicaid Services. 2016 measure updates and specifications report: hospital visits after hospital outpatient sur-

gery measure (risk-standardized hospital visits within 7 days after hospital outpatient surgery measure). 2016. URL: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/Hospital-Visits-after-Hospital-Outpatient-Surgery-Measure.pdfl.

19. Fry DE, Pine M, Locke D, Pine G. Prediction models of Medicare 90-day postdischarge deaths, readmissions, and costs in bowel operations. Am J Surg 2015;209:509–14.
20. Krell RW, Girotti ME, Dimick JB. Extended length of stay after surgery: complications, inefficient practice, or sick patients? JAMA Surg 2014;149:815–20.
21. Van Arendonk KJ, Tymitz KM, Gearhart SL, Stem M, Lidor AO. Outcomes and costs of elective surgery for diverticular disease: a comparison with other diseases requiring colectomy. JAMA Surg 2013;148:316–21.
22. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol 2011;64:749–59.
23. United States Census Bureau. 2009–2013 5-Year American Community Survey. URL: https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=dataset&id=dataset.en.ACS_13_5YR#.
24. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. Circulation 2016;133:601–9.
25. Spreeuwenberg MD, Bartak A, Croon MA, Hagenaars JA, Busschbach JJ, Andrea H, et al. The multiple propensity score as control for bias in the comparison of more than two treatment arms: an introduction from a case study in mental health. Med Care 2010;48:166–74.
26. Feng P, Zhou XH, Zou QM, Fan MY, Li XS. Generalized propensity score for estimating the average treatment effect of multiple treatments. Stat Med 2012;31:681–97.
27. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Stat Med 2013;32:3388–414.
28. Austin PC, Schuster T. The performance of different propensity score methods for estimating absolute effects of treatments on survival outcomes: a simulation study. Stat Methods Med Res 2016;25:2214–37.
29. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol 2008;168:656–64.
30. Goodman SM. Rheumatoid arthritis: perioperative management of biologics and DMARDs. Semin Arthritis Rheum 2015;44:627–32.
31. Bongartz T, Halligan CS, Osmon DR, Reinalda MS, Bamlet WR, Crowson CS, et al. Incidence and risk factors of prosthetic joint infection after total hip or knee replacement in patients with rheumatoid arthritis. Arthritis Rheum 2008;59:1713–20.
32. Scherrer CB, Mannion AF, Kyburz D, Vogt M, Kramers-de Quervain IA. Infection risk after orthopedic surgery in patients with inflammatory rheumatic diseases treated with immunosuppressive drugs. Arthritis Care Res (Hoboken) 2013;65:2032–40.
33. Michaud K, Fehringer EV, Garvin K, O'Dell JR, Mikuls TR. Rheumatoid arthritis patients are not at increased risk for 30-day cardiovascular events, infections, or mortality after total joint arthroplasty. Arthritis Res Ther 2013;15:R195.
34. Somayaji R, Barnabe C, Martin L. Risk factors for infection following total joint arthroplasty in rheumatoid arthritis. Open Rheumatol J 2013;7:119–24.
35. Lyman S, Fields KG, Nocon AA, Ricciardi BF, Boettner F. Prolonged length of stay is not an acceptable alternative to coded complications in assessing hospital quality in elective joint arthroplasty. J Arthroplasty 2015;30:1863–7.
36. Jämsen E, Varonen M, Huhtala H, Lehto MU, Lumio J, Konttinen YT, et al. Incidence of prosthetic joint infections after primary knee arthroplasty. J Arthroplasty 2010;25:87–92.
37. Yun H, Xie F, Delzell E, Levitan EB, Chen L, Lewis JD, et al. Comparative risk of hospitalized infection associated with biologic agents in rheumatoid arthritis patients enrolled in Medicare. Arthritis Rheumatol 2016;68:56–66.
38. Strangfeld A, Eveslage M, Schneider M, Bergerhausen HJ, Klopsch T, Zink A, et al. Treatment benefit or survival of the fittest: what drives the time-dependent decrease in serious infection rates under TNF inhibition and what does this imply for the individual patient? Ann Rheum Dis 2011;70:1914–20.
39. Dixon WG, Symmons DP, Lunt M, Watson KD, Hyrich KL, British Society for Rheumatology Biologics Register Control Centre Consortium, et al, on behalf of the British Society for Rheumatology Biologics Register. Serious infection following anti–tumor necrosis factor α therapy in patients with rheumatoid arthritis: lessons from interpreting data from observational studies. Arthritis Rheum 2007;56:2896–904.
40. Orencia (abatacept) [package insert]. Bristol-Myers Squibb; 2005. URL: https://packageinserts.bms.com/pi/pi_orencia.pdf.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1234–1242
DOI 10.1002/acr.23758

Patient Perspectives on Intravenous Biologics for Rheumatologic Disease

Lucas Grisanti,[1] Alysia Kwiatkowski,[2] Peter Dyrda,[3] Ellen Field,[4] Joseph Grisanti,[1] James Hatem,[1] Raphael J. Dehoratius,[5] and Norman Gaylis[6]

Objective. Two surveys were conducted with patients with rheumatologic diseases to evaluate perceptions of different routes of administration (intravenous [IV] or subcutaneous [SC]) for biologic therapy.

Methods. In Survey I, patient preferences toward biologic treatment were evaluated at a rheumatology practice in Buffalo, New York. In Survey II, Canadian patients enrolled in the BioAdvance patient support program and scheduled to receive IV biologic therapy were asked about their opinions of IV treatment.

Results. In Survey I, 243 rheumatology patients participated. Median patient age was 60 years, 76% were female, and 44% were naive to treatment with biologic agents. Among biologic-naive patients, the majority (56%) were open to either SC or IV treatment; biologic-naive women were more likely than men to express a preference for the route of administration. In Survey II, 1,598 patients from the BioAdvance program (including 306 rheumatology patients) completed the full survey. Among the rheumatology patients, the median age was 49 years, 58% were female, and 61% had not previously taken biologics before enrolling in the BioAdvance program. The median rating of IV favorability (on a 10-point scale, with higher numbers indicating increased favorability) recalled by rheumatology patients was 5 prior to their first program infusion, which increased to 9 after multiple treatment infusions.

Conclusion. These survey results indicate that patients with rheumatoid arthritis are generally open to IV treatment and express high satisfaction with IV therapy. Additional patient and provider education may improve shared decision-making regarding biologic therapy administration options.

INTRODUCTION

Biologic therapies are the treatment of choice for moderate-to-severe cases of many rheumatologic diseases, including rheumatoid arthritis (RA), psoriatic arthritis (PsA), and ankylosing spondylitis (AS), particularly in patients with an inadequate response to conventional disease-modifying antirheumatic drugs (DMARDs) (1–4). For RA patients with active disease despite conventional DMARDs, current recommendations do not specify a treatment of choice from among approved anti–tumor necrosis factor (anti-TNF) inhibitors (e.g., adalimumab, certolizumab pegol, etanercept, golimumab, or infliximab) or the non-TNF options (e.g., abatacept, anakinra, rituximab, sarilumab, tocilizumab, and tofacitinib) in some cases (1,2). The choice of biologic therapy for each patient is generally based on consideration of patient-related factors, disease-related factors, the mechanism of action of the prescribed medication, and patient preferences for treatment (5–9).

One of the key factors influencing patient preference for biologic therapy is the route of administration (i.e., intravenous [IV] or subcutaneous [SC]) (9–11). Biologics generally exhibit comparable efficacy, despite differences in route of administration (12). Thus, patients' and rheumatologists' preferences for route of administration may play a role in driving the choice of biologic treatment; however, rheumatologists' and patients' perspectives on various routes of administration may differ. In 1 study in the US, evaluating perceptions of biologic therapy among patients with RA and

Survey II was sponsored by Janssen Inc.

[1]Lucas Grisanti, BA, Joseph Grisanti, MD, James Hatem, BSc: Buffalo Rheumatology, Orchard Park, New York; [2]Alysia Kwiatkowski, DO, MS: Rush University Medical Center, Chicago, Illinois; [3]Peter Dyrda, BSc, MBA: Janssen Scientific Affairs, LLC, Horsham, Pennsylvania; [4]Ellen Field, MD: Private practice, Lehigh Valley, Pennsylvania; [5]Raphael J. Dehoratius, MD: Janssen Scientific Affairs, LLC, Horsham, and Sidney Kimmel School of Medicine, Thomas Jefferson University, Philadelphia, Pennsylvania; [6]Norman Gaylis, MD: Arthritis and Rheumatic Disease Specialties, Aventura, Florida.

Mr. Dryda is an employee of Janssen Canada and owns stock in Johnson & Johnson. Dr. Field has received consulting and/or speaking fees from Amgen, Genentech, and AbbVie (less than $10,000 each), and UCB and Janssen (more than $10,000 each). Dr. Joseph Grisanti has received speaking fees from Amgen and Janssen (more than $10,000 each). Dr. Dehoratius is a former employee of Janssen Scientific Affairs, LLC, and owns stock in Johnson & Johnson, of which Janssen Scientific Affairs, LLC, is a wholly owned subsidiary; and he has received consulting fees from Janssen (greater than $10,000). No other disclosures relevant to this article were reported.

Address correspondence to Joseph Grisanti, MD, 3055 Southwestern Boulevard #100, Orchard Park, NY 14127. E-mail: buffalorheum@aol.com.

Submitted for publication October 23, 2017; accepted in revised form September 11, 2018.

SIGNIFICANCE & INNOVATIONS

- A general preference for subcutaneous biologic therapies has previously been reported among rheumatology patients.
- Results of the current surveys indicate that rheumatoid arthritis patients, including biologic-naive patients, are generally open to intravenous or subcutaneous treatment.
- Patients receiving intravenous therapy express high satisfaction with this route of administration.

rheumatologists, 53% of patients reported being open to both IV infusion and SC injections, while only 41% of rheumatologists reported that patients would be open to either route of administration (10). Approximately 28% of surveyed patients reported preferring IV treatment, while the rheumatologists reported that only 16% of their patients would prefer IV therapy (10).

Although a general preference for SC biologic therapies has been reported among rheumatology patients, among those who have experienced IV therapy, there is some evidence that most have favorable opinions of the IV route of administration (13–15). In a Danish study, IV therapy was preferred by 85% of patients already receiving IV treatment (15). In another recent study evaluating the characteristics of US patients with inflammatory arthritis (e.g., RA, PsA, or AS) who had been receiving IV biologic treatment (n = 100) ranging in duration from 0.08 to 16 years, approximately 80% of patients reported being very satisfied with their IV infusions (16). In that study, patients' perception of IV therapy improved considerably after starting treatment; only 33% of patients reported an extremely favorable view of IV therapy prior to treatment, while 71% reported an extremely favorable view after receiving therapy. Advantages of IV therapy reported by patients included medication administration by a health care provider in the infusion setting, as well as having regular on-site monitoring of health status and potential side effects. In general, patients did not consider the time required for travel to infusion centers, infusion durations, or time missed from work or school to be major disadvantages of IV therapy.

Two surveys were conducted to further assess trends regarding rheumatology patient preferences for SC and IV biologic treatment. The aim of Survey I was to evaluate preferences for route of administration of biologics among rheumatology patients at a clinical practice in the US, while the aim of Survey II was to evaluate perceptions of receiving IV biologic therapy among rheumatology patients in a Janssen-sponsored infusion clinic and treatment support program (BioAdvance). Results of both surveys are presented here.

PATIENTS AND METHODS

Survey I: In-office patient preferences for biologic route of administration in a rheumatology practice. Survey I was a 20-item survey assessing patient preferences and adherence to treatment (see Supplementary Appendix 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23758/abstract) and was administered to patients seen in a suburban rheumatology practice in Greater Buffalo, New York from January through March 2015. The survey was distributed at a single-specialty rheumatology practice consisting of 4 rheumatologists and 4 nurse practitioners at 2 offices. The survey was distributed to patients with rheumatic diseases, including RA, PsA, AS, systemic lupus erythematosus, and other diseases. This survey included biologic-naive patients as well as patients who were currently or previously receiving an IV or SC treatment. Patients completed the survey voluntarily in the waiting room prior to their visit. Surveys were reported anonymously; patients provided informed consent for demographic information to be captured from their chart (based on an identifying number) if they did not provide that data in their survey.

This survey was approved by the Mercy Hospital of Buffalo Institutional Review Board. Data analyses were performed using Excel software with the Data Analysis ToolPak. For demographic characteristics, continuous outcomes were summarized using descriptive statistics (median, range, mean, and SD), and categorical outcomes were summarized using percentages. For responses to survey questions about patient preference and adherence, results were summarized as percentages reporting each response. Statistical testing included *t*-tests and analysis of variance. *P* values less than 0.05 were considered significant.

Survey II: Patient perceptions of IV therapy in the BioAdvance treatment support program. During an approximately 2.5-month period (May 5 to July 18, 2014), nearly 10,000 patients were invited to participate in an online survey (Survey II) regarding their experiences with IV infusion and the Janssen-sponsored BioAdvance program (17). All invited patients were receiving infliximab treatment at 192 clinics across Canada and were enrolled in the BioAdvance patient support program, which provides support for Canadian patients scheduled to receive IV infusions of infliximab, golimumab, or ustekinumab for any approved indication (17). The BioAdvance program streamlines the care process for both the provider and patient, incorporating a prebiologic checklist, standard administration protocol for infusion clinic staff, injection/infusion education, continuous monitoring, and reminder calls for patients, while allowing patients to have a flexible clinic selection. Written informed consent was provided by all participants or their legal guardians (17).

Survey II included a total of 28 questions, including 11 related to patient characteristics and disease, health status, and treatment characteristics, and 17 questions related to patient perceptions of their experience prior to and after multiple IV administrations in the BioAdvance program (17).

The results of the full survey, which largely included patients with inflammatory bowel disease (e.g., Crohn's disease and ulcerative colitis) who were receiving IV treatment in the BioAdvance program, have been published previously (17). In the full survey, the majority of patients (75.5%) had received ≥1 year of infliximab treatment (to be taken every 8 weeks after induction doses at 0, 2, and 6 weeks) in the BioAdvance program (17). Patients completed the surveys online, and all identifying information was removed from the survey results (17). This survey was reviewed by an institutional review board. Data analyses were performed using SAS software, version 9.2 (17). For demographic characteristics, continuous outcomes were summarized using descriptive statistics (median, range, mean, and SD), and categorical outcomes were summarized using percentages. For responses to survey questions about patient preference and adherence, results were summarized as percentages reporting each response.

RESULTS

Survey I. *Patients.* A total of 243 rheumatology patients completed the patient preferences survey. Baseline and demographic characteristics are shown in Table 1. The majority of patients were female (76%), 49% were ages >60 years, 81% had a diagnosis of RA, and 44% of the 243 were biologic naive.

Route of administration preference. When biologic-naive patients (n = 107) were asked whether they would be open to SC biologic therapy to be self-injected every 1 to 2 weeks, 63% reported being open to some degree: 31% reported being somewhat open, 13% were very open, and 18% were extremely open. When biologic-naive patients were asked whether they would be open to an IV biologic infusion every 1 to 2 months, 75% reported being at least somewhat open: 36% somewhat open, 21% very open, and 18% extremely open. In all, 56% of biologic-naive patients reported being open (i.e., somewhat, very, or extremely) to both SC and IV biologic therapy. A total of 7% of patients were open to SC treatment only, and 18% of patients were open to IV treatment only. Among biologic-naive patients who expressed a preference (n = 54), a significantly higher percentage of patients preferred IV biologic therapy (65%) compared with SC biologic therapy (35%; $P < 0.01$).

When route of administration preferences were evaluated by sex among biologic-naive patients, the proportion of patients who did not express a preference was higher among men (64%) than among women (43%), although the difference did not reach statistical significance ($P = 0.10$). Among biologic-naive women who expressed a preference, a significantly higher percentage preferred IV biologic therapy (65%) compared with SC biologic therapy (35%; $P < 0.01$). Among biologic-naive men expressing a preference, the percentage who preferred IV biologic therapy (63%) did not differ significantly from the percentage who preferred SC biologic therapy (38%; $P = 0.35$) (Figure 1A). Among all biologic-naive patients (both sexes) who expressed a route of administration preference, the proportion who preferred IV therapy increased with age, as nearly equal numbers of patients age ≤60 years endorsed SC and IV routes, while approximately 3.4 times as many patients age >60 years preferred IV (Figure 1B).

When route of administration preferences were evaluated for patients already receiving biologic therapy (i.e., biologics experienced), 50% of patients who were currently receiving SC preferred SC therapy, with 33% expressing no preference, while 58% of patients who were currently receiving IV therapy preferred IV therapy, with 29% expressing no preference.

In the overall cohort of patients who completed the survey and reported a preference, a significantly higher percentage of patients with RA preferred IV therapy (62%) compared with SC therapy (38%; $P = 0.0004$). In contrast, among patients with PsA, a significantly higher percentage of patients preferred SC therapy (82%) compared with IV therapy (18%; $P < 0.0001$). Only 6 patients with AS reported a preference, with 50% preferring SC therapy and 50% preferring IV infusion.

Adherence with SC therapy was assessed with the following question: "Assuming you are not having surgery and do not have an infection, how frequently do you give yourself your shot exactly as prescribed?" Patients who were not currently receiving home-based SC treatment were instructed to skip this question; of the 73 patients who did respond, most reported taking their shot (injection) exactly as they were

Table 1. Survey I: patient baseline and demographic characteristics*

Characteristic	Values (n = 243)
Age, years	
Median (range)	60 (19–92)
Mean ± SD	59.8 ± 12.8
Age group	
<40	7
40–60	44
>60	49
Sex	
Female	76
Male	24
Disease state†	
Rheumatoid arthritis	81
Psoriatic arthritis	13
Ankylosing spondylitis	5
Current medications	
Methotrexate	55
Biologics	50

* Values are the percentage, unless indicated otherwise.
† Patients could report >1 disease state. Percentages do not add up to 100%; 5 patients reported having systemic lupus erythematosus only, 6 patients reported having undifferentiated spondyloarthropathies, and 34 patients reported multiple diagnoses.

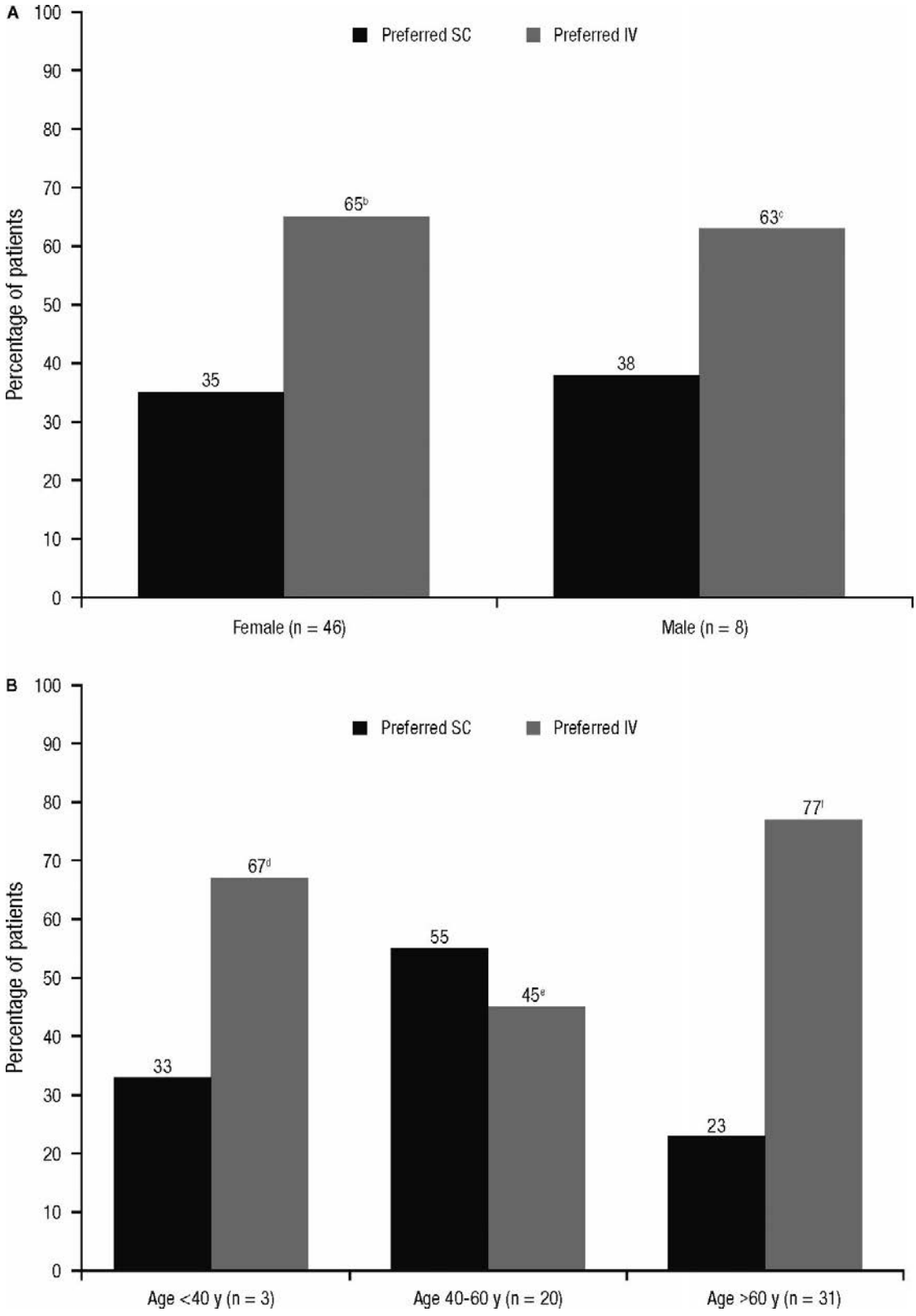

Figure 1. Survey I: biologic-naive patient preferences for subcutaneous (SC) or intravenous (IV) biologic therapy among patients who expressed a preference by sex (**A**) and age group (**B**). Patients who reported that they did not have a preference and would consider either option or were unwilling to consider either option were excluded from these analyses. b indicates $P < 0.01$ versus proportion preferring SC biologic therapy; c indicates $P = 0.35$ versus proportion preferring SC biologic therapy; d indicates $P = 0.52$ versus proportion preferring SC biologic therapy; e indicates $P = 0.54$ versus proportion preferring SC biologic therapy; f indicates $P < 0.01$ versus proportion preferring SC biologic therapy.

Table 2. Survey II: patient baseline and demographic characteristics*

Characteristic	AS (n = 113)	PsA (n = 74)	RA (n = 119)	Overall rheumatic diseases (n = 306)
Age, years				
Median (range)	44 (18–69)	49 (23–69)	54 (20–81)	49 (18–81)
Mean ± SD	45.0 ± 11.4	49.0 ± 11.5	51.8 ± 14.0	48.6 ± 12.8
Age group				
<40	33	27	24	28
40–60	56	55	47	52
>60	12	18	29	20
Sex				
Male	68	35	21	42
Female	32	65	79	58
Health rating†				
Median (IQR)	7 (6–9)	7 (6–8)	7 (6–9)	7 (6–9)
Mean ± SD	7.2 ± 1.9	6.8 ± 2.0	6.9 ± 2.3	7.0 ± 2.1
Category				
1–4	6	12	14	11
5–6	26	35	15	24
7–8	39	35	44	40
9–10	29	18	27	25
Infliximab treatment, duration‡				
0–2 months	5	5	7	6
3–6 months	7	10	4	7
7–11 months	8	7	7	7
1–2 years	13	14	10	12
>2 years	66	64	72	68
No. of prior biologic therapies§				
None⁋	66	42	69	61
1	18	27	16	20
2	12	19	8	12
3	2	10	3	4
>3	3	3	4	3
Employment status				
Full-time	49	27	37	39
Part-time	12	15	8	11
Student	3	1	3	2
Retired	12	15	29	19
Long-term disability	20	28	15	20
Unemployed	5	12	8	8

* Values are the percentage, unless indicated otherwise. Percentages may not total 100%, due to rounding. AS = ankylosing spondylitis; PsA = psoriatic arthritis; RA = rheumatoid arthritis; IQR = interquartile range.
† Possible health rating ranged from 0 to 10, with higher scores indicating better health.
‡ AS: n = 112; PsA: n = 73; RA: n = 118; overall rheumatic diseases: n = 303.
§ AS: n = 111; PsA: n = 72; RA: n = 116; overall rheumatic diseases: n = 300.
⁋ Patients had received no prior biologics before initiating intravenous treatment in the BioAdvance program.

supposed to all of the time (58%), or most of the time (41%), with only 1 patient (1%) reporting a lack of adherence with SC treatment ("I'm somewhat casual about giving myself a shot, and I take it primarily when I feel I need it"). When the data were analyzed by age, all patients age <40 years (n = 11) and >60 years (n = 31) reported adherence with treatment most or all of the time; 97% of patients ages 40–60 years (n = 31) reported adherence with treatment most or all of the time.

Survey II. *Patients.* Of 10,000 invitations, 1,712 responses to Survey II were provided by patients receiving IV therapy for any indication from clinics enrolled in the BioAdvance program (17). A total of 1,598 patients completed the full survey. In all, 306 patients with rheumatic diseases completed the full survey (RA: n = 119; AS: n = 113; PsA: n = 74). Nearly all of these patients (99%) were receiving IV infliximab after enrolling in the BioAdvance program. Baseline and demographic characteristics for patients with rheumatic diseases who completed full surveys are shown in Table 2. Among patients with rheumatic diseases, the population with RA was the oldest (median age 54 years [range 20–81 years]) and had the smallest proportion of male patients (21%) (Table 2). The population with AS was the youngest (median age 44 years [range 18–69 years]) and had the highest proportion of male patients (68%). Fewer than 40% of patients with rheumatic diseases were employed full-time, 19% were retired, and 20% were on disability.

The proportion of patients with rheumatic diseases with a health rating of 7 or greater (possible score 0 to 10, with higher scores indicating better health) ranged from 53% (for patients with PsA) to 71% (for patients with RA) (Table 2). The majority of patients (80%) with rheumatic diseases had been receiving infliximab treatment for 1 year or more.

Patient perceptions of IV therapy in the BioAdvance treatment program. The proportion of patients with rheumatic diseases who felt that their time commitment to obtain IV biologic therapy was highly worthwhile (score of 9 of a possible total score of 10) increased after having received multiple IV infusions. Based on recall, 38.2% of patients reported that their time commitment to IV biologic therapy was highly worthwhile prior to therapy, compared with 66% who rated their time commitment as highly worthwhile after multiple IV infusions (Figure 2A). Patients with rheumatic diseases who believed that the time commitment required for IV therapy was highly worthwhile (score of at least 9) typically maintained that belief, while 84% of patients who did not see their time commitment as worthwhile (score of 1 to 6) prior to therapy increased their rating by at least 1 point, with 77% of patients increasing their rating to 7 or greater, after undergoing IV infusions in the BioAdvance program.

Prior to their first IV infusion, the median patient rating of IV favorability (based on recall) was 5 of a possible score of 10 (with higher numbers indicating increased favorability). After multiple treatment infusions, the median rating increased to 9 of 10. The proportion of patients with rheumatic diseases with a highly positive impression of IV infusions (score of 9) increased after having received multiple IV infusions. Based on recall, 16% of patients reported a highly favorable rating for their impression of IV therapy in general (score of at least 9) prior to treatment, while 57.2% reported a highly favorable rating after multiple IV infusions (Figure 2B). The proportion of patients with rheumatic diseases with a highly positive perception (score of at least 9) of IV therapy prior to initiating IV treatment mainly remained positive, while 94% of patients who had a negative perception (score of 1 to 6) prior to therapy increased their rating by at least 1 point, with 88% of patients increasing their rating to 7 or greater, after undergoing infusions in the BioAdvance program. On average, patients reported that having a health care practitioner on site at the BioAdvance clinic was very important to them (median score 9 [range 1–9]), based on a Likert scale of 1 (not important at all) to 10 (extremely important). Using the same Likert scale, patients rated the importance of spending time with other patients in the BioAdvance clinic as only moderately important (median score 5 [range 1–9]).

DISCUSSION

The 2 surveys presented here directly evaluated patient preferences regarding mode of administration (SC or IV) for biologic treatment among patients with rheumatic diseases in the US and Canada. Based on the results of Survey I, most biologic-naive patients (56%) were open to either SC or IV biologic therapy. These results were in keeping with findings of a previous study evaluating the perceptions of biologic therapy among US patients with RA, which showed that 53% of patients were open to either SC or IV therapy (10). In Survey I, a higher proportion of biologic-naive patients overall expressed a preference for IV therapy than for SC therapy, although preferences varied by age, sex, and disease state. Younger patients (age <40 years) showed a strong preference for SC therapy; however, younger patients with rheumatic diseases have previously been shown to be less adherent to SC biologic treatment than older patients (18,19). In general, female patients were more likely to state a preference for the route of administration of their biologic treatment than male patients, suggesting that men were more likely to allow their physician to decide the best route of administration. The specific disease state also appeared to affect patient preferences for route of administration; IV therapy was preferred by a higher proportion of patients with RA, while SC therapy was preferred by a higher proportion of patients with PsA. In a survey reported recently by Gaylis et al (16), patients with inflammatory arthritis reported a number of perceived advantages with IV therapy, including additional monitoring by health care staff, the immediate availability of health care resources, less frequent dosing, ease of scheduling administrations, and reduced fear of self-injection.

Results of Survey II were also consistent with previous data in the literature (16) regarding patients' acceptance of IV therapy; Canadian patients with rheumatic diseases surveyed generally reported very favorable perceptions of IV therapy and felt that the time commitment to obtain biologic therapy was highly worthwhile, and perceptions of IV therapy and the associated time commitment improved favorably after receiving multiple infusions in the program.

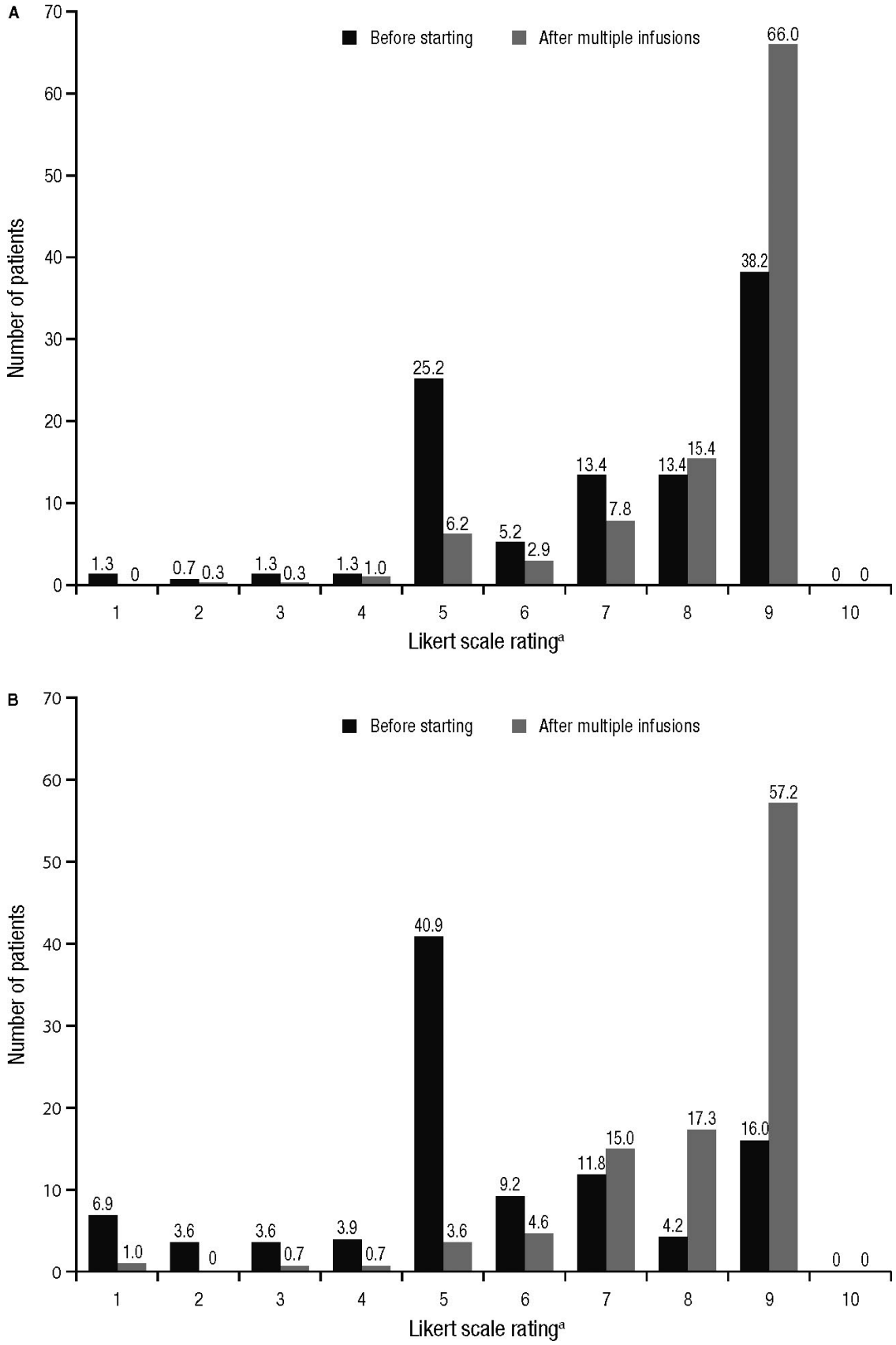

Figure 2. Survey II: proportion of patients with rheumatic diseases, with their impressions of time commitment required for intravenous infusion (**A**), and intravenous infusions in general (**B**). a = Scale of 0 to 10, with higher scores indicating a higher rating.

Results of both surveys reported here indicate that patients with inflammatory arthropathies generally have favorable perceptions of IV therapy. Based on the authors' clinical experience, patient education and convenience appear to be factors driving the choice of treatment modality; biologic-naive patients may be unaware of the option of IV biologic therapy or perceive home-based SC biologic therapy as a safer option than IV therapy. In a previous study of 500 patients with RA, less than half of patients reported receiving information about alternative biologic therapies (specifically, anti-TNF agents) from their physicians (20). Thus, patient education regarding IV biologics as an option for treatment could represent a key unmet need. Results from the previous literature on patient and physician preferences for biologic therapy suggest a disconnect in patient and physician perceptions (21). Results of the 2 surveys reported here, along with the results of previous patient preference studies (8–10,15,20–22) may address this disconnect by providing physicians with information regarding patients' attitudes and concerns around different rheumatology treatments. This understanding of the patient perspective may help guide physicians' discussions with their patients about different biologic therapy options.

To facilitate patient-centered, collaborative care of rheumatology patients, the choice of IV versus SC therapy should be discussed as part of a shared decision-making process. In Survey II, most patients, particularly those who had received multiple infusions, perceived the time commitment required for IV therapy as highly worthwhile, which may be related both to the effects of consistent biologic treatment and regular contact with a health care provider.

As previously noted, there is often a disconnect between patient and physician perceptions of IV therapy, with physicians presenting a more negative perception of IV therapy than that of patients who have been receiving IV therapy (10,21). Bridging this gap in patient and provider communication to ensure that the benefits, as well as risks, of both IV and SC options are presented to patients, to allow for a more balanced decision-making process, could assist with improving biologic treatment adherence and outcomes.

Another key aspect of biologic therapy to be considered is adherence to therapy. Adherence to treatments for RA remains problematic; in a recent meta-analysis, an overall adherence rate of 66% for all evaluated therapeutics (biologics, conventional DMARDs, steroids, and nonsteroidal antiinflammatory drugs) was reported (23). Adherence to treatment could be more readily monitored for patients receiving IV therapy at an infusion center than for those self-administering SC therapy; however, studies of the impact of the different routes of administration on adherence are generally lacking. In addition to the perceived benefit for potentially improving patient adherence to treatment, increased oversight of patients by health care providers in an infusion center may allow better management of patients' overall health and the ability to rapidly address any potential side effects of treatment. This concept is supported by the results of Survey II, in which patients reported that having a health care provider on-site at the infusion center was important to them. In addition, a recent study by Gaylis et al (16) showed that patients believed that medication administration by a health care provider in the infusion setting, as well as having regular on-site monitoring of health status and potential side effects, were key benefits of IV therapy. However, while patients perceived an advantage to receiving treatment at an IV center, more research will be needed to demonstrate an association with general health or rheumatology outcomes.

The findings of the current surveys in patients with rheumatic diseases could potentially be applied to other specialties. For example, in patients with inflammatory bowel disease, findings regarding patient preference for IV or SC treatment are conflicting (24,25). In 1 survey evaluating preferences for route of administration of anti-TNFs in 78 patients with Crohn's disease, 42% of patients preferred the IV option while 24% preferred the SC option (25). In a separate survey of 100 patients with Crohn's disease, 64% preferred an SC anti-TNF, while 25% preferred an IV anti-TNF (24). These conflicting results suggest that patients with inflammatory bowel disease may experience similar challenges in treatment decision-making as those identified for patients with rheumatic diseases, and that broad education about IV and SC biologic options could be valuable for this patient population as well.

The results of these 2 surveys were subject to certain limitations. Some of the questions in these surveys were based on patient recall. Patient preferences for mode of administration may have been affected by experiences with prior or current treatments for other conditions. In Survey I, the results for the overall cohort of rheumatology patients may have been influenced by other factors (e.g., sex, disease state). Survey II included only patients in the BioAdvance program. Their perceptions may have been influenced by multiple different factors specifically related to the program, such as interactions with the BioAdvance coordinator and participating physicians. Furthermore, there may be systemic differences in how health care is administered in Canada that may not be applicable in other countries.

Taken together, results of the current surveys indicate that rheumatology patients are generally open to IV treatment and express high satisfaction with IV therapy. Based on these data and the authors' many years of combined clinical experience, we believe that additional patient and provider education may contribute to an improved shared decision-making process for the patient.

ACKNOWLEDGMENTS

The authors thank the participants of an advisory board (including rheumatologists, rheumatology nurses, and practice managers), at which the results of these 2 surveys were discussed. This advisory board included the following participants: Douglas Conway, Ellen Field, Norman Gaylis, Joseph Grisanti, Jeremy Lavender, Stephanie Ott, Dianne Petrone, Victoria Ruffing,

James Salliotte, Elizabeth Schultz, Kathy Wittemann, and Andrea Zlatkus. This panel's discussion of the results contributed to the concepts, perceptions, and interpretations expressed in the Discussion section of this article. The authors would particularly like to thank Victoria Ruffing for her ongoing contributions and insights during the development of this article. Writing and editorial support were provided by Megan Knagge, PhD, of MedErgy.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. J. Grisanti 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. L. Grisanti, Kwiatkowski, Dyrda, J. Grisanti, Gaylis.

Acquisition of data. L. Grisanti, Dyrda, J. Grisanti, Dehoratius, Gaylis.

Analysis and interpretation of data. L. Grisanti, Kwiatkowski, Dyrda, Field, J. Grisanti, Hatem, Dehoratius, Gaylis.

ROLE OF THE STUDY SPONSOR

Employees of Janssen, Inc., were involved in the design of Survey II and the collection, analysis, and interpretation of the data from that survey. All authors participated in the writing of the manuscript and decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Janssen, Inc.

REFERENCES

1. Singh JA, Saag KG, Bridges SL Jr, Aki AE, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheumatol 2016;68:1–26.
2. Smolen JS, Landewé R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis 2014;73:492–509.
3. Gossec L, Smolen JS, Gaujoux-Viala C, Ash Z, Marzo-Ortega H, van der Heijde D, et al. European League Against Rheumatism recommendations for the management of psoriatic arthritis with pharmacological therapies. Ann Rheum Dis 2012;71:4–12.
4. 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 2015;68:282–98.
5. Blake T, Rao V, Hashmi T, Erb N, O'Reilly SC, Shaffy S, et al. The perplexity of prescribing and switching of biologic drugs in rheumatoid arthritis: a UK regional audit of practice. BMC Musculoskelet Disord 2015;15:290.
6. Keyser FD. Choice of biologic therapy for patients with rheumatoid arthritis: the infection perspective. Curr Rheumatol Rev 2011;7:77–87.
7. Monti S, Klersy C, Gorla R, Sarzi-Puttini P, Atzeni F, Pellerito R, et al. Factors influencing the choice of first- and second-line biologic therapy for the treatment of rheumatoid arthritis: real-life data from the Italian LORHEN Registry. Clin Rheumatol 2017;36:753–61.
8. Alten R, Krüger K, Rellecke J, Schiffner-Rohe J, Behmer O, Schiffhorst G, et al. Examining patient preferences in the treatment of rheumatoid arthritis using a discrete-choice approach. Patient Prefer Adherence 2016;10:2217–28.
9. Louder AM, Singh A, Saverno K, Cappelleri JC, Aten AJ, Koenig AS, et al. Patient preferences regarding rheumatoid arthritis therapies: a conjoint analysis. Am Health Drug Benefits 2016;9:84–93.
10. Bolge SC, Goren A, Brown D, Ginsberg S, Allen I. Openness to and preference for attributes of biologic therapy prior to initiation among patients with rheumatoid arthritis: patient and rheumatologist perspectives and implications for decision making. Patient Prefer Adherence 2016;10:1079–90.
11. Poulos C, Hauber AB, González JM, Turpcu A. Patients' willingness to trade off between the duration and frequency of rheumatoid arthritis treatments. Arthritis Care Res (Hoboken) 2014;66:1008–15.
12. Schwartzman S, Morgan GJ Jr. Does route of administration affect the outcome of TNF antagonist therapy? Arthritis Res Ther 2004;6 Suppl 2:S19–23.
13. Barton JL. Patient preferences and satisfaction in the treatment of rheumatoid arthritis with biologic therapy. Patient Prefer Adherence 2009;3:335–44.
14. Williams EL, Edwards CJ. Patient preferences in choosing anti-TNF therapies-R1. Rheumatology (Oxford) 2006;45:1575–6.
15. Huynh TK, Ostergaard A, Egsmose C, Madsen OR. Preferences of patients and health professionals for route and frequency of administration of biologic agents in the treatment of rheumatoid arthritis. Patient Prefer Adherence 2014;8:93–9.
16. Gaylis NB, Sagliani J, Black S, Tang KL, DeHoratius B, Kafka WA, et al. Patient-reported outcome assessment of inflammatory arthritis patient experience with intravenously administered biologic therapy. Patient Prefer Adherence 2017;11:1543–53.
17. Jones J, Borgaonkar M, Siffledeen J, O'Reilly R, Dajnowiec D, Williamson M, et al. BioAdvance patient support program survey: positive perception of intravenous infusions of infliximab. Manag Care 2017;26:41–8.
18. Curkendall S, Patel V, Gleeson M, Campbell RS, Zagari M, Dubois R. Compliance with biologic therapies for rheumatoid arthritis: do patient out-of-pocket payments matter? Arthritis Rheum 2008;59:1519–26.
19. Borah BJ, Huang X, Zarotsky V, Globe D. Trends in RA patients' adherence to subcutaneous anti-TNF therapies and costs. Curr Med Res Opin 2009;25:1365–77.
20. Sylwestrzak G, Liu J, Stephenson JJ, Ruggieri AP, DeVries A. Considering patient preferences when selecting anti-tumor necrosis factor therapeutic options. Am Health Drug Benefits 2014;7:71–81.
21. De Mits S, Lenaerts J, Vander Cruyssen B, Mielants H, Westhovens R, Durez P, et al. A nationwide survey on patient's versus physician's evaluation of biological therapy in rheumatoid arthritis in relation to disease activity and route of administration: the Be-Raise study. PLoS One 2016;11:e0166607.
22. Scarpato S, Antivalle M, Favalli EG, Naccu F, Frigelli S, Bartoli F, et al. Patient preferences in the choice of anti-TNF therapies in rheumatoid arthritis: results from a questionnaire survey (RIVIERA study). Rheumatology (Oxford) 2010;49:289–94.
23. Scheiman-Elazary A, Duan L, Shourt C, Agrawal H, Ellashof D, Cameron-Hay M, et al. The rate of adherence to antiarthritis medications and associated factors among patients with rheumatoid arthritis: a systematic literature review and metaanalysis. J Rheumatol 2016;43:512–23.
24. Vavricka SR, Bentele N, Scharl M, Rogler G, Zeitz J, Frei P, et al. Systematic assessment of factors influencing preferences of Crohn's disease patients in selecting an anti-tumor necrosis factor agent (CHOOSE TNF TRIAL). Inflamm Bowel Dis 2012;18:1523–30.
25. Allen PB, Lindsay H, Tham TC. How do patients with inflammatory bowel disease want their biological therapy administered? BMC Gastroenterol 2010;10:1.

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1243–1248
DOI 10.1002/acr.23755

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Adverse Events and Resource Use Before and After Treat-to-Target in Rheumatoid Arthritis: A Post Hoc Analysis of a Randomized Controlled Trial

Daniel H. Solomon,[1] Zhi Yu,[1] Jeffrey N. Katz,[1] Asaf Bitton,[1] Cassandra Corrigan,[1] Liana Fraenkel,[2] Leslie R. Harrold,[3] Josef S. Smolen,[4] Elena Losina,[1] and Bing Lu[1]

Objective. Treat-to-target (TTT) is an accepted paradigm for care of patients with rheumatoid arthritis (RA). Because TTT can be associated with more medication switches, concerns arise regarding whether implementing TTT may increase adverse events and/or resource use. The aim of this study was to examine adverse events and resource use during the preintervention and intervention periods of the TTT intervention trial.

Methods. We used data from 6 practices enrolled in an 18-month cluster-randomized controlled trial to compare adverse events and resource use before (months 1–9) and during (months 10–18) a TTT intervention. The outcomes of interest, adverse events and resource use, were based on medical record review of all rheumatology visits for RA patients before and during the intervention.

Results. We examined records for 321 patients before the intervention and 315 during the intervention. An adverse event was recorded in 10.2% of visits before the intervention and 8.8% of visits during the intervention ($P = 0.41$). Biologic disease-modifying antirheumatic drugs were taken by 53.6% of patients before the intervention and 49.8% of patients during the intervention ($P = 0.73$). Rheumatology visits were more frequent before the intervention (mean ± SD 4.0 ± 1.4) than during the intervention (mean ± SD 3.6 ± 1.2; $P = 0.02$). More visits were accompanied by monitoring laboratory tests before the intervention (90.0%) compared with during the intervention (52.7%; $P < 0.001$). A greater percentage of visits before the intervention included diagnostic imaging (15.4%) versus during the intervention (8.9%; $P < 0.001$).

Conclusion. We observed similar rates of adverse events before and during the implementation of TTT for RA. Rheumatology visits, use of laboratory monitoring, and diagnostic imaging did not increase during the TTT intervention.

INTRODUCTION

Treat-to-target (TTT) has become a widely endorsed paradigm for treatment of rheumatoid arthritis (RA). This approach involves the provider and patient setting a target disease activity, measurement of disease activity at each visit, and adjustment of treatments until target disease activity is met. Shared decision-making is used both to set the target and to determine treatments (1). We worked with 11 rheumatology practice sites in the US in a cluster-randomized controlled trial to test a learning collaborative intervention for improving implementation of TTT (2). During the trial, we measured the implementation of TTT using a standardized medical record review, and each component was noted as present or absent and the percent of components at each visit was averaged. The intervention increased TTT implementation from a baseline of 11% to 57% after the 9-month learning collaborative (3). A very similar level of improvement was seen in a second phase of the work among the sites originally randomized to wait-list control (4).

ClinicalTrials.gov identifier: NCT02260778.

Dr. Solomon's work was supported by the NIH (grants P60-AR-047782 and K24-AR-055989). Dr. Fraenkel's work was supported by the NIH (grant K24-AR-060231) and a Rheumatology Research Foundation Innovative Research Grant.

[1]Daniel H. Solomon, MD, Zhi Yu, MPH, Jeffrey N. Katz, MD, Asaf Bitton, MD, Cassandra Corrigan, BA, Elena Losina, PhD, Bing Lu, PhD: Brigham and Women's Hospital, Boston, Massachusetts; [2]Liana Fraenkel, MD: Yale School of Medicine, New Haven, Connecticut; [3]Leslie R. Harrold, MD: University of Massachusetts Medical Center, Worcester; [4]Josef S. Smolen, MD: University of Vienna, Vienna, Austria.

Dr. Solomon has received research support from Amgen, Pfizer, Bristol-Myers Squibb, Genentech, and AbbVie, and has received consulting fees from Corrona (less than $10,000). Dr. Harrold has received consulting fees from Bristol-Myers Squibb and Roche (less than $10,000 each). Dr. Losina has received consulting fees from TissueGene and Samumed (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Daniel H. Solomon, MD, Division of Rheumatology, Brigham and Women's Hospital, 75 Francis Street, Boston MA 02115. E-mail: dsolomon@bwh.harvard.edu.

Submitted for publication April 3, 2018; accepted in revised form September 11, 2018.

SIGNIFICANCE & INNOVATIONS

- Among 6 rheumatology practices enrolled in a prospective randomized trial, we observed no increase in adverse events when the sites implemented treat-to-target for rheumatoid arthritis.
- No increase was observed in health care resource use when implementing treat-to-target for rheumatoid arthritis.

In other clinical areas, the implementation of a TTT paradigm has been associated with an increase in adverse events. One trial that tested a TTT intervention for diabetes mellitus showed an increased risk of hypoglycemia and death when using a target glycosylated hemoglobin level (HbA_{1C}) (5). While some hypertension trials have shown improvements in clinical outcomes when targeting a goal for blood pressure, a large meta-analysis demonstrated an increased risk of severe hypotension when targeting blood pressure (6–8). One medical society recently recommended a less aggressive target threshold for HbA_{1C} (9). There are also concerns that implementing a TTT paradigm might increase resource use. We conducted a post hoc analysis of the TTT intervention trial to examine adverse events and resource use during the preintervention and intervention periods of the trial.

MATERIALS AND METHODS

Study design. The current analyses examined 6 rheumatology sites that were part of a cluster-randomized controlled trial (the Treat-to-Target in RA: Collaboration to Improve Adoption and Adherence [TRACTION] trial) testing a learning collaborative to improve implementation of TTT (Figure 1 shows the overall study design). Details of the learning collaborative have been previously described (see Supplementary Appendix 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23755/abstract) (2). Briefly, the learning collaborative expert faculty provided guidance to teams through learning sessions. Teams worked on process improvement through tests of small changes in their practice. They subsequently evaluated these changes and adopted those that worked. Tests of change were conducted over the course of several-day cycles. The faculty reviewed results and provided feedback. Teams focused on routine metrics collected across all sites and attempted to spread successful interventions to the broader provider group. Implementing TTT required a modified RA treatment discussion for some providers and a change in documentation. Providers worked with patients to choose a disease target, typically low disease activity or remission. The practices were required to select a measure and then use it routinely at all visits. Providers were asked to respond to the disease activity measure when the target had not been reached. This response required adjustment of treatment or documentation of why no changes were made. The learning collaborative also involved learning sessions; the first was a 1-day face-to-face meeting. Subsequent learning sessions were conducted via webinar, approximately once per month. All team members from each site were expected to attend these calls, but this goal was not always

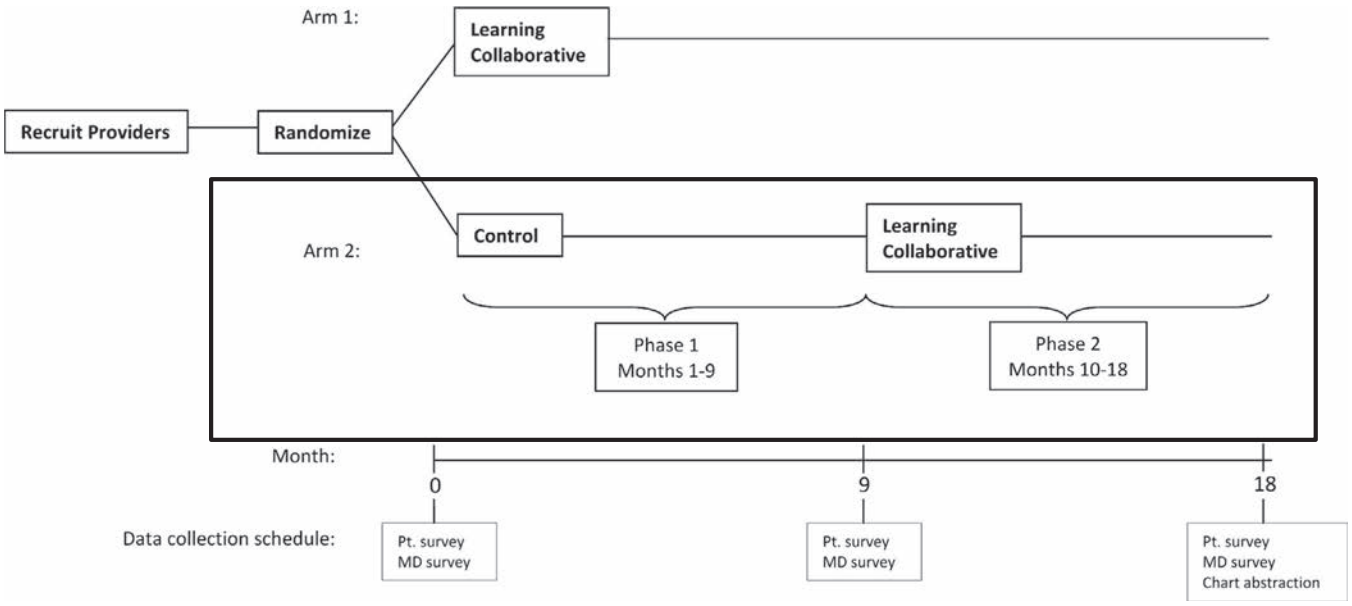

Figure 1. Design of the TRACTION trial. Phase I was the cluster-randomized controlled clinical trial comparing the learning collaborative intervention with a wait-list control group. Phase II provided the learning collaborative intervention to the phase I control group while the phase I intervention group was observed. The current article compares the same 6 control sites during phase I to phase II (outlined with the black box), with respect to adverse events and resource utilization.

possible. All learning sessions were recorded and made available on the web-based collaborative tool.

For the current analyses, we used data from before and during the intervention from these sites (3). The original trial comprised two 9-month periods: during the first 9 months, the 6 sites were in the wait-list control group (before intervention) and during the second 9 months, they received the intervention. We used these 2 periods to examine adverse events and resource use before and during TTT implementation. The appropriate institutional review boards approved all study activities.

Study population. Each site chose at least 30 patients with RA seen during the periods of interest, 3 months prior to intervention and the last 3 months of the intervention (a period of 3 months was chosen as the sampling frame since most patients with RA will have ≥1 visit every 3 months). The intervention for these 6 sites occurred during 9 months, from November 2015 to July 2016. The patients were chosen randomly from these two 3-month periods. The sites were all rheumatology practices from across the US. They all had at least 2 rheumatology providers and 2 sites included nonphysician providers. Four sites had an academic affiliation.

Outcomes: adverse events and resource use. We examined all visits for the selected patients during the 2 study periods described above, assessing for possible medication-related adverse events and resource use. Adverse events of interest included rashes, oral ulcers or mouth pain, alopecia, infections requiring antibiotics, liver toxicity as manifested by abnormal liver function tests (above the upper limits of normal) and/or abnormal liver imaging, cytopenias as manifested by complete blood counts below the lower limits of normal, renal insufficiency defined as a 50% decrease in creatinine clearance, a new cancer diagnosis, gastrointestinal symptoms (e.g., nausea, vomiting, diarrhea, unexplained weight loss/gain, abdominal pain, or dyspepsia), and other miscellaneous side effects. Three trained research assistants reviewed all the medical records using a standardized data abstraction form.

The same set of visits was assessed for resource use. We inspected visit notes, laboratory records, prescription lists, and imaging data for the following categories: prescription of biologic and nonbiologic disease-modifying antirheumatic drugs (DMARDs); orders and completion of monitoring laboratory tests, such as complete blood counts, liver function tests, serum creatinine, erythrocyte sedimentation rate, and C-reactive protein level; and orders and completion of diagnostic imaging, such as plain radiographs, computed tomography scans, magnetic resonance imaging, or dual energy radiograph absorptiometry. Most sites did not do routine joint ultrasounds, and these were not included as part of the diagnostic imaging assessment. Each aspect of the data analyzed was determined as absent or present, based on a medical record review by trained study staff (interrater κ = 94% [95% confidence interval (95% CI) 90–99]), and intrarater reliability κ = 98% (95% CI 95–99).

Statistical analysis. To compare adverse events across periods, we assessed the following metrics: the percent of visits during the period with any of the above adverse events, the percent of patients with any of the above adverse events, and the mean number of any of the above adverse events per patient. We also calculated similar metrics for resource use, including the percent of visits with each resource used, the percent of patients with each resource used, and the mean number of resources used per patient.

Table 1. Characteristics of the patients before and during the treat-to-target intervention*

	Before (n = 321)	During (n = 315)	*P*
Age, mean ± SD years	59.7 ± 4.3	61.0 ± 13.5	0.28
BMI, mean ± SD kg/m^2†	30.1 ± 7.5	30.0 ± 8.1	0.90
Female	250 (77.9)	260 (82.5)	0.14
RA duration, years†			0.47
<2	22 (16.1)	19 (11.3)	–
2–5	39 (28.5)	52 (31.0)	–
6–10	30 (21.9)	31 (18.5)	–
>10	46 (33.6)	66 (39.3)	–
Serologic status†			0.94
Positive	193 (76.3)	180 (76.6)	–
Negative	60 (23.7)	55 (23.4)	–
Use of synthetic DMARDs	248 (77.3)	248 (78.7)	0.65
Use of biologic DMARDs	131 (40.8)	122 (38.7)	0.59
Comorbidity index, mean ± SD	1.33 ± 0.6	1.31 ± 0.7	0.62
Joint erosion			0.86
Yes	109 (53.4)	103 (52.6)	–
No	95 (46.6)	93 (47.5)	–
Total medications			0.14
0	0 (0.0)	0 (0.0)	–
1–5	42 (13.1)	26 (8.3)	–
6–10	104 (32.4)	105 (33.2)	–
>10	175 (54.5)	184 (58.4)	–

* Values are the number (%) unless indicated otherwise. Some percentages are larger because of missing data. BMI = body mass index; RA = rheumatoid arthritis; DMARDs = disease-modifying antirheumatic drugs.
† Data were missing for the following variables: age (n = 64), BMI (n = 150), RA duration (n = 331), serologic status (n = 148), and joint erosions (n = 236).

The samples of patients at each site across the 2 periods were different. We described and compared the baseline characteristics of patients during the 3 months prior to the control period and the 3 months prior to the intervention using 2 sample *t*-tests, chi-square tests, or nonparametric tests when applicable. The metrics for adverse events and resource use were then compared across the 2 periods. Due to the hierarchical structure of data, we used generalized linear mixed models (GLMMs) to adjust for site effect and within-provider correlation. GLMMs for binary, Poisson, or negative binomial outcomes were used based on distributions of adverse events and resource use in the analysis. All analyses were performed using SAS software, version 9.4.

RESULTS

We examined records for 321 patients included in the assessment for the 9 months before the intervention and 315 patients during the 9 months of the intervention; these numbers included 1,284 visits before and 1,134 visits during the intervention (Table 1). Patients were similar in all respects across the 2 periods: mean ± SD age 60 ± 14 years, 81% female, and 76% seropositive. There was a similar proportion of patients using biologic DMARDs in the period before the intervention (40.8%) compared with during the intervention (38.7%) ($P = 0.59$). The percentage with joint erosions was also similar: 53.4% before and 52.6% during the intervention ($P = 0.86$).

Adverse events were similar across periods under consideration (Table 2). Any adverse event was recorded in 10.2% of visits before the intervention and 8.8% during the intervention (adjusted $P = 0.41$). The percent of patients with an abnormal liver function test was slightly greater in the period before the intervention (0.8%) than during the intervention (0.3%) (adjusted $P = 0.12$). Mucocutaneous adverse events (alopecia, oral ulcers, or any rash) were also slightly more common in the period before the intervention (1.7%) than during the intervention (0.8%) (adjusted $P = 0.07$). The percent of patients who experienced an infection trended higher during the intervention (12.1%) compared with before the intervention (9.4%; adjusted $P = 0.18$). Gastrointestinal symptoms were experienced by a similar percentage of patients in both periods: 2.2% before and 2.2% during the intervention (adjusted $P = 0.79$).

Finally, we compared resource use across the 2 periods (Table 2). Biologic DMARDs were taken by 53.6% of patients before the intervention and 49.8% during the intervention ($P = 0.73$). Rheumatology visits were more frequent before the intervention (mean ± SD 4.0 ± 1.4) than during the intervention (mean ± SD 3.6 ± 1.2; $P = 0.02$). More visits were accompanied by monitoring laboratory tests before the intervention (90.0%) compared with during the intervention (52.7%; $P < 0.001$). A greater percentage of visits before the intervention included diagnostic imaging (15.4%) versus during the intervention (8.9%; $P < 0.001$).

DISCUSSION

Ample evidence supports TTT as an effective paradigm for managing RA (10). While there appear to be opportunities for enhancing the use of TTT in rheumatology practice (11,12), understanding the potential for excess adverse events and resource use is required to optimize implementation. Using data from a recently completed randomized controlled trial, we examined adverse event rates and resource use in 6 rheumatology practices prior to, and during, implementation of TTT. We did not observe clinically important or statistically significant increases in adverse events during the intervention. We saw no increase in the use of biologic DMARDs, laboratory monitoring, diagnostic imaging, or overall rheumatology visits during the intervention.

The lack of an increase in adverse events should provide some reassurance to providers and patients. Treating to target in other chronic disease areas, such as diabetes mellitus and hypertension, has been associated with excess adverse events (5,6,8). Our results are consistent with other trials of TTT that did not observe an increase in drug toxicity (10).

Resource use differed across time periods, but appeared to be slightly reduced during the TTT intervention. Possibly, using a more systematic TTT approach reduced the need for the use of the resources we measured. We believe that the important result is that no increase in resource use was observed. However, the slight improvement in disease activity that we observed during TTT was possibly associated with a reduced need for resources. This finding may also have been based on chance, because of a relatively small sample size or some degree of misclassification that differed across time periods, producing a biased result; this possibility seems unlikely.

We acknowledge several strengths and limitations. The fact that the same 6 sites were examined before and during the TTT intervention limits the possible confounding bias. However, secular trends during the 18 months of the study period could have impacted the results. Further, 6 rheumatology practices may not be representative of national trends. A standardized review of medical records was performed centrally to reduce interobserver variability, but medical records may not perfectly represent all adverse events and resources used during the 2 periods studied. A larger sample size may have yielded a statistically significant increase in infections.

In conclusion, we did not observe an increase in overall adverse events or resource use associated with TTT. Prior work demonstrates the clinical benefits of TTT, which may translate into a reduction in resource use. But most importantly, patients

Table 2. Adverse events before and during the treat-to-target intervention*

	Before	During	*P*†	Difference, % (95% CI)
Any adverse event‡				
Visits	10.2	8.8	0.41	−1.4 (−3.7, 1.0)
Patients	29.6	25.7	0.45	−3.9 (−10.8, 3.1)
No. per patient, mean	0.43	0.35	0.18	−0.1 (−0.2, 0.1)
Abnormal liver function tests				
Visits	0.8	0.3	0.12	−0.5 (−1.1, 0.1)
Patients	2.8	1.0	0.11	−1.8 (−4.0, 0.3)
No. per patient, mean	0.03	0.01	0.08	−0.02 (−0.1, 0.1)
Rash/oral ulcers/alopecia				
Visits	1.7	0.8	0.07	−0.9 (−1.7, 0.1)
Patients	5.9	2.9	0.07	−3.0 (−6.2, 0.1)
No. per patient, mean	0.07	0.03	0.04	−0.04 (−0.1, 0.00)
Infections				
Visits	2.7	3.8	0.07	1.1 (−0.3, 2.6)
Patients	9.4	12.1	0.18	2.7 (−2.1, 7.5)
No. per patient, mean	0.11	0.14	0.13	0.03 (0.0, 0.1)
Gastrointestinal symptoms				
Visits	2.2	2.2	0.79	0.0 (−1.2, 1.2)
Patients	6.5	7.9	0.37	1.4 (−2.6, 5.4)
No. per patient, mean	0.09	0.08	0.96	0.01 (−0.1, 0.1)
Rheumatology visits				
No. per patient, mean ± SD	4.0 ± 1.4	3.6 ± 1.2	0.02	−0.4 (−0.6, −0.2)
Biologic DMARDs used				
Percent of visits	46.5	42.5	0.76	−3.9 (−7.9, 0.0)
Percent of patients	53.6	49.8	0.73	−3.7 (−11.5, 4.0)
Monitoring laboratory tests§				
Patients	90.0	52.7	<0.001	−37 (−44, −31)
No. per patient, mean ± SD	10.6 ± 6.3	5.1 ± 6.3	<0.001	−5.5 (−6.5, −4.5)
Diagnostic imaging§				
Visits	15.4	8.9	0.005	−6.5 (−9.7, −3.2)
Patients	38.6	12.4	<0.001	−26 (−33, −20)
No. per patient, mean ± SD	1.0 ± 2.0	0.3 ± 1.3	<0.001	−0.7 (−1.0, −0.4)

* Values are the percentage unless indicated otherwise. 95% CI = 95% confidence interval; DMARDs = disease-modifying antirheumatic drugs.
† *P* values from generalized linear mixed models adjusted for site effect and within provider clustering.
‡ Rashes, oral ulcers, alopecia, infections requiring antibiotics, liver toxicity as manifested by abnormal liver function tests and/or abnormal liver imaging, cytopenias as manifested by complete blood counts below the lower limits of normal, renal insufficiency defined as a 50% decrease in creatinine clearance, cancer, and gastrointestinal symptoms (e.g., nausea, vomiting, diarrhea, unexplained weight loss/gain, abdominal pain, or dyspepsia).
§ Monitoring laboratory tests included complete blood count, liver function tests, serum creatinine, and acute phase reactants. Diagnostic imaging included dual x-ray absorptiometry, plain radiographs, computed tomography scans, and magnetic resonance imaging. We did not calculate the percent of visits with laboratory tests because many tests took place between visits.

treated with a TTT approach do not appear to be at risk of increased adverse events.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Solomon 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. Solomon, Katz, Bitton, Fraenkel, Harrold, Smolen, Losina, Lu.

Acquisition of data. Corrigan.

Analysis and interpretation of data. Solomon, Yu, Katz, Losina, Lu.

REFERENCES

1. 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.
2. Solomon DH, Lee SB, Zak A, Corrigan C, Agosti J, Bitton A, et al. Implementation of treat-to-target in rheumatoid arthritis through a learning collaborative: rationale and design of the TRACTION trial. Semin Arthritis Rheum 2016;46:81–7.
3. Solomon DH, Losina E, Lu B, Zak A, Corrigan C, Lee SB, et al. Implementation of treat-to-target in rheumatoid arthritis through a learning collaborative: results of a randomized controlled trial. Arthritis Rheumatol 2017;69:1374–80.
4. Solomon DH, Lu B, Yu Z, Corrigan C, Harrold LR, Smolen JS, et al. Benefits and sustainability of a learning collaborative for implementation of treat-to-target in rheumatoid arthritis: results of a cluster-randomized controlled phase II clinical trial. Arthritis Care Res (Hoboken) 2018;70:1551–6.
5. Gerstein HC, Miller ME, Genuth S, Ismail-Beigi F, Buse JB, Goff DC Jr, et al. Long-term effects of intensive glucose lowering on cardiovascular outcomes. N Engl J Med 2011;364:818–28.
6. Ogihara T. Practitioner's Trial on the Efficacy of Antihypertensive Treatment in the Elderly Hypertension (The PATE-Hypertension Study) in Japan. Am J Hypertens 2000;13 Pt 1:461–7.
7. Ogihara T, Kikuchi K, Matsuoka H, Fujita T, Higaki J, Horiuchi M, et al. The Japanese Society of Hypertension guidelines for the management of hypertension (JSH 2009). Hypertens Res 2009;32:3–107.
8. Xie X, Atkins E, Lv J, Bennett A, Neal B, Ninomiya T, et al. Effects of intensive blood pressure lowering on cardiovascular and renal outcomes: updated systematic review and meta-analysis. Lancet 2016;387:435–43.
9. Qaseem A, Wilt TJ, Kansagara D, Horwitch C, Barry MJ, Forciea MA. Hemoglobin A1c targets for glycemic control with pharmacologic therapy for nonpregnant adults with type 2 diabetes mellitus: a guidance statement update from the American College of Physicians. Ann Intern Med 2018;168:569–76.
10. Schoels M, Knevel R, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas DT, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. Ann Rheum Dis 2010;69:638–43.
11. Harrold LR, Reed GW, Harrington JT, Barr CJ, Saunders KC, Gibofsky A, et al. The rheumatoid arthritis treat-to-target trial: a cluster randomized trial within the Corrona rheumatology network. BMC Musculoskelet Disord 2014;15:389.
12. Tymms K, Zochling J, Scott J, Bird P, Burnet S, de Jager J, et al. Barriers to optimal disease control for rheumatoid arthritis patients with moderate and high disease activity. Arthritis Care Res (Hoboken) 2014;66:190–6.

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1249–1254
DOI 10.1002/acr.23769

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Risk for Herpes Zoster in Tofacitinib-Treated Rheumatoid Arthritis Patients With and Without Concomitant Methotrexate and Glucocorticoids

Jeffrey R. Curtis,[1] Fenglong Xie,[1] Shuo Yang,[1] Sasha Bernatsky,[2] Lang Chen,[1] Huifeng Yun,[1] and Kevin Winthrop[3]

Objective. Increased incidence of herpes zoster (HZ) has been observed with JAK inhibitors such as tofacitinib. However, whether concomitant methotrexate (MTX) and/or glucocorticoids confer additional (additive or multiplicative) risk is unclear. We evaluated HZ risk in tofacitinib users with and without MTX and glucocorticoids.

Methods. Within MarketScan and Medicare data (2011–2016), we identified all patients with rheumatologist-diagnosed rheumatoid arthritis initiating treatment with tofacitinib (index date); demographics and baseline covariates were evaluated in the year prior to the index date. HZ was ascertained using International Classification of Diseases, Ninth Revision or Tenth Revision codes with antiviral drug use (±7 days). Multivariable Cox regression was used to evaluate hazard ratios (HRs) for HZ in tofacitinib users with and without current concomitant MTX and glucocorticoids, controlling for baseline covariates.

Results. We studied 8,030 new tofacitinib users (83.3% women). The mean ± SD age was 60.3 ± 12.6 years. HZ incidence in tofacitinib users was numerically lowest in the absence of glucocorticoids (3.4 per 100 patient-years with MTX; 3.7 per 100 patient-years without MTX). An approximately 2-fold increased incidence of HZ was observed for tofacitinib users receiving either glucocorticoids alone (6.0 per 100 patient-years) or both MTX plus glucocorticoids (6.5 per 100 patient-years). The adjusted HR for HZ in tofacitinib users was unchanged (HR 0.99 [95% confidence interval (95% CI) 0.64–1.54]) when given only with MTX, but was increased (HR 1.96 [95% CI 1.33–2.88]) for tofacitinib plus glucocorticoids. Older age and female sex were also risk factors, while prior vaccination was associated with a strong trend for lower risk.

Conclusion. In tofacitinib users, HZ occurred at a rate of approximately 4% per year and was further doubled with glucocorticoid exposure. Concomitant MTX did not confer additional risk. Zoster vaccination may decrease risk.

INTRODUCTION

JAK inhibitors have proven effective for the treatment of rheumatoid arthritis (RA) (1). Two JAK inhibitor therapies (tofacitinib, baricitinib) are currently approved for use in RA and other conditions (e.g., tofacitinib for ulcerative colitis). The infection risk associated with these therapies is comparable to other biologics and targeted therapies commonly used for RA, with the exception of herpes zoster (HZ), for which risk is increased approximately 2-fold to 3-fold (2,3). Results from the large tofacitinib RA clinical trial program have suggested the possibility that HZ risk might not be increased if tofacitinib is taken without glucocorticoids, and if used as monotherapy (i.e., without methotrexate [MTX] or other disease-modifying antirheumatic drugs). For example, at a monotherapy dosage of 5 mg twice daily, the rate of HZ with tofacitinib was 0.6 per 100 person-years, appreciably lower than the overall HZ rate of 4.0 per 100 across the tofacitinib trial program (4). However, this finding was based on only

Supported by the Patient Centered Outcomes Research Institute.

[1]Jeffrey R. Curtis, MD, MS, MPH, Fenglong Xie, PhD, Shuo Yang, MS, Lang Chen, PhD, Huifeng Yun, PhD: University of Alabama, Birmingham; [2]Sasha Bernatsky, MD, PhD: McGill University, Montreal, Quebec, Canada; [3]Kevin Winthrop, MD, MPH: Oregon Health Sciences University, Portland.

Dr. Curtis has received research grants and consulting fees from AbbVie, Bristol-Myers Squibb, Janssen, Regeneron, Roche, Radius, and UCB (all less than $10,000), and from Amgen, Corrona, Eli Lilly, Myriad, and Pfizer (all greater than $10,000). Dr. Winthrop has received consulting fees from Amgen, Genentech, Bristol-Myers Squibb, Astellas, Janssen, and Galapagos (all less than $10,000), and from AbbVie, Eli Lilly, Pfizer, Roche, and UCB (all greater than $10,000). No other disclosures relevant to this article were reported.

Address correspondence to Jeffrey R. Curtis, MD, MS, MPH, 510 20th Street South, FOT 802, Birmingham, AL 35294. E-mail: jrcurtis@uabmc.edu.

Submitted for publication July 19, 2018; accepted in revised form October 2, 2018.

SIGNIFICANCE & INNOVATIONS

- Although tofacitinib has been demonstrated to increase risk for herpes zoster (HZ) in rheumatoid arthritis (RA), the incremental effect of concomitant methotrexate and glucocorticoid use is controversial.
- Using real-world data from large US health plans, we found that the rate of HZ in new tofacitinib users in the absence of glucocorticoid treatment was 3–4 per 100 person-years. This rate is roughly double the rate of HZ in patients with RA exposed to anti–tumor necrosis factor agents.
- Concomitant glucocorticoid but not methotrexate use further doubled the risk of HZ in tofacitinib users.
- Younger age, male sex, and vaccination with the live virus zoster vaccine were associated with a lower risk for HZ, which may suggest a role for newer vaccination strategies in patients with RA who are at risk.

2 cases occurring in 361 person-years, resulting in uncertainty regarding HZ risk for patients receiving tofacitinib monotherapy.

Unresolved questions in JAK inhibitor users include the practical challenges around the potentially mitigating effects of zoster vaccination (5,6) and uncertainty regarding the importance of other HZ-related risk factors (e.g. age, zoster vaccination). In light of these evidence gaps, we evaluated HZ risk in tofacitinib users, according to whether patients received concomitant glucocorticoids, MTX, both, or neither, in a real-world US RA population.

PATIENTS AND METHODS

Cohort eligibility and study design. RA patients were identified in both Medicare and MarketScan (2011–2016 for both) based on International Classification of Diseases, Ninth Revision or Tenth Revision diagnosis codes from rheumatologists. RA patients initiating treatment with tofacitinib for the first time (based on no prior use in the data set) were identified, and the date of first use was defined as the index date. Continuous insurance coverage in the 12 months prior to the index date was required for all individuals. Patients with diagnosis codes for HIV, malignancy, zoster, and prescriptions of antiviral drugs for zoster prior to initiation were excluded. Patients with diagnosis codes for inflammatory bowel disease, psoriatic arthritis, psoriasis, and ankylosing spondylitis in the baseline year were also excluded.

Exposures of interest. MTX and glucocorticoid treatment concomitant to tofacitinib use (before and after initiation) were evaluated in a time-varying manner and updated for each person-day of observation. All patients were required to continue taking tofacitinib to continue under observation, and the 4 exposure groups of interest were concomitant MTX without oral glucocorticoids, MTX with oral glucocorticoids, oral glucocorticoids without MTX, and neither glucocorticoids nor MTX. A 30-day extension to exposure was added to all 3 exposures (tofacitinib, glucocorticoid, and MTX), and follow-up time was censored if patients discontinued tofacitinib (see below).

Outcome. The study outcome was HZ diagnosis by a physician with initiation of antiviral drug treatment (acyclovir, valacyclovir, famciclovir) within 7 days. The positive predictive value of this approach for HZ case finding was approximately 98%, based on large validation studies that compared this strategy to medical record review and clinical adjudication (7).

Covariates. Based on clinical interest and prior HZ literature, we identified covariates potentially associated with zoster, including demographic factors (age, sex, race/ethnicity [available only in Medicare data]), zoster vaccination (with the live virus vaccine, the only form used in the calendar years in which the analysis was conducted), comorbidities (Table 1), comorbidity indices, and measures of health care utilization, including physician visits and hospitalization. These potentially confounding factors were assessed in the 12-month baseline period prior to the index date. During the ascertainment period, selected covariates (e.g., zoster vaccination) were assessed in the 12-month baseline period plus all available prior data, if more than 12 months of data were available.

Statistical analysis. Descriptive statistics were used to characterize baseline covariates according to exposure groups. Cox proportional hazards models were used to estimate adjusted hazard ratios (HRs), controlling for age, sex, live zoster vaccination, and other potential confounders. Clinically important factors that were statistically significant (*P* values less than 0.05) after multivariable adjustment were retained in the model. Separate results were initially stratified by data source (Medicare versus MarketScan) and conditional on similarity, combined together in a single analysis. The University Institutional Review Board approved the study, and use of the data was governed by data use agreements from the Center for Medicare and Medicaid Services and IBM Watson Health. All analyses were conducted in SAS software, version 9.4.

RESULTS

After satisfying inclusion and exclusion criteria (Figure 1), 8,030 unique patients with RA contributed 5,811 person-years to the analysis. There were a total of 5,369 patients enrolled in Medicare available for analysis, and 2,661 in MarketScan. Characteristics of patients according to concomitant MTX and glucocorticoid exposure are shown in Table 1. Although demographics varied slightly by exposure, there were few clinically meaningful differences. The mean age was approximately 60 years, and

Table 1. Baseline characteristics of tofacitinib patients according to background methotrexate and glucocorticoid use*

Characteristic	No MTX, no GC	With MTX, no GC	No MTX, with GC	With MTX, with GC
No. of person-years: MarketScan/Medicare	691/1,295	283/508	404/952	191/426
Age, mean ± SD years	60.5 ± 12.3	60.3 ± 12.2	60.6 ± 12.2	60.4 ± 12.3
Female	84.4	81.3	81.2	80.1
Race				
White	43.3	41.2	47.5	46.2
African American	5.9	7.4	7.7	8.3
Other/unknown†	50.8	51.4	44.8	45.5
Comorbidities				
Chronic pulmonary disease	22.1	15.7	26.2	19.0
Myocardial infarction	3.1	1.0	5.6	2.7
Coronary heart disease	12.4	9.6	14.5	11.5
Peripheral vascular disorder	5.1	5.3	6.1	5.4
Cerebrovascular disease	4.9	4.9	4.4	5.5
Peptic ulcer disease	1.1	0.8	1.1	0.9
Mild liver disease	0.6	0.3	0.8	0.1
DM without complication	18.8	17.8	18.4	17.0
DM with complication	6.1	4.5	4.4	3.2
Renal disease	6.9	3.5	7.2	5.3
Charlson comorbidity index, mean ± SD‡	0.92 ± 1.32	0.69 ± 1.14	0.99 ± 1.35	0.79 ± 1.16
Baseline medication				
Prednisone-equivalent daily dose, median (IQR)	NA	NA	6.7 (5.0–10.0)	5.0 (5.0–10.0)
Hydroxychloroquine	22.7	22.7	24.1	22.9
Sulfasalazine	9.3	10.5	10.5	9.8
Leflunomide	26.6	8.3	29.6	9.4
NSAIDs	53.1	53.9	49.5	53.4
Narcotics	68.3	63.3	78.4	74.2
Bisphosphonates	9.3	9.8	14.9	16.6
Proton-pump inhibitor	41.3	38.2	46.3	47.7
Statin	37.8	38.2	35.3	36.3
Other lipid lowering drugs	6.8	5.4	6.3	5.9
Visit				
Emergency room	33.9	26.1	42.1	35.6
Nursing home	3.6	3.1	5.0	5.2
Any hospitalization	19.2	16.5	25.1	21.7
No. of ambulatory visits				
<5	13.7	17.0	11.1	12.8
6–10	31.2	31.1	28.6	31.2
>10	55.1	51.9	60.3	56.0
Herpes zoster vaccination§	9.0	9.8	7.5	9.1

* Values are the percentage unless indicated otherwise. Data are shown according to the person-time distribution of the 4 time-varying exposure groups described in the column headers, assessed after the start of follow-up. MTX = methotrexate; GC = glucocorticoids; DM = diabetes mellitus; IQR = interquartile range; NA = not applicable; NSAIDs = nonsteroidal antiinflammatory drugs.
† Race data are not available within MarketScan data, and so all patients represented in this data source were included in this row.
‡ Ignoring rheumatoid arthritis in the Charlson score.
§ With live virus vaccine, using 12-month baseline and all preceding data.

80% of participants were women. Comorbidities were mostly well-balanced across exposure groups, although chronic pulmonary disease was more prevalent in glucocorticoid-exposed individuals. Baseline leflunomide treatment was less common in MTX-exposed individuals, and bisphosphonates were more commonly prescribed for glucocorticoid-exposed individuals.

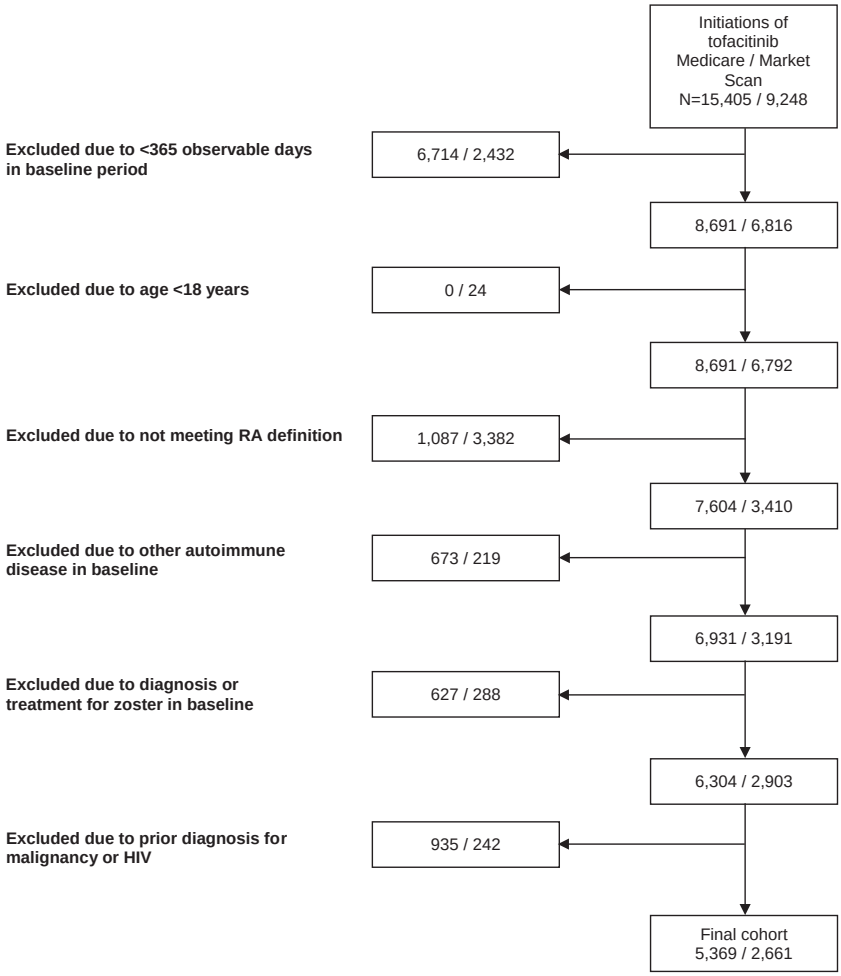

Figure 1. Cohort selection of rheumatoid arthritis (RA) patients initiating tofacitinib, after applying inclusion and exclusion criteria. Data reported as number of patients in Medicare/MarketScan.

In all 4 exposure groups, prior vaccination with the live zoster vaccine was low (<10%).

A total of 222 cases of HZ were observed (n = 156 Medicare, n = 66 MarketScan). Results stratified by data set were similar (data not shown) and were pooled to produce the incidence rates and HRs shown in Figure 2. In tofacitinib users exposed to concomitant glucocorticoids but not to MTX, the crude incidence rate was 6.0 (95% confidence interval [95% CI] 4.9–7.5) per 100 patient-years, which was similar to the rate with concomitant exposure to both glucocorticoids and MTX (6.5 per 100 patient-years). The corresponding crude rate associated with tofacitinib alone was 3.4 (95% CI 2.9–4.6) per 100 patient-years, which was similar to the rate in tofacitinib users who were exposed to MTX but not glucocorticoids (3.7 per 100 patient-years).

After multivariable adjustment, using tofacitinib monotherapy as the reference, exposure to glucocorticoids approximately doubled the HR, but there was no clear increased risk associated with MTX. Older age was a risk factor (adjusted HR 1.11 [95% CI 1.06–1.18]). Both female sex (adjusted HR 1.43 [95% CI 0.97–2.12]) and the live zoster vaccine (adjusted HR 0.60 [95% CI 0.34–1.05]) trended toward significance. No other zoster-related factors were significant, after adjusting for these 3 covariates, and these were the only factors (among all listed in Table 1) that were included in the final multivariable model.

DISCUSSION

In this large cohort of patients with RA, we found that the absolute rates of HZ associated with tofacitinib in the absence of glucocorticoid or MTX exposure were 3.4–3.7 per 100 person-years. Prior literature has shown that incidence rates of HZ associated with biologic therapy generally range from 1.7–2.7 per 100 patient-years (3,8), in RA cohorts where one-third to two-thirds of patients receive concomitant glucocorticoids. The HZ rate with tofacitinib is approximately double (adjusted HR 2.01 [95% CI 1.40–2.88], from a previously published direct comparison [3]) that of patients with RA receiving tumor necrosis factor (TNF) inhibitors or non–TNF inhibitor biologics. In the present analysis, where all

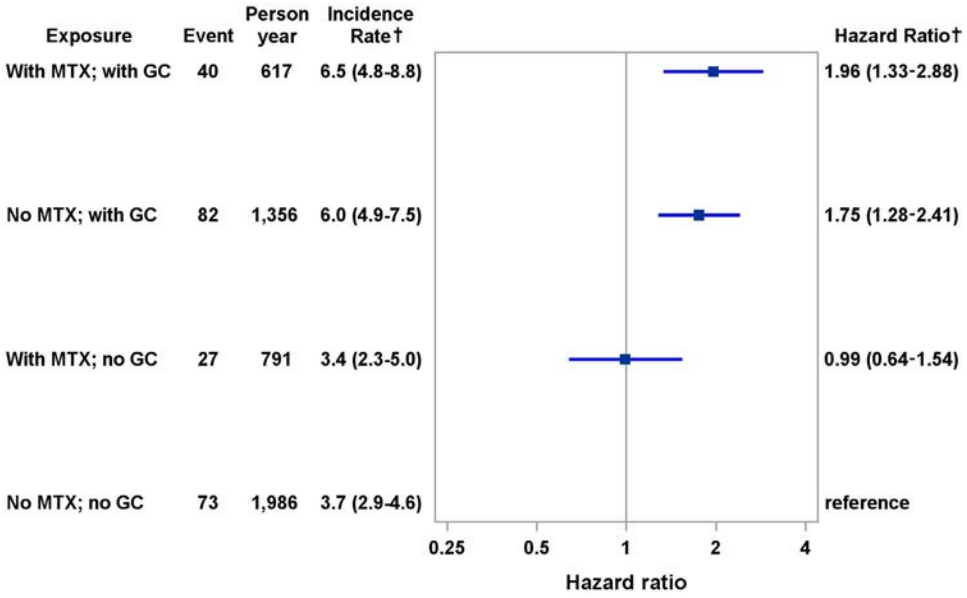

Figure 2. Association of glucocorticoid (GC) and methotrexate (MTX) use with herpes zoster among patients with rheumatoid arthritis treated with tofacitinib. Incidence rates expressed per 100 patient-years. † 95% confidence interval in parentheses.

patients with RA were receiving tofacitinib, concomitant glucocorticoid use further doubled HZ risk (incidence rates 6.0–6.5 per 100 patient-years; adjusted HR 1.96 [95% CI 1.33–2.88]), and there was no incremental risk associated with concomitant MTX use.

Our results add to the growing body of literature surrounding the risk for HZ associated with JAK inhibition and other risk factors in patients with RA. Although data from the tofacitinib clinical trial program suggested that tofacitinib might not elevate zoster risk in the absence of concomitant glucocorticoid and/or MTX exposure, the absolute rates of HZ in our real-world data set mirror the rate of 4.0 events per 100 patient-years from the tofacitinib trial program and its long-term extension (2,4). The effect of glucocorticoid exposure in the tofacitinib-treated patients in this study (a 1.96-fold relative increase) was quite similar to the corresponding effect in patients with RA receiving conventional synthetic disease-modifying antirheumatic drugs or biologic therapies (relative risk 1.4–2.4) (3,8). The effect of advancing age was a significant risk factor, as was sex. Indeed, prior HZ literature in RA has shown age to be a much stronger zoster risk factor than we observed (9), consistent with the interpretation that JAK inhibitor use may somewhat dominate other zoster risk factors. Female sex has been associated with higher HZ rates (possibly related to differences in susceptibility to viruses or in vaccination responses), and more pain complications, possibly driven by an estrogen-sensitive pain mechanism.

We saw a strong trend for decreased risk related to vaccination with the live agent (Zostavax [Merck]); the concern with this form of vaccination is that any live vaccination is potentially dangerous in patients receiving potent immunosuppression. An ongoing randomized controlled trial of the live virus vaccine in patients with RA and other autoimmune and inflammatory conditions (e.g., psoriasis, psoriatic arthritis) receiving anti-TNF therapy is being undertaken (https://clinicaltrials.gov/ct2/show/NCT02538341), with no safety signals identified in >500 patients randomized to date. However, this trial is not enrolling patients receiving JAK inhibitor therapy. A new adjuvant-based HZ vaccine (Shingrix [GlaxoSmithKline]) has recently become available and could be given even to JAK-inhibitor users, although it has not been tested in RA or in patients with other autoimmune conditions receiving immunosuppressive or immunomodulatory therapies. The risks for disease flare, and potentially problematic tolerability related to a relatively high incidence of grade 3 (severe) systemic reactogenicity, may limit enthusiasm until specific data in an RA population are available. Efforts to maximize zoster vaccination uptake are critical once these 2 open questions with this new vaccine have been resolved.

Strengths of our study include a large sample of 8,030 patients with RA and the characterization of exposure at a person-day level, which accounts for short-term variation in adherence and medication use. Limitations of our study include use of administrative health plan data without medical record confirmation of HZ cases. Our approach for HZ case ascertainment has been shown to have extremely high validity, with a PPV of 98% (7). We required concomitant antiviral treatment to satisfy our zoster case definition, consistent with the approach used in the validation study for our zoster reactivation case definition. This requirement may have somewhat decreased the absolute HZ incidence rate in our cohort. However, we preferred erring on the side of specificity rather than sensitivity for our case definition. We excluded

patients with risk factors that would put them at unique risk for HZ reactivation (e.g., malignancy, HIV), and we had a relatively short follow-up time available (given that tofacitinib was approved in the US in late 2012).

As another potential limitation, we did not have clinical information on RA severity and related variables; however, prior evaluations of HZ in RA registry data have identified few strong HZ risk factors except for age, sex, comorbidities, and treatment with RA medication (all of which were available in the data source used in this study). Clinical factors not available in this data source, such as functional status, duration of RA, and RA disease activity, have not been shown to confound the association between RA treatments and HZ reactivation (10). Additionally, we classified glucocorticoid and MTX exposure as time varying and updated it for each person-day of observation. We did not quantify MTX and glucocorticoid dose; previous work has examined the dose of both medications in tofacitinib-treated patients, and the dose of neither medication had a clear association with HZ incidence. We recognize that this finding was possibly a statistical power issue, or a misclassification of glucocorticoid dose that in the tofacitinib trial program was well characterized at baseline but may not have been systematically updated, if patients changed dose in the long-term extension (4). Except for our main medication exposures that were time varying, we controlled only for baseline factors, recognizing that time-varying factors (e.g., disease activity) might be influential. However, adjusting for factors that change after the start of follow-up can result in the important problem of over-adjustment and inappropriately controlling for causal intermediates (11), and is generally avoided. We note that zoster vaccination was assessed in the baseline period, which had a minimum of 12 months, but used more antecedent data if they were available. However, left censoring may have misclassified the use of the zoster vaccination prior to this period of time, although based on very limited uptake in rheumatic disease populations through 2011 (12), this under-ascertainment was likely small and would not be expected to vary by exposure group.

In conclusion, in the absence of glucocorticoids or MTX, we observed an incidence rate for HZ in tofacitinib users of almost 4% per year; the use of glucocorticoids approximately doubled that risk. Female sex and older age were other nonmodifiable risk factors. Given the strong trend for a protective role for vaccination, this strategy should be considered in all at-risk patients without contraindications. However, enthusiasm for use of this vaccine may be somewhat tempered until data are available to support the safety and tolerability of the new zoster vaccine in patients who have autoimmune conditions like RA and who may be at risk for flare related to the potent adjuvant.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Curtis 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. Curtis, Xie, Bernatsky, Chen, Yun, Winthrop.

Acquisition of data. Curtis, Xie, Yang, Chen.

Analysis and interpretation of data. Curtis, Xie, Yang, Bernatsky, Chen, Yun, Winthrop.

REFERENCES

1. Fleischmann R, Cutolo M, Genovese MC, Lee EB, Kanik KS, Sadis S, et al. Phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) or adalimumab monotherapy versus placebo in patients with active rheumatoid arthritis with an inadequate response to disease-modifying antirheumatic drugs. Arthritis Rheum 2012;64:617–29.
2. Winthrop KL, Yamanaka H, Valdez H, Mortensen E, Chew R, Krishnaswami S, et al. Herpes zoster and tofacitinib therapy in patients with rheumatoid arthritis. Arthritis Rheumatol 2014;66:2675–84.
3. Curtis JR, Xie F, Yun H, Bernatsky S, Winthrop KL. Real-world comparative risks of herpes virus infections in tofacitinib and biologic-treated patients with rheumatoid arthritis. Ann Rheum Dis 2016;75:1843–7.
4. Winthrop KL, Curtis JR, Lindsey S, Tanaka Y, Valdez H, Hirose T, et al. Herpes zoster and tofacitinib: clinical outcomes and the risk of concomitant therapy. Arthritis Rheumatol 2017;69:1960–8.
5. Winthrop KL, Wouters AG, Choy EH, Soma K, Hodge JA, Nduaka CI, et al. The safety and immunogenicity of live zoster vaccination in patients with rheumatoid arthritis before starting tofacitinib: a randomized phase II trial. Arthritis Rheumatol 2017;69:1969–77.
6. Levin MJ, Bresnitz E, Popmihajlov Z, Weinberg A, Liaw KL, Willis E, et al. Studies with herpes zoster vaccines in immune compromised patients. Expert Rev Vaccines 2017;16:1217–30.
7. Baxter R, Bartlett J, Fireman B, Marks M, Hansen J, Lewis E, et al. Long-term effectiveness of the live zoster vaccine in preventing shingles: a cohort study. Am J Epidemiol 2018;187:161–9.
8. Yun H, Xie F, Delzell E, Chen L, Levitan EB, Lewis JD, et al. Risks of herpes zoster in patients with rheumatoid arthritis according to biologic disease-modifying therapy. Arthritis Care Res (Hoboken) 2015;67:731–6.
9. Strangfeld A, Listing J, Herzer P, Liebhaber A, Rockwitz K, Richter C, et al. Risk of herpes zoster in patients with rheumatoid arthritis treated with anti-TNF-α agents. JAMA 2009;301:737–44.
10. Pappas DA, Hooper MM, Kremer JM, Reed G, Shan Y, Wenkert D, et al. Herpes zoster reactivation in patients with rheumatoid arthritis: analysis of disease characteristics and disease-modifying antirheumatic drugs. Arthritis Care Res (Hoboken) 2015;67:1671–8.
11. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology 2009;20:488–95.
12. Zhang J, Xie F, Delzell E, Chen L, Winthrop KL, Lewis JD, et al. Association between vaccination for herpes zoster and risk of herpes zoster infection among older patients with selected immune-mediated diseases. JAMA 2012;308:43–9.

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1255–1263
DOI 10.1002/acr.23748

American College of Rheumatology
Empowering Rheumatology Professionals

Outpatient Engagement and Predicted Risk of Suicide Attempts in Fibromyalgia

Lindsey C. McKernan, Matthew C. Lenert, Leslie J. Crofford, and Colin G. Walsh

Objective. Patients with fibromyalgia (FM) are 10 times more likely to die by suicide than the general population. The purpose of this study was to externally validate published models predicting suicidal ideation and suicide attempts in patients with FM and to identify interpretable risk and protective factors for suicidality unique to FM.

Methods. This was a case–control study of large-scale electronic health record data collected from 1998 to 2017, identifying FM cases with validated Phenotype KnowledgeBase criteria. Model performance was measured through discrimination, including the receiver operating area under the curve (AUC), sensitivity, and specificity, and through calibration, including calibration plots. Risk factors were selected by L1 penalized regression with bootstrapping for both outcomes. Secondary utilization analyses converted time-based billing codes to equivalent minutes to estimate face-to-face provider contact.

Results. We identified 8,879 patients with FM, with 34 known suicide attempts and 96 documented cases of suicidal ideation. External validity was good for both suicidal ideation (AUC 0.80) and attempts (AUC 0.82) with excellent calibration. Risk factors specific to suicidal ideation included polysomatic symptoms such as fatigue (odds ratio [OR] 1.29 [95% confidence interval (95% CI) 1.25–1.32]), dizziness (OR 1.25 [95% CI 1.22–1.28]), and weakness (OR 1.17 [95% CI 1.15–1.19]). Risk factors specific to suicide attempt included obesity (OR 1.18 [95% CI 1.10–1.27]) and drug dependence (OR 1.15 [95% CI 1.12–1.18]). Per utilization analyses, those patients with FM and no suicidal ideation spent 3.5 times more time in follow-up annually, and those without documented suicide attempts spent more than 40 times more time face-to-face with providers annually.

Conclusion. This is the first study to successfully apply machine learning to reliably detect suicidality in patients with FM, identifying novel risk factors for suicidality and highlighting outpatient engagement as a protective factor against suicide.

INTRODUCTION

Every day, 120 people die from suicide in the US (1,2). At a minimum, the presence of chronic pain doubles the suicide risk (3), and evidence suggests that specific pain disorders, such as fibromyalgia (FM) further elevate suicide risk (4,5). FM is characterized by the presence of widespread pain with cognitive dysfunction, fatigue, and sleep difficulty (6). Collectively, patients with FM are up to 10.5 times more at risk of death from suicide than the general population (7), and 3.3 times more at risk than other chronic pain patients (8). Similarly, patients with FM have high rates of suicidal ideation, thoughts, and behaviors (SITBs), including suicidal ideation (33–48%), ideation with active intent (6–8%) (9,10), and nonfatal suicide attempts by poisoning (17%) (9). The risk factors for SITBs in patients with FM are difficult to study prospectively because of underreporting worldwide (11), stigma (12), or lack of health care access (12). Moreover, SITBs in patients with FM may be misclassified as accidental deaths (8) (e.g., car accidents) if they are reported at all (11).

A recent comprehensive review indicated that the presence of chronic pain alone, regardless of demographics, pain severity, or mental health, doubles suicide risk. Further, this review suggested

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Center for Advancing Translational Sciences or the National Institutes of Health.

Supported by the Vanderbilt Patient Centered Outcomes Research Education and Training Initiative and the US Department of Health and Human Services Agency for Healthcare Research and Quality (grant 6K12-HS-022990-04). The Synthetic Derivative is supported by the NIH/National Center for Advancing Translational Sciences (grant UL1-RR-024975/RR/NCRR).

Lindsey C. McKernan, PhD, MPH, Matthew C. Lenert, BA, Leslie J. Crofford, MD, Colin G. Walsh, MD, MA: Vanderbilt University Medical Center, Nashville, Tennessee.

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

Address correspondence to Lindsey C. McKernan, PhD, MPH, Departments of Psychiatry and Behavioral Sciences, Physical Medicine and Rehabilitation, Vanderbilt University Medical Center, Osher Center for Integrative Medicine, 3401 West End Avenue, Suite 380, Nashville, TN 37203. E-mail: lindsey.mckernan@vumc.org.

Submitted for publication April 10, 2018; accepted in revised form September 4, 2018.

SIGNIFICANCE & INNOVATIONS

- To the best of our knowledge, this is the first study to successfully apply machine learning to suicidality in fibromyalgia, identifying novel risk factors for both suicidal ideation and suicide attempts.
- Risk factors for suicidal ideation included polysomatic symptoms such as fatigue, dizziness, and weakness.
- Risk factors for suicide attempt include drug dependence and obesity.
- Fibromyalgia patients without documented suicidality spent up to 40× more time with providers annually, highlighting the importance of outpatient engagement as a protective factor against suicide.

that general demographic risk factors for suicide (e.g., sex, age, marital status, education level) may not translate to chronic pain populations, and that other modifiable factors specific to pain may increase the suicide risk (3). Very little, and conflicting, information exists detailing risk factors for suicidality in FM patients, in part because of reliance on small, prospective cohorts as the mainstay of study in this domain. In FM populations specifically, risk factors identified thus far include pain severity (5,8,13,14), widespread pain (15), disease severity (7), younger age (8), depression/anxiety (9,10,16), sleep dysfunction (4,9), and mood disorder (17). A point of debate is the relative contribution of pain severity versus psychiatric comorbidity to suicidality in patients with FM. Initial differing findings may suggest that the presence of psychiatric comorbidity does not fully explain the increased risk of suicide in patients with FM and may only apply to some patients, and that both pain-specific and general risk factors for elevated suicide risk need to be considered when examining suicide risk in patients with FM (18). Limitations of existing research include generalizability, small sample sizes, self-reported symptoms and diagnoses, low response rates, or inability to assess individuals over time. Similar to other conditions, risk factors for suicidal ideation and suicidal attempts may differ (19). The preponderance of studies occur in tertiary specialty clinic settings and may not reflect settings in which large quantities of health care are delivered, such as primary care. Last, population characteristics that are common in patients with FM and known to elevate suicide risk, such as posttraumatic stress (20) and the presence of multiple pain conditions (17), have yet to be investigated.

One path to study SITBs remains large-scale retrospective analyses of clinical electronic health record (EHR) data, including predictive modeling. For example, we have validated predictive models of suicide attempt risk in a broad, heterogeneous population of adults (21) and adolescents (22) at a large academic medical center. Generalized models like these may be personalized to high-risk populations (e.g., FM) to predict risk before harm occurs and to identify risk and protective factors specific to these groups. Research to date has yet to assess risk factors for SITBs concurrently or longitudinally in routinely collected EHR data in patients with FM. Whether these general algorithms will accurately identify risk in FM, or whether the resultant risk patterns differ in patients with FM compared to other groups, is unknown. Such risk patterns, once quantified, may suggest targets of clinical intervention.

Studies of clinical EHR data are well suited to address these knowledge gaps. Moreover, they also present opportunities to develop clinical tools to identify and prevent SITBs using these same data. To bridge accurate risk identification with interpretable, actionable intervention, identifying who is at risk and considering why risk profiles look as they do is paramount. We hypothesize that in translating existing models of suicide risk to an FM population, novel predictors specific to this cohort will need to be considered. Existing evidence shows that general risk factors for suicide do not always translate to chronic pain populations (3), and that pain subpopulations may have different risk factors for suicidality and need to be studied separately to enhance prevention efforts (18).

Coupling literature and domain knowledge of SITBs in patients with FM with validated machine learning algorithms of suicide attempt risk (21), the purpose of this investigation was to assess the external validity of published models in predicting suicidal ideation and attempts in FM and to use novel analyses to identify interpretable risk profiles unique to FM. These latter findings will inform prevention strategies directly.

MATERIALS AND METHODS

Clinical predictive modeling/clinical phenotyping data collection (adapted from Walsh et al 2017 [21]). Clinical data were collected from the EHR at Vanderbilt University Medical Center (VUMC) using the de-identified clinical data repository known as the Synthetic Derivative (23). This repository includes clinical data such as diagnoses, demographics, clinical text, laboratory values, and other information collected during 20 years at VUMC on >2.8 million individuals, with rich data available on >1 million patient lives.

Model development of the general suicide attempt risk algorithm has been previously described (21). Briefly, candidate charts were identified using self-injury International Classification of Diseases, Ninth Revision codes (E95x.xx) for adults in the Synthetic Derivative and labeled by multiple experts to establish a reliable gold standard. These 3,250 cases of suicide attempt were compared to a control cohort of 12,695 adults drawn from the general population of VUMC. These charts were identified from a minimum of 3 visits in the medical center and patients age >18 years. This cohort defined the published algorithm reported previously with an internal validation C statistic as high as 0.92 to predict a suicide attempt in 7 days. The original algorithm was designed to predict a suicide attempt. In this investigation, we sought to extend its reach to predict both suicidal ideation without

an attempt as well as a suicide attempt in an FM cohort with no additional model training or updating.

External validation data collection for this study. We defined the FM cohort through a validated phenotype publicly available in Phenotype KnowledgeBase (PheKB) (24). The phenotype uses a combination of diagnostic codes and text phrases to identify cases of true FM. We applied this phenotype to the VUMC population and selected only those meeting PheKB criteria for FM and with ≥3 visits to VUMC during a period of ≥6 months.

To ensure true external validity testing on the FM cohort, we returned to the initial modeling experiment and removed any patients in the FM cohort from the general model algorithm training set. The general algorithm was then refitted and internally validated after assurance that there were no patients in common in the general cohort.

FM-specific feature selection. We combined domain knowledge and existing research to inform feature selection in an FM-specific model of SITBs. First, we conducted a review of existing literature to extrapolate known risk factors (features), and with informatics coauthors (MCL and CGW), incorporated those features not in the existing model as new risk factors. Second, we used clinical expertise from authors who had direct experience working with FM and SITBs (LCM and LJC) to include additional features derived from patient-provider interactions not accounted for in previous research or the existing algorithm. Briefly, we added model features using regular expressions from notes and diagnostic codes relevant to FM. These features included post-traumatic stress, trauma exposure, violence exposure, abuse exposure (sexual and nonsexual), sleep dysfunction, marijuana use, abdominal pain, and polysomatic symptoms. Supplementary Appendix 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23748/abstract, shows a table of novel features added to the model, their basis, and specific codes used to derive them.

Data preprocessing and missing data handling. Clinical data were preprocessed as previously reported and as described in Supplementary Appendix 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.23748/abstract, to support external validity testing and replication here. Missing data were rare because the variables measured as counts (diagnoses, medications, and visits) were imputed to zeroes if not present. Zip codes were missing in 7.5% of charts and race was missing in 0.05% of charts. Multiple imputation was used to impute missing values in those instances (25).

External validity testing of the general algorithm on the FM cohort. Data on the FM cohort were preprocessed identically to the internal validation sample and included 2 outcomes and multiple time points of prediction: suicidal ideation and attempts at 30 days from the last clinical encounter. The general suicide attempt algorithm was then applied to these data to obtain a posterior probability of suicide attempt risk. This predicted probability was used to validate both suicidal ideation and suicide attempt outcomes in this cohort. External validity testing not only included testing an algorithm on a new set of input data but also testing its generalizability to predict different outcomes. For example, the Charlson Comorbidity Index, a common risk score originally validated to assess mortality risk, has been used in a panoply of new predictive tasks, such as hospital readmission risk. We would not clinically equate suicidal ideation and behaviors such as suicide attempts, but we hypothesized that some shared risk factors between ideation and attempts suggest that an algorithm designed to predict suicide attempts specifically could also be generalized to predict suicidal ideation.

Recalibration in external validation. The general model development cohort was enriched to a ratio of 4 controls for every case, in order to optimize model performance. Because outcome prevalence in FM (approximately 1%) was different than that in the enriched, internal development set (approximately 25%), recalibration of external predictions was performed using logistic calibration as we have used in other predictive domains (26). This method passes the predictions through a logit function trained on the prevalence in the new setting, in this case the FM cohorts. The resultant predictions are subsequently calibrated properly to indicate that a 40% risk of an outcome correlates with 4 of 10 similar individuals in the new setting actually having that outcome. This latter example is the definition of good calibration, whether or not predictions reflect real outcome rates.

Development and validation of the novel explanatory FM suicide risk algorithm. We used the bootstrapped L-1 penalized regression (BoLASSO algorithm [27]) with 2 levels of bootstrapping to gain insight into which factors may have the most influence on suicide risk for FM patients. Briefly, L-1 penalized regression (LASSO algorithm) is well accepted for its ability to select a small number of important predictors across complex data. The BoLASSO enhances this technique with resampling to yield a set of influential predictors and an ability to obtain interpretable test statistics for those same predictors. We conservatively tuned the BoLASSO to select only those features that were chosen in 80% of bootstraps. Full details can be reviewed in Supplementary Appendix 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.23748/abstract.

Performance evaluation and utilization analyses. Performance was measured through discrimination, including the receiver operating area under the curve (AUC), sensitivity/recall, specificity, and precision and also through calibration metrics, including calibration plots, calibration slope/intercept, and scoring

rules. With preliminary results identifying differential health care utilization as a protective factor of SITBs, we conducted a secondary analysis of health care encounters in study cohorts. We counted evaluation and management Current Procedural Terminology (CPT) codes, 99211–99215, health and behavior codes, 96150–96154, and outpatient psychiatry CPT codes, 90791–90792, 90832–90840, 90846–90849, and 90853, for each study cohort. We linked evaluation and management codes to equivalent minutes in time-based billing to estimate time spent in follow-up.

RESULTS

Using the validated PheKB definition of FM (24), we identified 14,430 patients from January 1998 to November 2017 with the phenotype. After censoring only those patients with ≥3 visits during a 6-month period, there were 8,879 patients with 34 known attempts, 0.4% outcome prevalence, and 96 documented cases of suicidal ideation, 1.1% outcome prevalence. The baseline characteristics of these cohorts are shown in Table 1.

External validation of the published model. The general suicide attempt prediction model predicted both suicidal ideation and suicide attempts in a novel FM cohort, with good discrimination. The AUCs were 0.82 for suicide attempts and 0.8 for suicidal ideation (Figure 1). Sensitivity and specificity varied, based on the threshold of case positivity, and ranged from 0.01 to 1 for specificity for both outcomes, from 0 to 1 for sensitivity for attempts, and from 0 to 0.99 for ideation. Precision and recall/sensitivity were also assessed, and precision was low

Table 1. Baseline patient characteristics*

Characteristic	Suicide attempts		Suicidal ideation	
	Cases (n = 34)	Controls (n = 8,845)	Cases (n = 96)	Controls (n = 8,788)
Sex				
Male	3 (9)	805 (9)	15 (16)	796 (9)
Female	31 (91)	8,040 (90)	81 (84)	7,992 (91)
Race				
White	30 (88)	7,768 (88)	86 (90)	7,719 (88)
African American	4 (12)	796 (9)	10 (10)	788 (9)
Asian	0 (0)	141 (1)	0 (0)	42 (0.5)
Alaskan/Native American	0 (0)	0 (0)	0 (0)	20 (0.2)
Declined to respond	0 (0)	0 (0)	0 (0)	140 (1)
Unknown/not recorded	0 (0)	140 (1)	0 (0)	79 (0.8)
Age, median ± SD years	45 ± 9.0	57 ± 14.2	50 ± 13.7	57 ± 14.1
Utilization mix in preceding year				
Outpatient visits, mean (%)	7.1 (73)	14.6 (66)	23.3 (60)	14.5 (65)
Inpatient visits, mean (%)	5.8 (62)	1.3 (84)	10.7 (72)	1.2 (84)
Comorbidity mix				
ADD with hyperactivity	0 (0)	149 (1.7)	10 (10)	466 (5)
Post-traumatic stress disorder	6 (18)	525 (6)	31 (32)	498 (6)
Oppositional defiant disorder	0 (0)	3 (0.03)	0 (0)	2 (0.02)
Generalized anxiety disorder	3 (9)	679 (8)	13 (14)	633 (7)
Asthma	8 (24)	1,693 (19)	25 (26)	1,676 (19)
Episodic mood disorder	3 (9)	446 (5)	96 (100)	431 (5)
Bipolar	6 (18)	522 (6)	35 (36)	496 (6)
Schizophrenia	1 (3)	91 (1)	8 (8)	82 (1)
Congestive heart failure	1 (3)	643 (7)	12 (13)	639 (7)
Diabetes mellitus	0 (0)	84 (1)	2 (2)	80 (1)
COPD	1 (3)	467 (5)	11 (11)	464 (5)
Malignancy	0 (0)	89 (1)	0 (0)	89 (1)
Liver disease	0 (0)	29 (0.3)	0 (0)	29 (0.3)

* Values are the number (%) unless indicated otherwise. ADD = attention deficit disorder; COPD = chronic obstructive pulmonary disease.

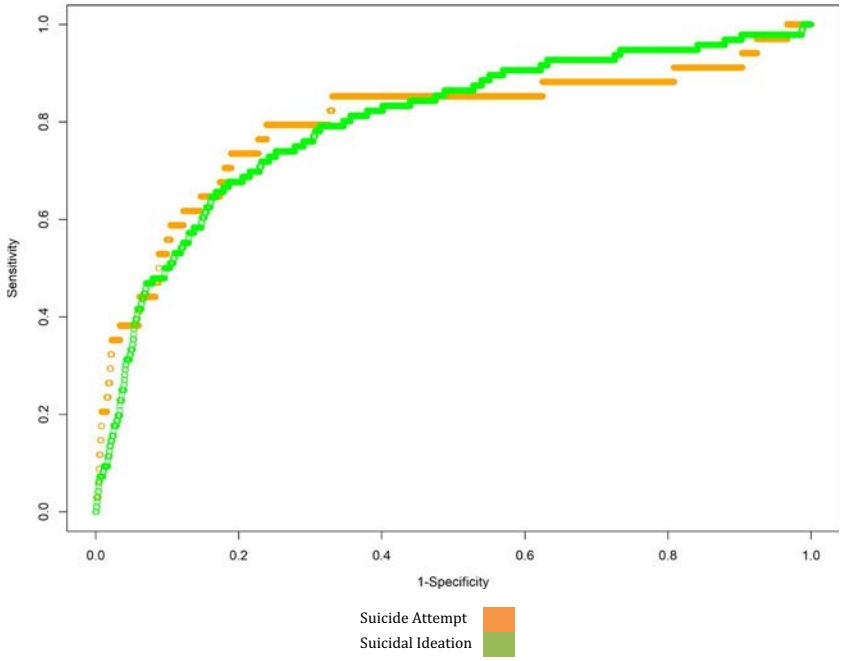

Figure 1. External validity discrimination performance: receiver operating characteristic curves.

for both outcomes, given the extreme case imbalance in this context. Maximum precision was 0.08 for attempts and 0.14 for suicidal ideation (Figure 2 shows precision-recall curves).

Calibration is an important metric to illustrate whether predicted probabilities reflect true prevalence in a population. The externally valid predictions demonstrated excellent calibration performance after recalibration to the outcome prevalence in the novel FM cohort. Risk concentration is the proportion of cases of ideation or attempts by binned quantile of risk. The proportions of cases of suicidal ideation by predicted bin of risk are shown in Figure 3 and indicate that the majority of cases of ideation fall into the highest predicted bins of risk, as anticipated.

Risk factors of suicidal ideation and suicide attempt in FM. The BoLASSO algorithm selected both risk and protective factors for both outcomes. Risk factors are summarized by category in Table 2. The risk categories for suicidal ideation included polysomatic symptoms (fatigue odds ratio [OR] 1.29 [95% confidence interval (95% CI) 1.25–1.32], dizziness OR 1.25 [95% CI 1.22–1.28], and weakness OR 1.17 [95% CI 1.15–1.19]), and serious and persistent mental

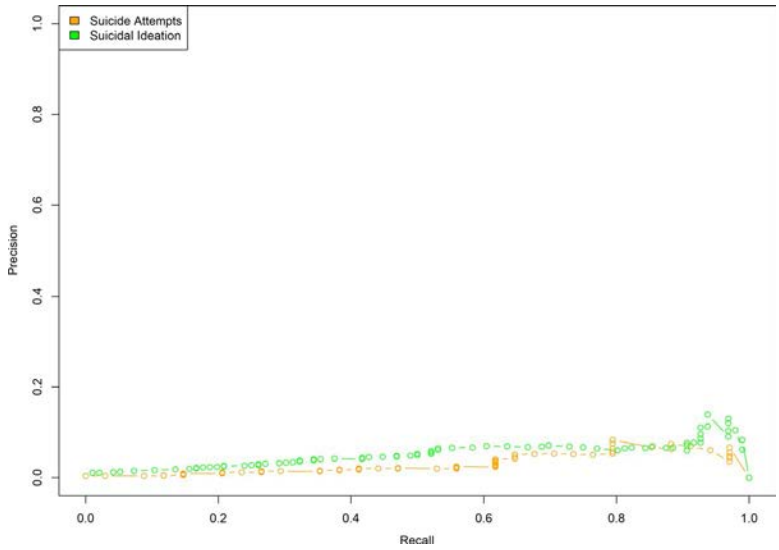

Figure 2. Precision-recall curves.

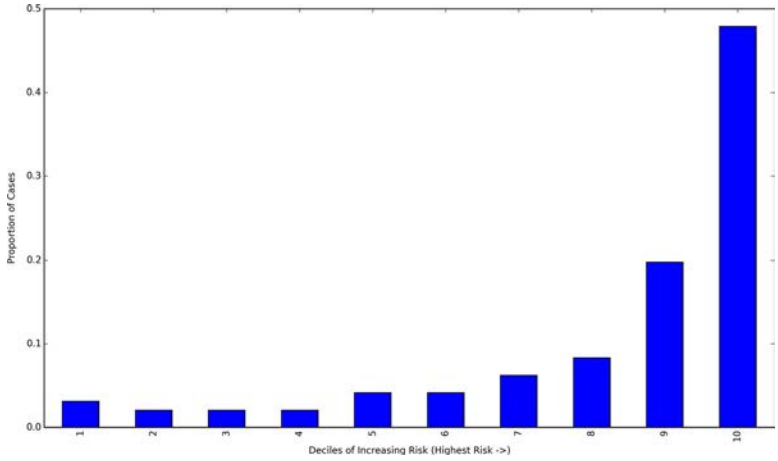

Figure 3. Proportions of cases of ideation by predicted bin of risk.

illness (e.g., bipolar disorder not otherwise specified OR 1.18 [95% CI 1.17–1.20] and inpatient utilization OR 1.5 [95% CI 1.46–1.53]). Concomitant categories for suicide attempt were drug dependence (e.g., cocaine dependence) OR 1.18 (95% CI 1.1–1.27), obesity (body mass index 50–59) OR 1.15 (95% CI 1.12–1.18), mental illness (e.g., recurrent depression with psychosis) OR 1.12 (95% CI 1.07–1.18), and inpatient utilization OR 1.32 (95% CI 1.27–1.36).

We note that commonly held risk factors such as posttraumatic stress disorder, histories of sexual abuse and trauma, and medications like benzodiazepines were all included as potential predictors of SITBs. However, because of the conservatism of our approach to only report those predictors selected >80% of the time, they were not finally selected in the models summarized here.

Utilization analysis. We tallied minutes spent in outpatient follow-up in the cohorts in our study and determined that for suicidal ideation, those patients with FM who did not have suicidal ideation spent 3.5× more time in follow-up per year than those with documented suicidal ideation. This ratio was even more pronounced for suicide attempters. Individuals with FM who did not have documented suicide attempts spent more than 40× more time with outpatient providers than those with documented attempts. We then assessed psychiatry and mental health behavior and intervention codes (CPT codes 90791, 90846, and 96150–96154) and determined that, while these were small proportions of the overall study cohort (0.1–2%), none of the patients with these encounters had a documented suicide attempt. Notably, the majority of mental health behavior and intervention codes were billed for those patients with ideation but none with subsequent attempts. These data may suggest a straightforward albeit nontrivial prevention strategy, enabled by predictive models that suggest patients for whom outpatient engagement should be established.

DISCUSSION

This study is the first to our knowledge to apply machine learning to suicidality in FM in the context of clinical domain expertise to obtain interpretable patterns of risk. We demonstrated that generalizable predictive models of SITB risk perform well in predicting SITBs (attempts: AUC approximately 0.82, maximum precision 0.08; ideation: AUC 0.80, maximum precision 0.14). Notably, the initial algorithm validated externally across a novel cohort and for 2 different outcomes with no further model refitting. In other words, a model predicting suicide attempts alone performed well to predict both suicidal ideation and suicide attempts.

Adding disease-specific risk factors in a rigorous statistical experimental design, the BoLASSO, highlighted different risk patterns for suicidal ideation versus suicide attempts in FM. Both ideation and attempt risk were conferred by younger age, serious and persistent mental illness, comorbid medical illness, and frequent inpatient admission. Polysomatic symptoms (e.g., fatigue, dizziness, and weakness) typified risk of suicidal ideation, while drug dependence and comorbid obesity increased the risk of suicide attempt. Of note, we did not have the capability to capture pain severity or duration in this context, though doing so remains a consideration for future work.

This analysis suggests that unique profiles of suicide risk exist in FM. In our sample, profiles of suicide risk in FM combine those risks indicated in previous investigations in the general population (i.e., obesity, younger age, frequent inpatient admission, severe and persistent mental illness), in chronic pain (illicit drug use, comorbid health conditions), and in FM (mood disorder) with novel risk factors identified in this study (polysomatic symptoms, including fatigue, dizziness, and weakness). Furthermore, our investigation shows that patterns of suicide risk differ for suicidal ideation and suicide attempt in FM, prompting further investigation.

Table 2. Risk factors for attempts and suicidal ideation, thoughts, and behaviors*

	Factor source	Suicide attempts	Suicidal ideation	Examples
Comorbidity				
Antiinfective drugs	Medication list	1.16–1.20	–	Ciprofloxacin, gentamycin
Thiazolidinedione	Medication list	1.12–1.15	–	Pioglitazone
Non-nucleoside reverse transcriptase	Medication list	1.11–1.14	–	Etravirine
Selective estrogen receptor modulator	Medication list	1.05–1.11	–	Raloxifene
Antiepileptic drugs, hydantoin derivatives	Medication list	1.01–1.06	–	Phenytoin, fosphenytoin
Atrial flutter	ICD-10 diagnosis	–	1.11–1.15	ICD-9: 427.32
Obstructive chronic bronchitis without exacerbation	ICD-10 diagnosis	–	1.09–1.13	ICD-9: e.g., 491.21–491.22
Diabetic retinopathy	ICD-10 diagnosis	–	1.08–1.14	ICD-9: e.g., 362.01
Chemotherapy, pyrimidine analogs	Medication list	–	1.08–1.11	Gemcitabine, fluorouracil, cytarabine
Ulcer of lower extremity	ICD-10 diagnosis	–	1.05–1.14	ICD-9: 707.10
Ulcer of ankle	ICD-10 diagnosis	–	1.05–1.11	ICD-9: 707.13
History of septic shock	ICD-10 diagnosis	–	1.04–1.12	ICD-9: 785.52
Diabetes mellitus w/other specified manifestations, type I	ICD-10 diagnosis	–	1.04–1.11	ICD-9: 250.81
Immune thrombocytopenic purpura	ICD-10 diagnosis	–	1.03–1.14	ICD-9: 287.31
Blood clots, AC DVT/embolism in lower extremities	ICD-10 diagnosis	–	1.03–1.07	ICD-9: 453.41
Cerebral embolism with infarction	ICD-10 diagnosis	–	1.01–1.09	ICD-9: 434.91
Hypersensitivity angiitis	ICD-10 diagnosis	–	1.01–1.09	ICD-9: 446.20
Cocaine dependence, unspecified	ICD-10 diagnosis	1.10–1.27	–	ICD-9: 304.20
Inpatient visits within the past year	Visit count	1.27–1.36	1.46–1.53	–
Mental illness				
Borderline personality disorder	ICD-10 diagnosis	1.16–1.20	–	ICD-9: 301.83
Indole derivatives (antipsychotic drugs)	ICD-10 diagnosis	1.10–1.15	–	Clomipramine, imipramine
Recurrent depression, w/psychotic features	ICD-10 diagnosis	1.07–1.18	–	ICD-9: e.g., 296.31, 296.16
Bipolar disorder, not otherwise specified	ICD-10 diagnosis	–	1.17–1.20	ICD-9: e.g., 296.80
Bipolar I disorder, manic w/psychotic features	ICD-10 diagnosis	–	1.13–1.17	ICD-9: e.g., 296.43–296.44
Monoamine oxidase inhibitors	ICD-10 diagnosis	–	1.04–1.06	Tranylcypromine, phenelzine
Obesity				
Body mass index 50.0–59.9, kg/m^2	ICD-10 diagnosis	1.12–1.18	–	–
Morbid obesity	ICD-10 diagnosis	1.01–1.12	–	–
Polysomatic symptoms				
Fatigue	ICD-10 diagnosis	–	1.25–1.32	ICD-9: 780.7; ICD-10: R53%
Dizziness	ICD-10 diagnosis	–	1.22–1.28	ICD-9: 780.4, 438.85; ICD-10: R42%
Weakness	ICD-10 diagnosis	–	1.15–1.19	ICD-9: 728.87; ICD-10: M62.81

* Values are the 95% confidence interval of the odds ratio unless indicated otherwise. ICD-10 = International Classification of Diseases, Tenth Revision; ICD-9 = International Classification of Diseases, Ninth Revision; AC = acute; DVT = deep vein thrombosis.

Notably, frequent outpatient utilization (clinic follow-up) and increased rates of outpatient prescriptions for both mental and medical illnesses served as protective factors in both groups. Additionally, preventive medications and vaccinations, typical of longitudinal outpatient engagement, lowered the risks of SITBs in FM. Subsequent utilization analyses showed a dramatic difference in follow-up time (up to 40× increased time spent with providers in follow-up for the low-risk group compared to those with SITBs)

across outpatient settings, including primary care, medical specialty, and mental health clinics. There was a concomitant increase in the use of outpatient resources like health and behavioral interventions in the low-risk cohort. Outpatient health and behavior codes were more likely to occur for those with suicidal ideation without evidence of subsequent attempts, potentially indicating a preventive effect of health and behavior intervention in this high-risk cohort. These findings suggest that further research in patterns of outpatient engagement with respect to suicidality may be indicated.

This work extends existing research by quantifying, characterizing, and predicting SITB risk in a population, using clinical data science for the first time. Our study confirms and builds upon known risk factors of SITBs in FM based on both literature review and clinical expertise. Building on existing research, we also highlight actionable foci of risk management strategies (e.g., polysomatic symptoms, pharmacologic therapies) and the buffering effect of outpatient engagement to lower predicted risk.

Strengths of our study include using validated models applied to a valid phenotype of FM in a large EHR cohort. The models were designed to scale to any clinical setting with EHR data, facilitating external validation in this study. Applying these methods to a large academic medical center allowed us to sample patients at all points of care, assessing both known general and disease-specific risk factors concurrently. In addition to reviewing investigations to date, we combined expertise in machine learning, rheumatology, and psychology to identify additional patient characteristics to clinically inform risk prediction and interpret results.

These findings should be interpreted in light of study limitations. We relied on a single major academic medical center for study data. Our overall sample size was relatively small; however, this size is reasonable given the low base-rate phenomenon of SITBs in FM (in our cohort, approximately 0.4%). External validity results of this investigation are encouraging, but studies of reproducibility and generalizability in new settings are important steps of future work. In working with EHR data, there is always a risk of misclassification. Our reliance on the suicidal ideation codes is typical of this literature, but codes are an imperfect surrogate for true SITBs. Suicidal ideation remains at risk of underreporting. We report a 1.1% prevalence of documented suicidal ideation in this cohort. Under-documentation occurs from multiple potential sources, including patient hesitancy to report symptoms, lack of provider inquiry, and billing workflows failing to document diagnostic codes even if the latter 2 have occurred. Other studies in FM that were reliant on patient self-report have been associated with higher rates. Future analyses should address whether these differences in prevalence result from differences in self-report compared to retrospective EHR analyses, underreporting, incomplete documentation, or innate differences in our cohort compared to those in other health systems or countries. An existing limitation of replicable machine learning methods is the reliance on structured data within the health record to assess for patient characteristics that inform outcomes. While this reliance permits replicability/reproducibility and the potential for larger-scale investigations across networks, nuance can be lost in additional risk factors that may exist in unstructured data, such as the text of patient notes (versus a diagnostic code, for example). Experts are addressing this limitation by processing clinical text through natural language (28), which remains a future direction for this work.

While our current efforts focus on identifying risk, the ultimate goal is to translate these findings into actionable methods in clinical settings to enhance suicide prevention. A clear signal from this investigation is the importance of simply maintaining outpatient contacts over time to reduce the risk of SITBs. Predictive models like ours may play a role in identifying those patients who are both at risk of SITBs and who have been lost to follow-up. Enhancing outpatient continuity with at-risk patients is an active area of prevention in military and civilian settings and in diverse diseases (22,29,30). The gold standard for pain treatment is multimodal therapy, including psychologic approaches to pain management (31). Cognitive behavioral therapy in particular has been shown to improve outcomes in FM by improving mood, pain-related disability, and pain severity at follow-up (32). Given our findings that outpatient engagement of any type, including mental health engagement, may attenuate the risk of suicide attempt in those with suicidal ideation, we suggest a connection to mental health resources such as cognitive-behavioral therapy for FM patients with suicidal ideation, to enhance outpatient engagement and provider connection. Providers have expressed helplessness and frustration with being unable to intervene in complex situations for patients with FM (33). This work shows that the contact itself may have intrinsic benefits that decrease the likelihood of suicidality in this population.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. McKernan 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. McKernan, Crofford.
Acquisition of data. McKernan, Lenert, Walsh.
Analysis and interpretation of data. McKernan, Lenert, Crofford, Walsh.

REFERENCES

1. Singh GK. Area deprivation and widening inequalities in US mortality, 1969–1998. Am J Public Health 2003;93:1137–43.
2. National Institute of Mental Health. Suicide. URL: https://www.nimh.nih.gov/health/statistics/suicide.shtml.
3. Racine M. Chronic pain and suicide risk: a comprehensive review. Prog Neuropsychopharmacol Biol Psychiatry 2018;87:269–80.
4. Tang NK, Crane C. Suicidality in chronic pain: a review of the prevalence, risk factors and psychological links. Psychol Med 2006;36:575–86.
5. Ilgen MA, Kleinberg F, Ignacio RV, Bohnert AS, Valenstein M, McCarthy JF, et al. Noncancer pain conditions and risk of suicide. JAMA Psychiatry 2013;70:692–7.

6. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10.
7. Dreyer L, Kendall S, Danneskiold-Samsøe B, Bartels EM, Bliddal H. Mortality in a cohort of Danish patients with fibromyalgia: increased frequency of suicide. Arthritis Rheumatol 2010;62:3101–8.
8. Wolfe F, Hassett AL, Walitt B, Michaud K. Mortality in fibromyalgia: a study of 8,186 patients over thirty. Arthritis Care Res (Hoboken) 2011;63:94–101.
9. Calandre EP, Vilchez JS, Molina-Barea R, Tovar MI, Garcia-Leiva JM, Hidalgo J, et al. Suicide attempts and risk of suicide in patients with fibromyalgia: a survey in Spanish patients. Rheumatology (Oxford) 2011;50:1889–93.
10. Triñanes Y, González-Villar A, Gómez-Perretta C, Carrillo-de-la-Peña MT. Suicidality in chronic pain: predictors of suicidal ideation in fibromyalgia. Pain Pract 2015;15:323–32.
11. Tøllefsen IM, Helweg-Larsen K, Thiblin I, Hem E, Kastrup MC, Nyberg U, et al. Are suicide deaths under-reported? Nationwide re-evaluations of 1800 deaths in Scandinavia. BMJ Open 2015;5:e009120.
12. Gould MS, Munfakh JL, Kleinman M, Lake AM. National suicide prevention lifeline: enhancing mental health care for suicidal individuals and other people in crisis. Suicide Life Threat Behav 2012;42:22–35.
13. Kikuchi N, Ohmori-Matsuda K, Shimazu T, Sone T, Kakizaki M, Nakaya N, et al. Pain and risk of completed suicide in Japanese men: a population-based cohort study in Japan (Ohsaki Cohort Study). J Pain Symptom Manage 2009;37:316–24.
14. Smith MT, Edwards RR, Robinson RC, Dworkin RH. Suicidal ideation, plans, and attempts in chronic pain patients: factors associated with increased risk. Pain 2004;111:201–8.
15. Macfarlane GJ, McBeth J, Silman AJ. Widespread body pain and mortality: prospective population based study. BMJ 2001;323:662–5.
16. Jimenez-Rodríguez I, Garcia-Leiva JM, Jimenez-Rodriguez BM, Condés-Moreno E, Rico-Villademoros F, Calandre EP. Suicidal ideation and the risk of suicide in patients with fibromyalgia: a comparison with non-pain controls and patients suffering from low-back pain. Neuropsychiatr Dis Treat 2014;10:625–30.
17. Ratcliffe GE, Enns MW, Belik SL, Sareen J. Chronic pain conditions and suicidal ideation and suicide attempts: an epidemiologic perspective. Clin J Pain 2008;24:204–10.
18. Hassett AL, Aquino JK, Ilgen MA. The risk of suicide mortality in chronic pain patients. Curr Pain Headache Rep 2014;18:436.
19. Aaltonen K, Näätänen P, Heikkinen M, Koivisto M, Baryshnikov I, Karpov B, et al. Differences and similarities of risk factors for suicidal ideation and attempts among patients with depressive or bipolar disorders. J Affect Disord 2016;193:318–30.
20. Pompili M, Sher L, Serafini G, Forte A, Innamorati M, Dominici G, et al. Posttraumatic stress disorder and suicide risk among veterans: a literature review. J Nerv Ment Dis 2013;201:802–12.
21. Walsh CG, Ribeiro JD, Franklin JC. Predicting risk of suicide attempts over time through machine learning. Clin Psychol Sci 2017;5:457–69.
22. Walsh CG, Riberio JD, Franklin JC. Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning. J Child Psychol Psychiatry 2018;59:1261–70.
23. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther 2008;84:362–9.
24. Doss J, Mo H, Carroll RJ, Crofford LJ, Denny JC. Phenome-wide association study of rheumatoid arthritis subgroups identifies association between seronegative disease and fibromyalgia. Arthritis Rheumatol 2017;69:291–300.
25. Harrell FE Jr. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag; 2001. p. 53–85.
26. Walsh CG, Sharman K, Hripcsak G. Beyond discrimination: a comparison of calibration methods and clinical usefulness of predictive models of readmission risk. J Biomed Inform 2017;76:9–18.
27. Bach FR. Bolasso: model consistent lasso estimation through the bootstrap. In: Proceedings of the 25th International Conference on Machine Learning; 2008 July 5–9; Helsinki, Finland. p. 33–40.
28. McCoy TH Jr, Castro VM, Roberson AM, Snapper LA, Perlis RH. Improving prediction of suicide and accidental death after discharge from general hospitals with natural language processing. JAMA Psychiatry 2016;73:1064–71.
29. Kessler RC, Warner CH, Ivany C, Petukhova MV, Rose S, Bromet EJ, et al. Predicting suicides after psychiatric hospitalization in US Army soldiers: the Army Study To Assess Risk and rEsilience in Servicemembers (Army STARRS). JAMA Psychiatry 2015;72:49–57.
30. Barak-Corren Y, Castro VM, Javitt S, Hoffnagle AG, Dai Y, Perlis RH, et al. Predicting suicidal behavior from longitudinal electronic health records. Am J Psychiatry 2017;174:154–62.
31. Department of Health and Human Services. National pain strategy: a comprehensive population health strategy for pain. 2015. URL: https://iprcc.nih.gov/docs/HHSNational_Pain_Strategy.pdf.
32. Bernardy K, Klose P, Welsch P, Häuser W. Efficacy, acceptability and safety of cognitive behavioural therapies in fibromyalgia syndrome: a systematic review and meta-analysis of randomized controlled trials. Eur J Pain 2018;22:242–60.
33. Hayes SM, Myhal GC, Thornton JF, Camerlain M, Jamison C, Cytryn KN, et al. Fibromyalgia and the therapeutic relationship: where uncertainty meets attitude. Pain Res Manag 2010;15:385–91.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1264–1269
DOI 10.1002/acr.23739

BRIEF REPORT

How Do Health Literacy, Numeric Competencies, and Patient Activation Relate to Transition Readiness in Adolescents and Young Adults With Rheumatic Diseases?

Samuel M. Lazaroff,[1] Alexa Meara,[1] Mary Kate Tompkins,[1] Ellen Peters,[1] and Stacy P. Ardoin[2]

Objective. To evaluate how demographics, health literacy, numeracy, and patient activation are related to transition readiness in adolescent and young adult (AYA) patients and to describe how parent/guardian (PG) performance on these metrics predicts AYA patients' transition readiness.

Methods. In this single center, cross-sectional study, consecutive English-speaking AYA patients ages 17–21 years and PGs were recruited from outpatient rheumatology clinics. Participants completed the following self-reported instruments: demographic questionnaire, Short Test of Fundamental Health Literacy, Objective Numeracy Scale, Subjective Numeracy Scale, Symbolic-number mapping, Patient Activation Measure, and Transition Readiness Assessment Questionnaire (TRAQ; AYA patients only).

Results. Ninety-one AYA patients participated in the study, of whom 64 of 91 (70%) had juvenile idiopathic arthritis, and 54 PGs. Mean ± SD TRAQ score was 4.0 ± 0.65, correlating with "I am starting to do this" stage of change. Most participants (98%) had adequate health literacy. Multivariable regression analysis showed that AYA patients of female sex, older age, and higher patient activation significantly predicted higher TRAQ scores ($P < 0.05$). No PG characteristics were linked to higher AYA patient TRAQ scores.

Conclusions. Transition readiness in AYA patients as measured by TRAQ is associated with female sex, older age, and higher patient activation. Though sex and age are nonmodifiable, interventions to boost patient activation represent a promising opportunity to improve transition readiness and outcomes.

INTRODUCTION

Since the mid-1980s, leaders in the fields of pediatrics and adolescent, internal, and family medicine have emphasized the need to study and improve the chronic condition transition process, defined as the "purposeful, planned movement of adolescents and young adults with chronic physical and medical conditions from child-centered to adult-oriented health-care systems" (1). Three decades later, our knowledge on the subject remains incomplete, with insufficient quantitative data on the factors that contribute to successful transition in order to inform best practices and improve outcomes (2). Knowledge gaps in transition from pediatric to adult care include understanding the relationships among patient and parent health literacy, numeracy, patient activation, and transition readiness. These skills are vital in provider-patient communication and health care system navigation.

Health literacy is the "capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions" (3). Poor health literacy is linked to medication non-adherence, poor appointment keeping, poorer health knowledge, and poorer self-management of medical conditions in several disease states (3,4). Likewise, lower numeracy, "the ability to use and understand numbers in daily life," in the medical setting is associated with higher rates of comorbidities, unrealistic expectations for treatment options, and poorer disease control (5,6). Both health literacy and numeracy are key to comprehending health care information, but a study by Hibbard and colleagues demonstrated that comprehension alone is insufficient

Supported by the Ohio State University College of Medicine Barnes Research Scholarship awarded to Mr. Lazaroff.

[1]Samuel M. Lazaroff, BA, Alexa Meara, MD, Mary Kate Tompkins, MS, Ellen Peters, PhD: Ohio State University, Columbus; [2]Stacy P. Ardoin, MD, MS: Ohio State University and Nationwide Children's Hospital, Columbus.

Dr. Peters owns stock or stock options in Insignia Health, which owns the Patient Activation Measure used in this research. Dr. Ardoin has received consulting fees from the American Board of Pediatrics, Novartis, and Aurinia Pharmaceuticals (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Stacy P. Ardoin, MD, MS, 700 Children's Drive, Columbus, Ohio 43205. E-mail: stacy.ardoin@osumc.edu.

Submitted for publication January 4, 2018; accepted in revised form August 21, 2018.

SIGNIFICANCE & INNOVATIONS

- Female sex, older age, and higher patient activation scores predicted transition readiness in adolescent/young adult patients, as measured by the Transition Readiness Assessment Questionnaire.
- Parent competencies and adolescent/young adult health literacy and numeracy scores did not predict transition readiness.
- Interventions shown to increase patient activation may improve transition readiness.

for patients to make informed choices (7). Patient activation (taking an active role in managing one's own health and health care) is an additional measure of patients' abilities to weigh choices and make health care decisions. Increases in patient activation are associated with increased engagement in healthier lifestyle behaviors and fewer hospital visits (7).

However, the transition process involves not only adolescent and young adult (AYA) patients but also their parents/guardians (PGs), with control over health care decisions gradually transferred from PGs to the AYA patient. Health literacy, numeracy, patient activation, and transition readiness have been studied in adolescent patients, but relationships between AYA patients and PGs in the setting of transition remain understudied (3,4,7,8). A study by Chisolm et al demonstrated the prevalence of health literate congruent parent-child dyads (i.e., both parent and patient have adequate health literacy) and incongruent dyads within a low socioeconomic (SES) population, suggesting the need for providers to consider both the abilities of the patient as well as their PG in health care navigation. Notably, 23% of teens with sub-adequate health literacy also had a PG with poor health literacy (9), and low parental numeracy has been linked to over- and under-weight children (10).

Focusing on pediatric rheumatology, a study by Bingham et al (2) showed that older patient age, younger PG age, having a family member with a similar disease, longer disease duration, having other comorbidities, and having had a summer job correlated with increased self-reported autonomy in accessing medical care. These findings highlight important personal characteristics that may contribute to successful transition in AYA patients. The goal of this study was to evaluate the association of AYA and PG health literacy, numeracy, and patient activation with AYA transition readiness in a cohort of AYA patients with chronic rheumatic conditions.

MATERIALS AND METHODS

In this single center, cross-sectional study, consecutive AYA patients and 1 PG of each AYA patient were recruited from the outpatient rheumatology clinic from May to August of 2016; AYA patient inclusion criteria included those who were ages 17–21, English-proficient, and capable of completing the questionnaires. Participants ages ≥18 years provided verbal consent to participate; those younger than 18 provided verbal assent. The Nationwide Children's Hospital Institutional Review Board approved the protocol in regard to ethics.

Data collection. AYA patients completed demographic questionnaires and the instruments on paper (described below). When possible, PGs completed all measures except the Transition Readiness Assessment Questionnaire (TRAQ). PGs were instructed to assess themselves on the measures, not the AYA patient. Data collected on paper were entered into a secure database for analysis. No personal health information was collected or recorded.

Measures. *Demographics.* All participants provided demographic information including sex, race, ethnicity, marital status, insurance coverage, employment status, household income, and education. Patients also reported their rheumatologic diagnosis and month/year of diagnosis.

Health literacy. Health literacy was assessed using the Short Test of Fundamental Health Literacy (sTOFHLA), a 36-item assessment designed to measure reading comprehension (11). Continuous sTOFHLA scores were used in analyses. Scores in the range 0–16 are defined as inadequate health literacy, 17–22 as marginal, and 23–36 as adequate. This test contains 2 English passages written at 4th grade and 10th grade reading levels and has been widely validated across disease states (11).

Transition Readiness Assessment Questionnaire (TRAQ). The TRAQ (version 5.0), a 20-item survey that measures proficiency in 5 domains (including managing medications, appointment keeping, tracking health issues, talking with providers, and managing daily activities), was used to measure the readiness of AYA patients to transition to adult providers. Each item is scored on a scale of 1 ("No, I do not know how") to 5 ("Yes, I always do this when I need"), based upon the stages of change model. Item scores were averaged to produce an overall score (12). The TRAQ can be used for any chronic medical condition and has been validated or studied in populations including healthy individuals and patients with cystic fibrosis, congenital heart disease, sickle cell disease, and those with rheumatologic, gastroenterologic, and endocrinous disorders (12,13–15). The TRAQ has been utilized as a key measure in interventional and transition observational studies (13–15). To date, the TRAQ is the most extensively validated disease-neutral transition readiness assessment tool.

Objective Numeracy Scale (ONS). The ONS includes 8 math questions that are focused primarily on percentages and proportions. The number of correct questions is the total score, with questions left blank scored as incorrect. Possible scores range from 0 to 8; higher scores indicate higher levels of numeric

ability. This tool was developed to have a broader range of difficulty relative to other tests available to researchers (16).

Subjective Numeracy Scale (SNS). The SNS is an 8-question survey assessing self-perceptions of math abilities that relate positively to objective numeracy scores (17). Patients rate themselves from 1 to 6 on each of 8 items, and an average is computed. Patients with high overall scores believe they have high math capabilities and generally prefer to use numbers instead of words. Higher SNS scores have been related to greater confidence and perseverance in decision tasks (17).

Symbolic-number Mapping (SMap). The SMap is a 22-question assessment associated with improved ability to discriminate and remember numbers (18). Participants are told to "please draw a 'hatch mark' or little vertical up-and-down line to indicate how big is the number shown" on a scale from 0 to 1,000. On each page, participants place a number on the number line as instructed. As per previously published methods, this measure was reverse scored (i.e., scores closer to 0 represent more exact mapping) on a log scale based on the mean absolute differences between participants' answers and the target number (18). More exact symbolic-numbering mapping is thought to be related to the development of more proficient objective math skills (17).

Patient Activation Measure (PAM). The 13-question PAM has been used to predict outcomes, including healthy behaviors, chronic self-management, maintaining a health/blood pressure diary, controlling chronic illness, and health care costs (8). The PAM has been studied and validated in adults and adolescents; questions assess preventive and health-oriented behaviors, self-management, and health information seeking. Each question is scored on a 4-point scale (1 = disagree strongly, 2 = disagree, 3 = agree, 4 = agree strongly), which are then translated into a 0–100 score. Higher scores represent increased activation.

Data analysis. Descriptive statistics were performed. Univariate analysis was performed to identify relations between TRAQ scores and AYA and PG demographic characteristics and other survey scores. Variables with P values < 0.05 in univariate analysis were included in multivariate modeling. Using TRAQ scores as the dependent variable, multivariate linear regression models were conducted using backwards, stepwise logistic regression. P values less than 0.05 were considered significant. The multivariable analyses did not control for education because education exerts a causal influence on numeric ability.

RESULTS

Ninety-one AYA patients and 54 PGs completed the study. Demographic data and performance on survey instruments are summarized in Table 1. The mean ± SD age of AYA respondents was 19 ± 1.3 years. Most patients in this cohort were female (80%), white (78.6%), and had completed some college/technical school (66.7%), with an additional 26.1% in 12th grade. All annual household income levels were represented, with the plurality (21.5%) in the $100,000–150,000 range. Among PGs, the mean ± SD age was 48 ± 7.9 years and 87% were female. The majority of PGs were white (88.5%) and had completed college or technical school (>60%). All levels of SES were represented in PGs, with $100,000–$150,000 being the most common response (26%).

The majority of AYA patients in this cohort (64 of 91) saw a rheumatologist for juvenile idiopathic arthritis (JIA). Four

Table 1. Baseline characteristics and survey results for transitioning adolescent/young adults and their parents/guardians*

	AYA (n = 91)	PG (n = 54)
Age, mean ± SD years	19 ± 1.3	48 ± 7.9
Women	72 (80)	45 (87)
Race		
White	70 (78.6)	46 (88.5)
African American	11 (12.4)	3 (5.8)
Other	8 (9.0)	3 (5.8)
Education		
10th grade	1 (1.5)	0
11th grade	4 (5.8)	0
12th grade	18 (26.1)	9 (16.7)
Some college/tech school	46 (66.7)	12 (22.2)
Graduated college/tech school	0	20 (37.0)
Graduate degree	0	13 (24.1)
Annual household income		
<$25,000	12 (18.5)	2 (4.0)
$25,000–$49,999	11 (16.9)	9 (18.0)
$50,000–$74,999	9 (13.6)	10 (20.0)
$75,000–$99,999	7 (10.8)	8 (16.0)
$100,000–$150,000	14 (21.5)	13 (26.0)
>$150,000	12 (18.5)	8 (16.0)
TRAQ score, mean ± SD†	4.0 ± 0.67	–
sTOFHLA score, mean ± SD‡	34.1 ± 3.5	34.7 ± 1.7
PAM, mean ± SD§	64.8 ± 17.6	68.0 ± 12.8
SNS, mean ± SD¶	3.8 ± 1.1	4.0 ± 1.0
ONS, mean ± SD#	3.7 ± 1.7	3.8 ± 1.7
SMAP, mean ± SD**	−0.87 ± 0.2	−0.85 ± 0.2

* Values are the number (%) of adolescent/young adults (AYAs) and their parents/guardians (PGs) unless indicated otherwise. TRAQ = Transition Readiness Assessment Questionnaire; sTOFHLA = Short Test of Fundamental Health Literacy in Adults; PAM = Patient Activation Measure; SNS = Subjective Numeracy Scale; ONS = Objective Numeracy Scale; SMAP = Symbolic Number Mapping.
† Possible range 1–5.
‡ Possible range 0–36.
§ Possible range 0–100.
¶ Possible range 1–6.
# Possible range 0–8.
** Possible range –2.14 to 0.

Table 2. Final model of predictors of AYA TRAQ scores from AYA surveys*

Variable†	B coefficient‡	b coefficient§	SE¶	*t* value#	*P***
Intercept	0.00	−0.47	0.86	−0.55	0.58
AYA PAM	0.40	0.02	0.003	4.60	<0.0001
AYA age	0.33	0.17	0.05	3.72	0.0004
AYA female	0.22	0.37	0.15	2.49	0.01

* AYA = adolescent/young adult; SE = standard error; PAM = Patient Activation Measure.
† R^2 = 0.38.
‡ Coeffient variable = 13.6.
§ Root mean square error = 0.55.
¶ Transition Readiness Assessment Questionnaire (TRAQ) score (mean) = 4.03
# *F* value = 16.93.
** $P < 0.0001$.

patients had lupus, 4 had mixed connective tissue disease, 2 had juvenile dermatomyositis, and the remainder had other conditions, including psoriasis, Behçet's disease, and scleroderma.

The mean ± AYA patient TRAQ score was 4.0 ± 0.67, which reflects an "I am starting to do this" stage of change (12). Most AYA patients (98%) and all PGs had adequate health literacy, as defined by sTOFHLA scores ≥22. Broad concordance existed between AYA patients and PGs on the numeracy measures; average scores of 3.8 and 4.0 (on the 1–6 scale), respectively, were recorded on the SNS. The average score of AYA patients on the ONS was 3.7 and was 3.8 for PGs (out of 8 possible). On the SMap, the average score of AYA patients was –0.87 and –0.85 for PGs. Similarly, the average AYA patient score on the PAM was 64.8, (on a 0–100 scale) signifying "Agree" with items such as, "When all is said and done, I am the person who is responsible for taking care of my health" and "I understand my health problems and what causes them," which is a level 3 score on the PAM (55.2–72.4), indicative of beginning to engage in recommended health behaviors (7,15). The corresponding average score of PGs was 68.0.

In AYA patients, female sex ($P = 0.01$), older age ($P = 0.0004$), and higher PAM scores ($P < 0.0001$) were related to higher transition readiness (TRAQ score) in multivariate regression (Table 2). Results in Table 2 indicated that, for every year increase in age, TRAQ scores increased by 0.17 units. Table 2 standardized regression results indicated that patient activation was the strongest predictor of AYA patient TRAQ scores, relative to age and sex. Specifically, 1 SD increase in PAM corresponded to a 0.40 SD increase in AYA patient TRAQ scores. Other variables, including scores on the sTOFHLA and numeracy measures as well as demographic data, were not predictive of TRAQ scores in AYA patients. In separate modeling, PG health literacy, numeric competency, and patient activation were not independently associated with AYA TRAQ scores (see Table 3). Higher PG sTOFHLA scores were marginally associated with higher AYA TRAQ scores ($P = 0.06$). Specifically, 1 SD increase in PG health literacy corresponded with a 0.26 SD increase in AYA patient TRAQ scores.

DISCUSSION

To our knowledge, this study is the first to attempt to quantitatively examine transition readiness in the context of AYA and PG health literacy, numeracy, and patient activation. We have shown that AYA patient activation, older age, and female sex predicted higher performance on the TRAQ, whereas other AYA patient characteristics, PG demographics, and other responses did not relate to AYA patient transition readiness in this cohort. The present study also represents the first known study to use the PAM in a pediatric population. Although the measured PG competencies were not significantly associated with AYA transition readiness, certainly

Table 3. Final model of predictors of AYA TRAQ scores from PG surveys*

Variable†	B coefficient‡	b coefficient§	SE¶	*t*#	*P***
Intercept	0.00	−0.05	2.06	−0.02	0.98
PG sTOFHLA	0.26	0.11	0.06	1.90	0.06

* AYA = adolescent/young adult; PG = parent/guardian; SE = standard error; sTOFHLA = Short Test of Functional Health Literacy in Adults.
† R^2 = 0.07.
‡ Coeffient variable = 18.68.
§ Root mean square = 0.72.
¶ Transition Readiness Assessment Questionnaire (TRAQ) score (mean) = 3.86.
# *F* = 3.60.
** *P* = 0.06.

PGs importantly influence AYA beliefs, values, behaviors, and illness experience in many ways that these instruments do not measure.

The association between the PAM and the TRAQ is not surprising because these 2 validated instruments include similar domains. For example, 1 item on the PAM is "I am confident that I can tell a doctor my concerns, even when he or she does not ask"(7). Similarly, TRAQ asks "Do you tell the doctor or nurse what you are feeling?"(12). The strong concordance between these 2 measures is thus reassuring and provides validity to the TRAQ instrument in light of the extensive clinical validation of the PAM (7,8). Higher PAM scores have been linked to improved health behaviors, including seeking preventative care (e.g., cancer screenings, immunizations), healthy eating, exercising regularly, and avoiding smoking. In chronic disease states, highly activated patients have been shown to have better chronic disease control and long-term outcomes, as well as decreased utilization of inpatient and emergency department services and lower health care costs (8).

Importantly, various interventions are effective in increasing adult patient activation. In a study by Hibbard and Greene, these interventions are broadly classified into the following 3 groups: skill development, problem solving, and peer support; changing the social environment; and tailoring support to the person's activation level (8). Skill-based interventions may be community based and rely on teaching patients about specific aspects of disease management, as well as how to communicate effectively with their providers. Often such a program is combined with tailored care, wherein providers use the patient's PAM score to dictate priorities in care and set realistic goals (i.e., smaller steps and more frequent follow-up for less-activated patients). Along these lines, some health care systems have turned to using the PAM as a "vital sign" to be checked regularly to aid clinical decision-making (8). Similar interventions targeted toward transitioning AYA patients can likely improve transition readiness. Using patient activation as a vital sign could help providers to identify patients at increased risk of poor transition outcomes. Further study and use of the PAM in the setting of transition is recommended.

With respect to the associations between older age and female sex on TRAQ scores in AYA patients, these trends support the published developmental literature for adolescents and existing transition literature. Older age and being female are associated with increased self-reported autonomy in pediatric patients, likely allowing those patients to engage more fully in their own medical care (2).

As hypothesized, health literacy and numeracy were not linked to improved TRAQ performance. Nearly every participant in our cohort had adequate health literacy as defined by sTOFHLA ≥23 (11). Our population was largely female, white, and of higher SES and the majority of AYA patients had JIA. Additional study is warranted in other populations in which differences in health literacy are greater and/or could play a more important role in differentiating those with high and low transition readiness. Indeed, additional data on the roles of health literacy and numeracy could provide actionable guidance for providers on the most effective forms of communication to present health-related information to AYA patients during the transition process, though population-specific effects may exist as noted above.

Our study has several limitations. This small cohort that was enrolled at a single center was racially homogenous and skewed toward white AYA patients (and PGs) who lived in high-income households and were well-educated. Data on education and household income were missing for a substantial number of patients. The cohort's composition may reflect selection bias. Some potential enrollees (both AYA patients and PGs) refused participation when informed that the surveys included completing math problems and a reading test. Thus, numeracy and health literacy scores may be inflated; AYA patients who struggle in math or reading may have been more likely to refuse participation, possibly corresponding to AYA patients of lower education levels and/or SES. It is not possible to know whether patient refusal affected the sex or racial makeup of the cohort. Our limited ability to enroll PG participants (whether because of their refusal or because the AYA patient was unaccompanied to their medical appointment) restricted statistical power to examine correlations between PG and AYA patient characteristics. Finally, our small study used the TRAQ as a measure of transition readiness. Though clinically validated and widely used as a marker of transition readiness, the TRAQ serves as an imperfect surrogate marker compared to longitudinal data on the actual transition outcomes of these patients. It is our hope that larger, multicenter studies will follow up to examine the validity of our findings (and expand upon them) in more diverse populations, particularly concerning the role of PGs in the transition process.

In conclusion, this study has shown that AYA patient activation, older age, and female sex predict higher TRAQ performance in AYA patients with chronic rheumatologic diseases. In this cohort, no demonstrated association existed between PG demographics or survey performance with AYA patient TRAQ score. Our results identify patient activation as an important modifiable factor in the transition process. Improving patient activation in AYA through targeted interventions represents an opportunity to greatly decrease the morbidity and mortality associated with prolonged transition.

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. Lazaroff 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. Lazaroff, Peters, Ardoin.

Acquisition of data. Lazaroff, Ardoin.

Analysis and interpretation of data. Lazaroff, Meara, Tompkins, Peters, Ardoin.

REFERENCES

1. Blum RW, Garell D, Hodgman CH, Jorissen TW, Okinow NA, Orr DP, et al. Transition from child-centered to adult health-care systems for adolescents with chronic conditions: a position paper of the Society for Adolescent Medicine. J Adolescent Health 1993;14:570–6.
2. Bingham CA, Scalzi L, Groh B, Boehmer S, Banks S. An assessment of variables affecting transition readiness in pediatric rheumatology patients. Pediatr Rheumatol Online J 2015;13:42.
3. Wong PK, Christie L, Johnston J, Bowling A, Freeman D, Joshua F, et al. How well do patients understand written instructions? Health literacy assessment in rural and urban rheumatology outpatients. Medicine (Baltimore) 2014;93:e129.
4. Gazmararian JA, Williams MV, Peel J, Baker DW. Health literacy and knowledge of chronic disease. Patient Educ Couns 2003;51:267–75.
5. Garcia-Retamero R, Andrade A, Sharit J, Ruiz JG. Is patients' numeracy related to physical and mental health? Med Decis Making 2015;35:501–11.
6. Cavanaugh K, Huizinga MM, Wallston KA, Gebretsadik T, Shintani A, Davis D, et al. Association of numeracy and diabetes control. Ann Intern Med 2008;148:737–46.
7. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. Health Aff (Millwood) 2015;34:431–7.
8. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. Health Aff (Millwood) 2013;32:207–14.
9. Chisolm DJ, Sarkar M, Kelleher KJ, Sanders LM. Predictors of health literacy and numeracy concordance among adolescents with special health care needs and their parents. J Health Commun 2015;20 Suppl 2:43–9.
10. Dallacker M, Hertwig R, Peters E, Mata J. Lower parental numeracy is associated with children being under- and overweight. Soc Sci Med 2016;161:126–33.
11. Baker DW, Williams MV, Parker RM, Gazmararian JM, Nurss J. Development of a brief test to measure functional health litearacy. Patient Educ Couns 1999;38:33–42.
12. Wood DL, Sawicki GS, Miller MD, Smotherman C, Lukens-Bull K, Livingood WC, et al. The Transition Readiness Assessment Questiontionnaire (TRAQ): its factor structure, reliability, and validity. Acad Pediatr 2014;14:415–22.
13. Mackie AS, Rempel GR, Kovacs AH, Kaufman M, Rankin KN, Jelen A, et al. A cluster randomized trial of a transition intervention for adolescents with congenital heart disease: rationale and design of the Chapter 2 study. BMC Cardiovasc Disord 2016;6:127.
14. Jensen PT, Paul GV, LaCount S, Peng J, Spencer CH, Higgins GC, et al. Assessment of transition readiness in adolescents and young adults with chronic health conditions. Pediatr Rheumatol Online J 2017;15:70.
15. Hart LC, Pollock M, Hill S, Maslow G. Association of transition readiness to intentional self-regulation and hopeful future expectations in youth with illness. Acad Pediatr 2017;17:450–5.
16. Weller JA, Dieckmann NF, Tusler M, Mertz CK, Burns WJ, Peters E. Development and testing of an abbreviated numeracy scale: a Rasch analysis approach. J Behav Decis Mak 2013;26:198–212.
17. Peters E, Bjalkebring P. Multiple numeric competencies: when a number is not just a number. J Pers Soc Psychol 2015;108:802–22.
18. Thompson CA, Siegler RS. Linear numerical-magnitude representations aid children's memory for number. Psychol Sci 2010;21:1274–81.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1270–1275
DOI 10.1002/acr.23757

Incidence, Clinical Manifestations, and Severity of Juvenile Idiopathic Arthritis Among Maori and Pacific Island Children

Anthony Concannon,[1] Peter Reed,[2] and Genevieve Ostring[3]

Objective. To describe the incidence, demographics, diagnostic clinical manifestations, and severity of juvenile idiopathic arthritis (JIA) in Maori and Pacific Island children compared to European children.

Methods. A chart review was conducted of all children with JIA seen by Auckland pediatric and rheumatology services between the years 2000 and 2015. Demographic data and diagnostic clinical manifestations, including poor prognostic features, were collated. The incidence, diagnostic, clinical manifestations, and severity of JIA were determined and compared between ethnic groups, in particular Maori, Pacific Island, and European children.

Results. The overall incidence in a New Zealand cohort of children with JIA was 5.1/100,000 children per year, which was significantly higher among European children (7.2/100,000 children per year) compared to all other ethnic groups. Poor prognostic features at diagnosis were present in 36% of children with JIA, with significantly more Maori and Pacific Island children presenting with poor prognostic features compared to European children (58% versus 27%; $P = 0.0001$). Maori and Pacific Island children had significantly more poor prognostic features per child associated with JIA (1.10 versus 0.37; $P < 0.0001$) and in oligoarticular and polyarticular JIA (1.28 versus 0.40; $P < 0.0001$), which was independent of socioeconomic status. Significant features included cervical involvement (25% versus 9%; $P = 0.03$), erosive changes (22% versus 8%; $P = 0.05$), joint space narrowing (13% versus 2%; $P = 0.02$), and positive rheumatoid factor polyarticular disease (47% versus 14%; $P = 0.01$).

Conclusion. Maori and Pacific Island children were more likely to present with poor prognostic features at diagnosis, although the incidence of JIA was demonstrated to be significantly higher among European children compared to all ethnic groups.

INTRODUCTION

Demographic and clinical manifestations of juvenile idiopathic arthritis (JIA) are well established in European and North American populations (1), where the reported incidence ranges between 1.6 and 23 per 100,000 children. The incidence and prevalence of JIA and other rheumatologic diseases such as juvenile systemic lupus erythematous (SLE) have been shown to be higher in indigenous populations (2). An increased understanding of the genomics of rheumatologic conditions would suggest that at least some of this dissimilarity is due to ethnic and genetic variation (3).

The Maori are the indigenous population of New Zealand, with nearly one quarter (n = 52,000 [24%]) of all Maori children and nearly two-thirds (n = 63,000 [65%]) of Pacific Island children in New Zealand residing within the Auckland region. Juvenile SLE has been shown to convey a worse prognosis among Maori and Pacific Island children when compared to European children (4). Although a previous study determined the incidence of juvenile rheumatoid arthritis among Auckland children (1971–1980) to be 3.1/100,000 per year (n = 55), there was no recognizable influence of race on the course of the disease (5). Since then, there have been changes in the nomenclature, including the categorization of subtypes and delineation of poor prognostic features (6). The aim of this study was to establish the current ethnic incidence and determine the prognostic features at diagnosis of JIA among Maori and Pacific Island children and compare these to their European counterparts.

[1]Anthony Concannon, MBChB, FRACP: Starship Hospital and Kidz First Hospital, Auckland, New Zealand; [2]Peter Reed, DPhil: Children's Research Centre, Starship Hospital, Auckland, New Zealand; [3]Genevieve Ostring, MBChB, FRACP: Starship Hospital and Waitakere Hospital, Auckland, New Zealand.

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

Address correspondence to Anthony Concannon, MBChB, FRACP, Department of Pediatric Rheumatology, Starship Hospital, Level 10 Support Building, Auckland Hospital, 2 Park Road, Grafton, Auckland 1023, New Zealand. E-mail: AConcannon@adhb.govt.nz.

Submitted for publication March 24, 2018; accepted in revised form September 11, 2018.

SIGNIFICANCE & INNOVATIONS

- The incidence of juvenile idioipathic arthritis (JIA) in New Zealand is 5.1/100,000 children per year.
- Poor prognostic features at JIA diagnosis are significantly more common among Maori and Pacific Island children, independent of socioeconomic status.
- JIA is significantly more common in New Zealand European children (7.2/100,000) compared to all other ethnicities, including Maori or Pacific Island children (2.7/100,000).

PATIENTS AND METHODS

The public health care system in New Zealand is subsidized by the government, with all children able to access heavily subsidized or free primary care through their general practitioner and free access to emergency care. The estimated pediatric population (ages 0–15 years) was 121,000 in Counties Manukau (serviced by Kidz First Hospital), 82,000 in Auckland Central (serviced by Starship Hospital), and 114,000 in Waitemata (serviced by Waitakere Hospital) (317,000 in total). Of these populations, an estimated 51% (62,000) of children in Counties Manukau, 29% (24,000) in Auckland Central, and 25% (29,000) in Waitemata are of Maori or Pacific Island ethnicity (36% in total) (7). In Auckland, children with JIA were almost exclusively treated by the Starship Hospital rheumatology service until the establishment of a national service in 2006, and integrated pediatric rheumatologist–led services at Waitakere Hospital in 2010 and Kidz First Hospital rheumatology services in 2012. During the period of the study, a single pediatric rheumatologist provided a private service within the Auckland region and had a policy of referring all children with JIA to the Starship Hospital pediatric rheumatology service for review.

A chart review was conducted of all children <16 years of age who had been diagnosed with JIA within Counties Manukau and Auckland and Waitemata counties by the associated district health boards between January 2000 and December 2015. Ethical approval for this study was obtained from all hospitals involved.

Within New Zealand, all individuals are assigned a national unique identifier to public health care services at the time of first presentation of disease, which is linked to all subsequent health care encounters and obtained clinical information. Patient encounters at all 3 district health boards are assigned a clinical code at the time of discharge linked to the national unique identifier, which is stored electronically in each hospital database. Children were identified with the Counties Manukau and Auckland and Waitemata county databases using the key words "juvenile idiopathic arthritis," "juvenile chronic arthritis," "juvenile rheumatoid arthritis," "juvenile arthritis," and "arthritis." This method captured all children seen in the inpatient, emergency, and day-stay units. In addition, all outpatient encounters at Starship Hospital and local Kidz First Hospital and Waitemata Hospital outpatient databases were manually reviewed, reducing the potential for selection bias by severity.

Children were included if they met the International League Against Rheumatism diagnostic criteria and were classified as having oligoarticular, polyarticular (rheumatoid factor [RF] positive or negative), systemic, psoriatic, or enthesitis-related arthritis (ERA) (8). There were no children with undifferentiated JIA. Arthritis was defined as joint swelling or effusion, or ≥2 joints of limited range of motion, tenderness or pain on motion, or warmth. Demographic data (age, sex, and ethnicity) and the number and type of joints involved were determined within 1 month of diagnosis. Ethnicity was recorded at the time of the initial patient encounter within the public system at all 3 district health boards. Ethnicity was derived and assigned based on the ethnic background that each child's family declared at the time of initial presentation to public health care services, with combination ethnicities allocated to a single ethnicity based on the Ministry of Health ethnicity "prioritization" standards (9). The Maori have an official status as the indigenous people of New Zealand but also share a Polynesian ancestry. Consequently, the results for Maori are presented both separately and in combination with other Pacific Island children.

RF and the presence of anti–cyclic citrullinated peptide (anti-CCP) antibodies at any time were recorded. Baseline imaging (radiograph and/or magnetic resonance imaging) that was obtained within 6 months of diagnosis was reviewed, including the presence of joint space narrowing, synovitis, and erosions. This information was used to determine poor prognostic features at diagnosis for children with oligoarticular or polyarticular JIA, sacroiliitis, and systemic arthritis, as outlined by the American College of Rheumatology (6). This data included arthritis of the hip or cervical spine, arthritis of the ankle or wrist, marked or prolonged inflammatory marker elevation, radiographic damage (erosions or joint space narrowing), positive RF or anti-CCP, 6-month duration of significant active systemic disease (fever, elevated inflammatory marker levels), or treatment of systemic JIA with systemic corticosteroids.

Incidence denominators were estimated from the New Zealand census results of 2001, 2006, and 2013, and by simple linear interpolation for the non-census years. Count data were collated for 5-year age intervals for ages 0–14 years, with age 15 estimated as 20% of the age 15–19 years group. Ethnicity data were derived from "detailed single and combination" data (7) and "Ethnic group (single and combination) by age group and sex, for the census usually resident population count, 2001, 2006, and 2013" data, with combination ethnicities allocated to a single ethnicity in the same manner as patient ethnicity. Middle Eastern, Latin American, and African (MELAA) persons were not separated from "other" in the 2000 census, and a large number of Europeans were classified as "other" in 2006 ("New Zealander" was a popular response in 2006) (10). Therefore, to maintain consistency over the years, European ethnicity was combined

with "other" and MELAA for the incidence calculations. In 2013, "other" and MELAA groupings comprised only 1% and 2% of the pediatric population, respectively. Nevertheless, persons in the European/MELAA/"other" category are predominantly European.

Patient socioeconomic status (SES) was measured using the New Zealand Deprivation (NZDep) Indices of 2001, 2006, and 2013 (11). NZDep is an area-based index of deprivation based on national census variables (e.g., income, house ownership, and qualifications) and calculated for geographic units (mesh blocks) containing ~100 people (12). The index is divided nationally into deciles, where decile 10 represents the greatest deprivation (the lowest SES). Patient home addresses at the time of diagnosis were allocated an NZDep decile from the time-appropriate index, using the StatsNZ Geographic Data Service web-based geo mapping tools for 2001, 2006, and 2013 (13–15). For this study, the NZDep deciles were combined into 3 groups of high (1–3), mid (4–7), and low (8–10) SES categories and analyzed as a categorical variable. Data analyses were undertaken using JMP, version 13 (SAS) and StatsDirect, version 3 software. Incidence 95% confidence intervals (95% CI) and comparisons were undertaken using Poisson distribution–based methods. For statistical analyses, we combined Maori and Pacific Island patients, and compared them to European patients. Age at diagnosis and number of poor prognostic features were compared using the *t*-test. The proportions of patient subgroups were compared using Fisher's exact test and calculation of odds ratios (ORs) with 95% CIs. SES was incorporated as a covariable with patient ethnicity in order to investigate the effect of ethnicity (independent of SES) on disease subtype and poor prognosis factors, using multivariable logistic or least squares–regression as appropriate.

RESULTS

A total of 248 children were diagnosed with JIA with an overall incidence of 5.1/100,000 children per year. There was a significantly higher annual incidence of JIA among European/MELAA (7.2/100,000) children compared with Maori (3.3/100,000), Pacific Island (2.1/100,000), Maori or Pacific Island (2.7/100,000), and Asian (4.4/100,000) children (Table 1). The Pacific Island children were Tongan (n = 9), Samoan (n = 7), Niuean (n = 3), and unspecified Pacific (n = 1). The European/MELAA group were 97% European (n = 157), with the remainder Middle Eastern (n = 1) and African (n = 4). The Asian children within the study were predominantly Indian (n = 22) and Chinese (n = 8), and the remainder were Japanese (n = 1), Korean (n = 1), Filipino (n = 1), and unspecified Asian (n = 5). The average age at diagnosis was 8.9 years, with a female predominance (61%) (Table 2). SES is highly associated with ethnicity in New Zealand (9), and this was reflected in the findings of the present study, with European children being significantly less likely to be of low SES (10%) compared to Maori or Pacific Island children (69%) (Table 2). The frequency of oligoarticular JIA among children of Maori or Pacific Island ethnicity was lower than European children (31% versus 53%; OR 0.4 [95% CI 0.2–0.8]), although the difference was not statistically significant after adjusting for SES (OR 0.5 [95% CI 0.2–1.2]).

The frequency of polyarticular disease among children with JIA who are of Maori or Pacific Island ethnicity was higher than among European children, although not significantly (35% versus 24%; OR 1.8 [95% CI 0.9–3.6]) (Table 2), and the difference was similar after adjusting for SES (OR 1.5 [95% CI 0.6–3.6]). The frequency of RF positive polyarticular disease among children with JIA who are of Maori or Pacific Island ethnicity was also higher than among European children (17% versus 3%; OR 6.1 [95% CI 1.9–20]) (Table 2), and the difference remained significant after adjusting for SES (OR 5.1 [95% CI 1.2–22]). ERA, systemic, and psoriatic JIA and were the least common subtypes. ERA was present in 15% of children with JIA, with no significant ethnic difference (Table 2) or SES effect. The frequency of systemic JIA among children with JIA who are of Maori or Pacific Island ethnicity was higher than that among European children (17% versus 6%; OR 2.9 [95% CI 1.1–7.9]) (Table 2), although the difference was substantially weaker after adjusting for SES (OR 1.6 [95% CI 0.4–5.7]). Psoriatic arthritis was present in 3% of children with JIA in our cohort, with no significant ethnic difference (Table 2) or SES effect.

Table 1. Ethnic incidence of JIA in Auckland, New Zealand, 2000–2015

Ethnicity	No. (%)	Cases/ year	Incidence (per 100,000 persons/year)	95% CI	Rate ratio	*P*
European*	162 (65)	10.1	7.2	6.2–8.4	Ref.	–
Maori	28 (11)	1.8	3.3	2.2–4.8	0.45	<0.0001
Pacific Island	20 (8)	1.3	2.1	1.3–3.3	0.29	<0.0001
Maori or Pacific Island	48 (19)	3.0	2.7	2.0–3.6	0.37	<0.0001
Asian	38 (15)	2.4	4.4	3.1–6.1	0.61	0.006
Total	248	15.5	5.1	4.5–5.7	–	–

* Includes 5 patients who are Middle Eastern or African (see Patients and Methods). JIA = juvenile idiopathic arthritis; 95% CI = 95% confidence interval; Ref = reference.

Table 2. Demographics, subtype, and poor prognostic features at diagnosis of patients with JIA*

Demographics	Total (n = 248)	Asian (n = 38)	Maori (n = 28)	Pacific Island (n = 20)	European (n = 157)	Maori or Pacific Island vs. European, OR (95% CI); *P*
Age at diagnosis, mean ± SD years	8.9 ± 5	9.0 ± 4	9.2 ± 5	10.6 ± 4	8.6 ± 5	N/A; 0.13
Female	152 (61)	21 (55)	21 (75)	14 (70)	93 (59)	1.9 (0.9–3.8); 0.09
Socioeconomic status						
High	98 (40)	11 (29)	3 (11)	1 (5)	82 (52)	Ref.
Mid	95 (38)	22 (58)	8 (29)	3 (15)	59 (38)	3.8 (1.2–13); 0.03
Low	55 (22)	5 (13)	17 (61)	16 (80)	16 (10)	42 (13–136); <0.0001
ILAR diagnostic criteria						
Oligoarticular	117 (47)	16 (42)	10 (36)	5 (25)	83 (53)	0.4 (0.2–0.8); 0.01
Polyarticular	65 (26)	11 (29)	11 (39)	6 (30)	37 (24)	1.8 (0.9–3.6); 0.13
ERA	38 (15)	7 (18)	2 (7)	3 (15)	25 (16)	0.6 (0.2–1.7); 0.48
Systemic JIA	20 (8)	2 (5)	4 (14)	4 (20)	10 (6)	2.9 (1.1–7.9); 0.04
Psoriatic arthritis	8 (3)	2 (5)	1 (4)	2 (10)	2 (1)	5.2 (0.8–32); 0.09
Polyarticular (RF+)	17 (7)	4 (11)	4 (14)	4 (20)	5 (3)	6.1 (1.9–20); 0.003
No. children w/poor prognostic features						
Total	89 (36)	17 (45)	18 (64)	10 (50)	43 (27)	3.7 (1.9–7.3); 0.0001
Oligoarticular (n = 117)	28 (24)	7 (44)	5 (50)	1 (20)	14 (17)	3.3 (1.0–11); 0.07
Polyarticular (n = 65)	42 (65)	8 (73)	8 (73)	5 (83)	21 (57)	2.5 (0.6–12); 0.23
Systemic JIA (n = 20)	18 (90)	2 (100)	4 (100)	4 (100)	8 (80)	N/A; 0.48
Sacroiliitis (n = 9)	1 (11)	0 (0)	1 (100)	0 (0)	0 (0)	N/A; 0.25
No. poor prognostic features	132	21 (16)	32 (24)	21 (16)	58 (44)	
No. poor prognostic features per child, mean ± SD	0.53 ± 0.9	0.55 ± 0.7	1.14 ± 1.1	1.05 ± 1.3	0.37 ± 0.7	<0.0001

* Values are the number (%) unless indicated otherwise. JIA = juvenile idiopathic arthritis; OR = odds ratio; 95% CI = 95% confidence interval; N/A = not applicable; Ref. = reference; ILAR = International League of Associations for Rheumatology; ERA = enthesitis-related arthritis; RF+ = rheumatoid factor positive.

Poor prognostic features at diagnosis were present in 36% of children with JIA, including 65% (42 of 65) of children with polyarticular disease and 24% (28 of 117) of children with oligoarticular disease. The frequency of poor prognostic features among Maori or Pacific Island children with JIA was higher than among European children (58% versus 27%; OR 3.7 [95% CI 1.9–7.3]) (Table 2), and the difference was not largely reduced after adjusting for SES (OR 3.1 [95% CI 1.4–7.2]). The majority of children with systemic JIA (90%) presented with poor prognostic features with no significant ethnic differences, while a single Maori child with ERA presented with erosive sacroiliitis (Table 2). A total of 132 poor prognostic features were present, with an average of 0.53 per child. Maori or Pacific Island children with JIA (1.10 versus 0.37) (Table 2), including those with oligoarticular or polyarticular disease (1.28 versus 0.40), presented with significantly more poor prognostic features per child compared to European children. The effect of ethnicity was essentially unchanged when adjusted for SES. These included cervical spine involvement (25% versus 9%), erosive disease (22% versus 8%), joint space narrowing (13% versus 2%), and RF positive polyarticular disease (47% versus 14%) (Table 3).

DISCUSSION

While the overall incidence (5.1/100,000) and incidence of the varying subtypes of JIA in children in New Zealand were similar to rates in Europe and North America (16), there were significant ethnic differences of incidence, subtype, and prognostic features at presentation.

The incidence of JIA and the oligoarticular subtype was significantly higher among European children compared to Maori or Pacific Island children. This study confirmed the anecdotal suspicion that, despite this difference, Maori and Pacific Island children exhibit significantly more poor prognostic features at diagnosis. Maori and Pacific Island children were also more likely to present with RF polyarticular disease, radiographic changes (erosions and joint space narrowing), and cervical involvement associated with oligoarticular or polyarticular disease.

Previous studies have shown that early aggressive management of polyarticular JIA results in achievement of clinically inactive disease by a substantial proportion of patients (17). It is also well recognized that Maori and Pacific Island individuals may have difficulty with access to medical care (18) that may lead to a delay in diagnosis, which accounts for some of the ethnic

Table 3. Oligoarticular/polyarticular JIA and poor prognostic features at diagnosis*

	Total	Asian	Maori	Pacific	European	Maori or Pacific vs. European, OR (95% CI); *P*
Total	109/132 (83)	19/109 (17)	25/109 (23)	16/109 (15)	48/109 (45)	–
Hip†	18/182 (10)	3/27 (11)	3/21 (14)	3/11 (27)	8/120 (7)	3.2 (1.0–10); 0.08
Cervical spine†	23/182 (13)	4/27 (15)	6/21 (29)	2/11 (18)	11/120 (9)	3.3 (1.2–9.1); 0.03
Erosion†	18/182 (10)	1/27 (4)	4/21 (19)	3/11 (27)	10/120 (8)	3.1 (1.1–8.9); 0.05
Joint space narrowing†	8/182 (4)	2/27 (7)	3/21 (14)	1/11 (9)	2/120 (2)	8.4 (1.5–48); 0.02
Ankle, marked or prolonged inflammatory marker elevation‡	10/117 (9)	2/16 (13)	2/10 (20)	0/5 (0)	6/83 (7)	2.0 (0.4–11); 0.60
Wrist, marked or prolonged inflammatory marker elevation‡	2/117 (2)	0/16 (0)	1/10 (10)	0/5 (0)	1/83 (1)	5.9 (0.3–99); 0.28
RF+ §	17/65 (26)	4/11 (36)	4/11 (36)	4/6 (67)	5/37 (14)	5.7 (1.2–27); 0.01
Anti-CCP positive§	13/65 (20)	3/11 (27)	2/11 (18)	3/6 (50)	5/37 (14)	2.7 (0.5–14); 0.26
Average no. poor prognostic features/child, mean ± SD	0.60 ± 0.9	0.70 ± 0.8	1.19 ± 1.2	1.45 ± 1.6	0.40 ± 0.7	<0.0001

* Values are the number of patients with oligoarticular and polyarticular juvenile idiopathic arthritis (JIA) with poor prognosis feature divided by the applicable number or patients (%), unless indicated otherwise. OR = odds ratio; 95% CI = 95% confidence interval; RF+ = rheumatoid factor positive; Anti-CCP = anti–cyclic citrullinated peptide.
† Oligoarticular and polyarticular JIA.
‡ Oligoarticular JIA.
§ Polyarticular JIA.

differences. Maori and Pacific Island children in New Zealand also have a high incidence of osteomyelitis (19) and rheumatic fever (20). While it is possible that a presumed diagnosis with osteomyelitis, septic arthritis, or acute rheumatic fever may delay JIA diagnosis, more information is required to establish ethnic-specific diagnostic delay and further delineate the mechanisms of delay.

Maori and Pacific Island children face barriers of health care access, including socioeconomic factors, which may contribute to both a lower incidence and higher severity of disease at JIA diagnosis. The differences demonstrated between European and Maori and Pacific Island children with poor prognostic features at JIA diagnosis were, however, still significant when corrected for social deprivation, which demonstrates an ethnicity effect independent of SES. Therefore, other ethnic-specific factors, including genetic variability, may contribute to prognostic features at JIA diagnosis.

Multiple risk loci for JIA have been identified by candidate gene approaches and genetic studies; however, there are many complicating factors when searching for casual variants. Recently whole-genome sequencing has found new genetic variations and identified epigenetic landscapes surrounding the genetic mutations that may provide information regarding JIA disease mechanism (21).

There were a number of limitations to this study, including its retrospective nature and potential confounders. This study may have underestimated the incidence and prevalence of JIA, which is likely much higher if reviewed on a population-wide basis. Although almost all pediatric care is provided in the public hospitals in New Zealand, it is possible that a small number of young people with JIA may have been seen exclusively in private practice. However, an accessible public service combined with a single national private clinician who referred all patients with JIA to the public service significantly limits the number of children with JIA who may not have been captured by this review. Auckland, although ethnically diverse, does not have the highest ratio of Maori population compared to other parts of New Zealand such as Gisborne, where almost 50% of the population is Maori. Therefore, it is possible that our study may have underestimated some of the differences between Maori and European children.

In summary, we have established the current overall and ethnic incidence of JIA in a New Zealand cohort. We established ethnic differences regarding poor prognostic features at presentation between European and Maori and Pacific Island children. While environmental factors play a role in these differences, genetic comparisons are particularly important especially among individual ethnic groups. Future studies targeting the genetic basis for the disease may provide further important information.

ACKNOWLEDGMENTS

The authors thank Drs. Sue Rudge and Jacqueline Yan for review of the initial proof of concept and their general advice and 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 published. Dr. Concannon 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. Concannon.
Acquisition of data. Concannon, Ostring.
Analysis and interpretation of data. Concannon, Reed, Ostring.

REFERENCES

1. Thierry S, Fautrel B, Lemelle I, Guillemin F. Prevalence and incidence of juvenile idiopathic arthritis: a systematic review. Joint Bone Spine 2014;81:112–7.
2. McDougall C, Hurd K, Barnabe C. Systematic review of rheumatic disease epidemiology in the indigenous populations of Canada, the United States, Australia, and New Zealand. Sem Arthritis Rheum 2017;46:675–86.
3. Bolen J, Schieb L, Hootman JM, Helmick CG, Theis K, Murphy LB, et al. Differences in the prevalence and severity of arthritis among racial/ethnic groups in the United States, National Health Interview Survey, 2002, 2003, and 2006. Prev Chronic Dis 2010;7:A64.
4. Concannon A, Rudge S, Yan J, Reed P. The incidence, diagnostic clinical manifestations and severity of juvenile systemic lupus erythematosus in New Zealand Maori and Pacific Island children: the Starship experience (2000-2010). Lupus 2013;22:1156–61.
5. McGill NW, Gow PJ. Juvenile rheumatoid arthritis in Auckland: a long term follow-up study with particular reference to uveitis. Aust N Z J Med 1987;17:305–8.
6. Beukelman T, Patkar NM, Saag KG, Tolleson-Rinehart S, Cron RQ, DeWitt EM, et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. Arthritis Care Res (Hoboken) 2011;63:465–82.
7. Statistics New Zealand. NZ.STAT user guide. URL: http://nzdotstat.stats.govt.nz/wbos/index.aspx.
8. 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.
9. Ministry of Health. HISO 10001:2017 Ethnicity data protocols. URL: http://www.health.govt.nz/publication/ethnicity-data-protocols-health-and-disability-sector.
10. Statistics New Zealand. Review of the official ethnicity statistial standard 2009. URL: http://archive.stats.govt.nz/browse_for_stats/people_and_communities/households/review-of-the-official-ethnicity-statistical-standard-2009.aspx).
11. University of Otago Wellington. Socioeconomic deprivation indexes: NZDep, and NZiDep, Department of Public Health: a comparison of the NZDep index of socioeconomic deprivation and the Index of Multiple Deprivation (IMD). URL: https://www.otago.ac.nz/wellington/departments/publichealth/research/hirp/otago020194.html
12. Salmon C, Crampton P. The development of New Zealand's deprivation index (NZDep) and its uptake as a national policy tool. Can J Public Health 2012;103(8 Suppl 2):S7–11.
13. Statistics New Zealand. Meshblock 2001. URL: https://koordinates.com/from/datafinder.stats.govt.nz/layer/25744.
14. Statistics New Zealand. Meshblock 2006. URL: https://koordinates.com/from/datafinder.stats.govt.nz/layer/25746.
15. Statistics New Zealand. Meshblock 2013. URL: https://koordinates.com/from/datafinder.stats.govt.nz/layer/8347.
16. Barut K, Adrovic A, Sahin S, Kasapcopur O. Juvenile idiopathic arthritis. Balkan Med J 2017;34:90–101.
17. Wallace CA, Giannini EH, Spalding SJ, Hashkes PJ, O'Neil KM, Zeft AS, et al. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. Arthritis Rheum 2012;64:2012–21.
18. Ellison-Loschmann L, Pearce N. Improving access to health care among New Zealand's Maori population. Am J Public Health 2006;96:612–7.
19. Street M, Crawford H. Pediatric humeral osteomyelitis. J Pediatr Orthop 2015;35:628–33.
20. Gurney J. The incidence of acute rheumatic fever in New Zealand, 2010-2013. N Z Med J 2015;128:65–7.
21. Wong L, Jiang K, Chen Y, Jarvis JN. Genetic insights into juvenile idiopathic arthritis derived from deep whole genome sequencing. Sci Rep 2017;7:2657.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 71, No. 9, September 2019, pp 1276–1283
DOI 10.1002/acr.23765

Left Ventricular Systolic Myocardial Function in Ankylosing Spondylitis

Helga Midtbø,[1] Anne G. Semb,[2] Knut Matre,[3] Silvia Rollefstad,[2] Inger J. Berg,[2] and Eva Gerdts[3]

Objective. Subclinical left ventricular (LV) myocardial dysfunction is associated with an increased risk of cardiovascular disease (CVD), but it is not known whether subclinical LV myocardial dysfunction is present in patients with ankylosing spondylitis (AS) independent of CVD risk factors.

Methods. Conventional and speckle tracking echocardiography were performed in 106 patients with AS (mean ± SD age 48 ± 12 years; 59% men) and 106 matched controls (mean ± SD age 51 ± 12 years; 59% men). LV systolic myocardial function was assessed by peak systolic global longitudinal strain (GLS).

Results. CVD risk factors were similarly distributed in patients with AS and controls, but more controls received statin therapy ($P = 0.05$). GLS was significantly lower in patients with AS compared to controls (mean ± SD –17.7 ± 2.5% versus –18.4 ± 2.3%; $P = 0.03$). In univariable linear regression analyses in the total study population, lower GLS was associated with having AS, male sex, higher body mass index, higher LV mass index, and lower LV ejection fraction (all $P < 0.05$). Having AS retained an independent association with lower GLS when adjusted for these factors in multivariable analyses ($\beta = 0.16$, $P = 0.02$). In patients with AS, lower GLS was independently associated with larger aortic root diameter in multivariable analyses ($\beta = 0.24$, $P = 0.02$), while no association with AS disease activity, disease duration, or use of antirheumatic medication was observed.

Conclusion. Patients with AS had lower GLS compared with controls, independent of confounders. In AS patients, lower GLS was associated with larger aortic root diameter. Prospective studies should test whether lower GLS contributes to the observed higher CVD risk in patients with AS.

INTRODUCTION

Ankylosing spondylitis (AS) is an inflammatory joint disease primarily affecting the sacroiliac joints and the spine. Cardiac involvement in AS has been known for a long time, in particular alterations of the aortic root geometry and aortic valve dysfunction (1–3). As in other inflammatory joint diseases, there is an increased risk of premature atherosclerosis in AS (4,5). However, the risk of myocardial infarction has been shown to be only one-half as high in AS patients compared to that observed in patients with rheumatoid arthritis (RA) (4,6). Cardiovascular disease (CVD) risk factors, such as hypertension, diabetes mellitus, and metabolic syndrome, are prevalent in patients with AS (7,8) and have been suggested to largely explain the increased risk of CVD in such patients (6). However, it is well known that the presence of subclinical cardiac dysfunction such as left ventricular (LV) myocardial dysfunction, as measured by global longitudinal strain (GLS), predicts increased CVD risk independent of risk factors in unselected cardiac populations (9,10). Furthermore, GLS has been shown to outperform traditional measures of LV systolic function (e.g., ejection fraction) in the prediction of all-cause mortality and of a composite of CVD mortality, heart failure hospitalization, and malignant arrhythmias (9,10). Low GLS has also been shown to predict CVD events such as myocardial infarction, ischemic stroke, and CVD-related death in a population-based cohort (11). Assessment of GLS in clinical practice is now recommended by guidelines (12), since detection of low GLS is associated with increased CVD mortality and morbidity, irrespective of a normal ejection fraction and prevalent traditional CVD risk factors (11).

A recent study in patients with AS demonstrated that increased disease severity assessed by the modified Stoke Ankylosing Spondylitis Spine Score was associated with reduced GLS (13). However, important confounders of reduced GLS, such as hypertension and LV hypertrophy, were uncom-

Supported by the South-Eastern Norway and Western Norway Regional Health Authorities.

[1]Helga Midtbø, MD, PhD: Haukeland University Hospital, Bergen, Norway; [2]Anne G. Semb, MD, PhD, Silvia Rollefstad, MD, PhD, Inger J. Berg, MD, PhD: Diakonhjemmet Hospital, Oslo, Norway; [3]Knut Matre, MSc, PhD, Eva Gerdts, MD, PhD: University of Bergen, Bergen, Norway.

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

Address correspondence to Helga Midtbø, MD, PhD, Department of Heart Disease, Haukeland University Hospital, Jonas Lies vei 65, 5021 Bergen, Norway. E-mail: helgamit@online.no.

Submitted for publication March 26, 2018; accepted in revised form September 18, 2018.

SIGNIFICANCE & INNOVATIONS

- Patients with ankylosing spondylitis (AS) had lower left ventricular (LV) systolic myocardial function compared to controls, independent of cardiovascular risk factors.
- Lower LV systolic myocardial function was associated with a larger aortic root diameter in patients with AS, but not with disease activity, disease duration, or treatment with antirheumatic medications.
- The results of the present substudy may help explain the increased risk of cardiovascular disease in patients with AS.

mon in this Chinese study population, which is suggestive of the need for additional studies. The aim of the present substudy was to provide further information on presence and confounders of subclinical LV systolic myocardial function measured by GLS in patients with AS.

MATERIALS AND METHODS

Study population. Patients with AS were recruited from an established cohort at the Department of Rheumatology of Diakonhjemmet Hospital in Oslo, Norway, as previously described (14–16). All patients were diagnosed according to the modified New York criteria (17). A total of 257 patients were invited, and the response rate was 62% (159 patients). For the present substudy, 17 patients were excluded due to established CVD (defined as previous cardiac surgery or intervention, angina pectoris, myocardial infarction, cerebral infarction, transitory ischemic attack, or intermittent claudication), and an additional 36 patients were excluded because of insufficient quality of the echocardiographic images for speckle tracking echocardiography (STE) to be performed. Thus, a total of 106 patients with AS were eligible for the present STE substudy.

Control subjects. Statistics Norway randomly selected control subjects stratified for age, sex, and residential area to the patients with AS. The only exclusion criterion was having an inflammatory joint disease. The response rate was 40% among the invited control subjects (329 invited, 132 agreed to participate). Among these, 6 patients with established CVD were excluded, and an additional 20 were excluded because of insufficient echocardiographic quality for performance of STE, leaving 106 control subjects eligible for the present substudy.

All participants signed informed consent in accordance with the Declaration of Helsinki. The study protocol was approved by the South-Eastern Norwegian Regional Committee for Medical and Health Research Ethics.

AS disease characteristics. The disease duration was defined from the onset of symptoms, as recommended (18). AS disease activity was assessed by the Ankylosing Spondylitis Disease Activity Score using the C-reactive protein (ASDAS-CRP) level (19), and the Bath Ankylosing Disease Activity Index (BASDAI) (20). Functional limitation was assessed by the Bath Ankylosing Spondylitis Functional Index (BASFI) (21), and HLA–B27 status was obtained from medical records.

Assessment of CVD risk factors. Participant medical history, smoking status, and current medication were collected on a standardized questionnaire. Quality assurance of the information was provided by the consultant cardiologist (AGS) during the outpatient consultation.

Brachial blood pressure was measured using an Omron M7 apparatus. As recommended in the joint European Society of Hypertension/European Society of Cardiology guidelines for management of hypertension, the average of the last 2 measurements was reported as the clinic blood pressure (22). Hypertension was defined as a history of hypertension, use of antihypertensive medication, or elevated blood pressure (office systolic blood pressure ≥140 mm Hg and/or office diastolic blood pressure ≥90 mm Hg).

Obesity was defined as a body mass index ≥30 kg/m^2. Metabolic syndrome was defined according to The American Heart Association/National Heart, Lung, and Blood Institute criteria (23), if at least 3 of 5 criteria were present, including waist circumference ≥88 cm in women and ≥102 cm in men, triglycerides >150 mg/dl, high-density lipoprotein (HDL) cholesterol <50 mg/dl in women and <40 mg/dl in men, systolic blood pressure ≥130 mm Hg and/or diastolic blood pressure ≥85 mm Hg and/or treatment with antihypertensives, fasting blood glucose >100 mg/dl and/or antidiabetic treatment. Hypercholesterolemia was defined as low-density lipoprotein (LDL) cholesterol ≥190 mg/dl or statin treatment (24).

Laboratory measurements. Triglycerides, HDL cholesterol, and CRP levels were analyzed in fasting blood samples with a Cobas 6000 machine (Roche Diagnostics). LDL cholesterol was calculated using the Friedewald formula (25).

Echocardiography. Standardized transthoracic echocardiography was performed at the Preventive Cardio-Rheuma clinic in the Department of Rheumatology at Diakonhjemmet Hospital in 2008–2010 using a Vivid 7 ultrasound scanner (General Electric). All echocardiograms were stored digitally and transferred for expert analysis at the Echocardiography Core Laboratory at the University of Bergen, Bergen, Norway. Images were analyzed offline on dedicated workstations equipped with Image Arena software, version 4.4 (TomTec Imaging Systems). The digitally stored echocardiograms were all analyzed by the same reader (HM) and later proofread by the same highly experienced reader (EG). Readers were blinded to the presence or absence of AS. Conventional quantitative echocardiography was performed

in accordance with the joint guidelines from the European Association of Cardiovascular Imaging and American Society of Echocardiography (12). LV mass was indexed for height in the allometric power of 2.7 (height$^{2.7}$) (26). Aortic root diameter was measured as inner diameter in end-diastole at the sinus of Valsalva (27). LV systolic function was estimated by ejection fraction using biplane Simpson's method. Diastolic function was evaluated by the ratio of early transmitral–filling rate: late–filling rate (A), and pulsed Doppler tissue imaging at the septal mitral annulus in 4-chamber view for calculation of the ratio between early transmitral–filling rate and early septal mitral annulus velocity (e′, E/e′ ratio) (28).

LV systolic myocardial function was analyzed by STE with EchoPAC BT113 software (Vingmed ultrasound, General Electric) and automated function imaging. Since the patients/controls with known CVD or regional wall motion abnormalities were excluded, longitudinal strain from the 4-chamber and apical long-axis view was considered representative for the whole LV (29). The cardiac cycle with the best image quality was selected for the analyses. End-systole was defined by aortic valve closure visualized in the apical long-axis view. The endocardial border was traced automatically, and the region of interest was adjusted to include the entire LV myocardium, but avoid the pericardium. The software then automatically tracked the movement of speckles from frame to frame. Quality of tracking was assessed visually and if the tracking was poor, the segment with poor tracking was excluded. Peak systolic longitudinal strain was reported in the 4-chamber and apical long-axis view, and the average value was reported as GLS. GLS measures LV deformation during systole. GLS is a negative percentage value, since the LV length shortens during contraction in the apical views. Therefore, less negative GLS indicates lower LV systolic myocardial function (29).

All STE measurements were performed in 2017 by the same researcher (HM). Intraobserver variability was assessed in 30 randomly selected patients by repeated measurements on

Table 1. Clinical characteristics and CVD risk factors in the study populations*

	AS patients (n = 106)	Controls (n = 106)	*P*
Age, mean ± SD years	48.0 ± 12.2	51.1 ± 11.5	0.06
Men	63 (59)	63 (59)	1.0
CVD risk factors			
Hypertension	32 (30)	38 (36)	0.38
Current smoking	17 (16)	24 (23)	0.22
Diabetes mellitus	4 (4)	2 (2)	0.41
BMI, mean ± SD kg/m^2	24.7 ± 3.2	25.1 ± 3.5	0.41
Obesity	5 (5)	11 (10)	0.12
Metabolic syndrome	11 (12)	10 (10)	0.67
Hypercholesterolemia	8 (8)	14 (13)	0.25
CRP, median (IQR) mg/liter	3.0 (1.0, 8.5)	1.0 (1.0, 2.0)	<0.001
Medication			
Antihypertensive	18 (17)	16 (15)	0.71
Statin therapy	3 (3)	10 (9)	0.05
Prednisolone	9 (9)	2 (2)	0.03
DMARDs	16 (15)	0 (0)	<0.001
NSAIDs	69 (65)	14 (13)	<0.001
TNF inhibitor	21 (20)	0 (0)	<0.001
AS-specific characteristics			
Disease duration, mean ± SD years	22 ± 11	NA	NA
HLA–B27 positive	83 (93)	NA	NA
ASDAS, mean ± SD	2.2 ± 1.0	NA	NA
BASDAI score, mean ± SD	3.6 ± 1.9	NA	NA
BASFI score, mean ± SD	2.3 ± 2.1	NA	NA

* Values are the number (%) unless indicated otherwise. CVD = cardiovascular disease; AS = ankylosing spondylitis; BMI = body mass index; CRP = C-reactive protein; IQR = interquartile range; DMARDs = disease-modifying antirheumatic drugs; NSAIDs = nonsteroidal antiinflammatory drugs; TNF = tumor necrosis factor; NA = not applicable; ASDAS = Ankylosing Spondylitis Disease Activity Score; BASDAI = Bath Ankylosing Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index.

Table 2. Echocardiographic parameters in the study population*

	AS patients (n = 106)	Controls (n = 106)	*P*
Conventional echocardiography			
LV end-diastolic diameter, cm	4.94 ± 0.55	4.88 ± 0.52	0.35
Interventricular septum thickness at end-diastole, cm	0.97 ± 0.21	0.95 ± 0.18	0.29
LV mass index, gm/m$^{2.7}$†	35.8 ± 9.7	34.1 ± 10.2	0.22
Aortic root diameter, cm	3.2 ± 0.4	3.1 ± 0.4	0.44
Aortic regurgitation (any), no. (%)	16 (15)	17 (16)	0.85
Mitral regurgitation (any), no. (%)	52 (49)	53 (50)	0.89
Ejection fraction, %	66 ± 6	67 ± 5	0.05
Diastolic function			
E/A ratio	1.4 ± 0.4	1.4 ± 0.5	0.43
Septal e', cm/second	8.3 ± 2.6	8.4 ± 2.4	0.72
E/e'	8.3 ± 3.2	8.1 ± 2.7	0.63
Speckle tracking echocardiography, GLS, %	–17.7 ± 2.5	–18.4 ± 2.3	0.03

* Values are the mean ± SD unless indicated otherwise. AS = ankylosing spondylitis; LV = left ventricular; E/A = early/atrial transmitral peak velocities; e' = early septal mitral annular velocity; E/e' = early transmitral peak velocity/ early septal mitral annular velocity; GLS = global longitudinal strain.
† LV mass as LV mass indexed for height in the allometric power of 2.7 (height$^{2.7}$).

the same cine loop as the original measurements. Mean ± SD frame rate was 76 ± 7 frames/second.

Statistical analysis. The statistical analyses were performed using IBM SPSS statistics, version 23.0. Continuous data are expressed as the mean ± SD for normally distributed variables, and median and interquartile range for non-normally distributed variables. Categorical variables are presented as percentages and numbers. CRP level was non-normally distributed, and was therefore log transformed before comparison. Group comparisons were done using the chi-square test or Student's 2-sample *t*-test, as appropriate. Univariable associations of GLS were assessed with linear regression models, and results were reported as standardized beta coefficients and *P* values. Multivariable linear regression analyses were run with an enter method and collinearity tools, and results were reported as multiple R^2 for the models, and standardized beta coefficients and *P* values for the individual variables. Reproducibility of GLS measurement was assessed by the intraclass correlation coefficient. Two-tailed *P* values less than 0.05 were considered significant.

RESULTS

Clinical characteristics and CVD risk factors. The CVD risk factor burden was equally distributed between patients with AS and controls, but more controls received statin therapy (*P* = 0.05) (Table 1). Hypertension was the most prevalent CVD risk factor and was present in 30% of patients with AS and 36% of control subjects (*P* = 0.38) (Table 1). As expected, CRP levels were higher in patients with AS than controls (Table 1). Disease duration was long, at an average of 22 years in the patients with AS. Mean disease activity was moderate to high based on the mean ± SD ASDAS and BASDAI scores, and the functional capacity was good based on the mean ± SD BASFI score (Table 1).

LV myocardial function. Echocardiographic characteristics of patients with AS and controls are shown in Table 2. Notably, the prevalence of valvular regurgitation was similar between groups, and none of the patients with AS or controls had more than mild-moderate regurgitation. Mean LV ejection fraction was normal both in patients with AS and controls (Table 2). However, GLS was significantly lower in AS patients compared to controls (*P* = 0.03) (Table 2 and Figure 1). When excluding the 21 patients with AS who were treated with tumor necrosis factor alpha-inhibitors from the analyses, GLS still remained lower in patients with AS versus controls (mean ± SD –17.8 ± 2.3 versus –18.4

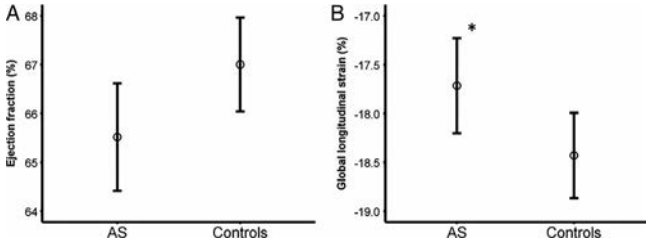

Figure 1. Left ventricular systolic function in patients with ankylosing spondylitis (AS) and controls. **A**, Ejection fraction. **B**, Global longitudinal strain. * = *P* < 0.05.

± 2.3; $P = 0.05$). The reproducibility of GLS measurements was excellent (intraclass correlation coefficient 0.95 [95% confidence interval 0.80–0.98]).

Covariables of GLS. In the total study population, lower GLS was associated with having AS, male sex, higher body mass index, higher LV mass index and larger aortic root diameter, and with lower LV ejection fraction (all $P < 0.05$) (Table 3). When analyzing patients with AS and controls separately, larger aortic root diameter was only associated with lower GLS in the patients with AS (β = 0.32, $P = 0.001$) and not in the control group (β = 0.10, $P = 0.32$). Having AS remained independently associated with lower GLS after adjustment for covariables in multivariable analyses (β = 0.16, $P = 0.02$) (Table 4).

Among the patients with AS, male sex, larger aortic root diameter, and higher LV mass index emerged as the strongest covariables of lower GLS in univariable analyses ($P < 0.05$) (Table 3). AS–specific factors, such as the use of prednisolone, tumor necrosis factor inhibitors, synthetic disease-modifying antirheumatic drugs or nonsteroidal antiinflammatory drugs, disease duration, ASDAS, and BASDAI and BASFI scores were not associated with lower GLS in univariable analyses (Table 3). In multivariable analysis

Table 3. Univariable associations of GLS with CVD risk factors, medication, echocardiographic parameters, and AS–specific characteristics*

	Total study population (n = 212)		AS patients (n = 106)	
	β	*P*	β	*P*
Having AS	0.15	0.03	NA	NA
Age, years	−0.11	0.11	−0.07	0.48
Men	0.22	0.002	0.23	0.02
CVD risk factors				
Hypertension	0.01	0.96	0.01	0.96
Smoking	−0.04	0.53	−0.09	0.34
Body mass index, kg/m^2	0.16	0.02	0.17	0.08
Metabolic syndrome	0.13	0.07	0.11	0.27
Hypercholesterolemia	0.05	0.52	0.02	0.87
Log-transformed CRP, mg/liter	0.13	0.06	0.09	0.35
Medication				
Antihypertensive	0.12	0.09	0.05	0.65
Statin therapy	0.12	0.08	0.14	0.16
NSAIDs	0.05	0.51	−0.02	0.84
Prednisolone	−0.01	0.87	−0.01	0.89
DMARDs	NA	NA	0.14	0.15
TNF inhibitor	NA	NA	0.06	0.53
Echocardiographic parameters				
Aortic root diameter, cm	0.21	0.002	0.32	0.001
Aortic regurgitation	−0.08	0.23	−0.18	0.23
LV mass index, gm/m$^{2.7}$†	0.19	0.005	0.23	0.02
Ejection fraction, %	−0.16	0.02	−0.16	0.11
Septal e', cm/second	−0.14	0.04	−0.19	0.06
E/e'	0.03	0.68	0.06	0.55
AS-specific characteristics				
Disease duration, years	NA	NA	−0.03	0.79
HLA–B27 positive	NA	NA	−0.05	0.66
ASDAS	NA	NA	0.06	0.55
BASDAI score	NA	NA	0.06	0.58
BASFI score	NA	NA	0.05	0.62
Uveitis	NA	NA	0.02	0.81

* GLS = global longitudinal strain; LV = left ventricular; e' = early septal mitral annular velocity; E/e' = early transmitral peak velocity/ early septal mitral annular velocity (See Table 1 for other definitions).
† LV mass as LV mass indexed for height in the allometric power of 2.7 (height$^{2.7}$).

Table 4. Multivariable linear regression analyses of GLS*

	Total study population (n = 212)†		AS patients (n = 106)‡	
	β	*P*	β	*P*
Having AS	0.16	0.02	–	–
Receipt of statin	0.12	0.09	–	–
Body mass index, kg/m^2	0.10	0.18	–	–
Men	0.14	0.04	0.09	0.39
Ejection fraction	−0.13	0.06	−0.12	0.22
LV mass index, gm/m$^{2.7}$	0.07	0.40	0.11	0.28
Aortic root diameter, cm	–	–	0.24	0.002

* GLS = global longitudinal strain; AS = ankylosing spondylitis; LV = left ventricular.
† Multiple R^2 = 0.12, *P* = 0.001.
‡ Multiple R^2 = 0.14, *P* = 0.004.

among patients with AS, lower GLS was independently associated with larger aortic root diameter independent of confounders (β 0.24, *P* = 0.02) (Table 4).

DISCUSSION

The present substudy demonstrates that patients with AS in whom CVD is unknown have lower LV systolic myocardial function than controls, as measured by GLS both in univariable analysis, and after adjustment for CVD risk factors and other confounders in multivariable analyses. Lower LV systolic myocardial function was associated with larger aortic root diameter in patients with AS, but not with AS disease activity or AS specific medication.

In the present study, the lower GLS found among patients with AS is in line with previous studies reporting subclinical myocardial dysfunction in other inflammatory joint diseases (30,31). In patients with AS, few previous studies have assessed GLS (13,32). In the study by Chen et al (13), which included 104 Chinese patients with axial spondylitis, GLS was lower in patients with AS compared to controls. However, only 79% of patients in the study had radiologically verified AS, and the prevalence of hypertension and hypercholesterolemia was low (13). A small Turkish study that included a cohort of 26 patients with AS also found lower GLS in patients with AS compared to controls (32). However, that study did not include multivariable analyses, due to limited statistical power (32). Taken together, the present substudy expands previous knowledge regarding subclinical myocardial dysfunction in AS by showing that lower GLS persists among patients with AS also in populations with a high prevalence of CVD risk factors, including hypertension and metabolic syndrome. Lower GLS is an independent risk factor for development of CVD, even in patients with normal ejection fraction (11), and the current results indicate that particular attention should be given to CVD risk screening in patients with AS.

Notably, having AS was associated with lower GLS independent of LV mass index in the present substudy. Higher LV mass index has previously been associated with increased focal cardiac fibrosis and reduced LV function in the general population (33). The gold standard method for detection of focal cardiac fibrosis is late gadolinium enhancement cardiac magnetic resonance imaging, but it has been demonstrated that GLS is a sensitive surrogate marker of focal myocardial fibrosis in cardiac magnetic resonance studies in patients with cardiomyopathies (34). A small exploratory study by Biesbroek et al (35), which used cardiac magnetic resonance in 14 patients with AS, demonstrated that 21% of AS patients had presence of late gadolinium enhancement, which was indicative of focal cardiac fibrosis. Further, fibrosis was not present in an ischemic pattern, but was localized in the midwall; this is a pattern of myocardial fibrosis that is also seen in other inflammatory diseases (36). In addition, larger myocardial extracellular volume, a marker of diffuse cardiac fibrosis, correlated with higher CRP level, suggesting that cardiac fibrosis occurred as a consequence of AS disease activity (35). Similar results were shown in a study by Kobayashi et al (37) of 60 patients with RA, in which the presence of late gadolinium enhancement by cardiac magnetic resonance imaging was associated with increased inflammatory disease activity (37). We have previously demonstrated that higher disease activity in patients with RA was associated with lower GLS (30). In contrast, we did not find an association between GLS and CRP level or AS disease activity in the present substudy.

In patients with AS, the strongest covariable of lower GLS in multivariable analyses was larger aortic root diameter. A similar association was not found in the control group. Pathophysiologic aortic root involvement with dilatation and thickening has been recognized in patients with AS for a long time (1), but the cause of aortic root abnormalities in AS is unknown. Further research is needed to better characterize the association between aortic root diameter and LV systolic myocardial function.

Some study limitations should be mentioned. The cross-sectional study design was unsuited to claim any causality between having AS and lower GLS. The current substudy did not have sufficient statistical power to compare GLS among subgroups of patients with AS. The low participation rate among the invited controls could have introduced a selection bias. However, a strength of the study is that we used a core laboratory for advanced imaging analysis, as recommended (38).

In conclusion, patients with AS had lower LV systolic myocardial function assessed by GLS than controls, independent of CVD risk factors, ejection fraction, and LV mass index. Lower GLS was particularly associated with larger aortic root diameter in patients with AS. The current results add to findings in previous studies on increased prevalence of subclinical cardiac disease in patients with AS, which may contribute to the increased CVD risk observed in such 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 published. Dr. Midtbø 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. Midtbø, Semb, Matre, Rollefstad, Berg, Gerdts.

Acquisition of data. Midtbø, Semb, Matre, Rollefstad, Berg,Gerdts.

Analysis and interpretation of data. Midtbø, Semb, Matre, Rollefstad, Berg, Gerdts.

REFERENCES

1. Roldan CA, Chavez J, Wiest PW, Qualls CR, Crawford MH. Aortic root disease and valve disease associated with ankylosing spondylitis. J Am Coll Cardiol 1998;32:1397–404.
2. Klingberg E, Sveälv BG, Täng MS, Bech-Hanssen O, Forsblad-d'Elia H, Bergfeldt L. Aortic regurgitation is common in ankylosing spondylitis: time for routine echocardiography evaluation? Am J Med 2015;128: 1244–50.
3. Midtbø H, Gerdts E, Berg IJ, Rollefstad S, Jonsson R, Semb AG. Ankylosing spondylitis is associated with increased prevalence of left ventricular hypertrophy. J Rheumatol 2018;45:1249–55.
4. Eriksson JK, Jacobsson L, Bengtsson K, Askling J. Is ankylosing spondylitis a risk factor for cardiovascular disease, and how do these risks compare with those in rheumatoid arthritis? Ann Rheum Dis 2017;76:364–70.
5. Wright KA, Crowson CS, Michet CJ, Matteson EL. Time trends in incidence, clinical features, and cardiovascular disease in ankylosing spondylitis over three decades: a population-based study. Arthritis Care Res (Hoboken) 2015;67:836–41.
6. Schieir O, Tosevski C, Glazier RH, Hogg-Johnson S, Badley EM. Incident myocardial infarction associated with major types of arthritis in the general population: a systematic review and meta-analysis. Ann Rheum Dis 2017;76:1396–404.
7. Han C, Robinson DW, Hackett MV, Paramore LC, Fraeman KH, Bala MV. Cardiovascular disease and risk factors in patients with rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. J Rheumatol 2006;33:2167–72.
8. Mathieu S, Gossec L, Dougados M, Soubrier M. Cardiovascular profile in ankylosing spondylitis: a systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2011;63:557–63.
9. Stanton T, Leano R, Marwick TH. Prediction of all-cause mortality from global longitudinal speckle strain: comparison with ejection fraction and wall motion scoring. Circ Cardiovasc Imaging 2009;2:356–64.
10. Kalam K, Otahal P, Marwick TH. Prognostic implications of global LV dysfunction: a systematic review and meta-analysis of global longitudinal strain and ejection fraction. Heart 2014;100:1673–80.
11. Russo C, Jin Z, Elkind MS, Rundek T, Homma S, Sacco RL, et al. Prevalence and prognostic value of subclinical left ventricular systolic dysfunction by global longitudinal strain in a community-based cohort. Eur J Heart Fail 2014;16:1301–9.
12. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. Eur Heart J Cardiovasc Imaging 2015;16:233–70.
13. Chen Y, Chung HY, Zhao CT, Wong A, Zhen Z, Tsang HH, et al. Left ventricular myocardial dysfunction and premature atherosclerosis in patients with axial spondyloarthritis. Rheumatology (Oxford) 2015;54:292–301.
14. Berg IJ, van der Heijde D, Dagfinrud H, Seljeflot I, Olsen IC, Kvien TK, et al. Disease activity in ankylosing spondylitis and associations to markers of vascular pathology and traditional cardiovascular disease risk factors: a cross-sectional study. J Rheumatol 2015;42:645–53.
15. Berg IJ, Semb AG, van der Heijde D, Kvien TK, Hisdal J, Olsen IC, et al. Uveitis is associated with hypertension and atherosclerosis in patients with ankylosing spondylitis: a cross-sectional study. Semin Arthritis Rheum 2014;44:309–13.
16. Berg IJ, Semb AG, van der Heijde D, Kvien TK, Olsen IC, Dagfinrud H, et al. CRP and ASDAS are associated with future elevated arterial stiffness, a risk marker of cardiovascular disease, in patients with ankylosing spondylitis: results after 5-year follow-up. Ann Rheum Dis 2015;74:1562–6.
17. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
18. Feldtkeller E, Erlendsson J. Definition of disease duration in ankylosing spondylitis. Rheumatol Int 2008;28:693–6.
19. Van der Heijde D, Lie E, Kvien TK, Sieper J, Van den Bosch F, Listing J, et al. ASDAS, a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:1811–8.
20. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
21. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994;21:2281–5.
22. Mancia G, Fagard R, Narkiewicz K, Redon J, Zanchetti A, Böhm M, et al. 2013 ESH/ESC guidelines for the management of arterial hypertension. Blood Press 2013;22:193–278.
23. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation 2005;112:2735–52.
24. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts) developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). Eur Heart J 2016;37:2315–81.
25. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18:499–502.
26. De Simone G, Daniels SR, Devereux RB, Meyer RA, Roman MJ, de Divitiis O, et al. Left ventricular mass and body size in normotensive children and adults: assessment of allometric relations and impact of overweight. J Am Coll Cardiol 1992;20:1251–60.
27. Bahlmann E, Nienaber CA, Cramariuc D, Gohlke-Baerwolf C, Ray S, Devereux RB, et al. Aortic root geometry in aortic stenosis patients (a SEAS substudy). Eur J Echocardiogr 2011;12:585–90.
28. Marwick TH, Gillebert TC, Aurigemma G, Chirinos J, Derumeaux G, Galderisi M, et al. Recommendations on the use of echocardiography in adult hypertension: a report from the European Association of Cardiovascular Imaging (EACVI) and the American Society of Echocardiography (ASE). Eur Heart J Cardiovasc Imaging 2015;16:577–605.
29. Voigt JU, Pedrizzetti G, Lysyansky P, Marwick TH, Houle H, Baumann R, et al. Definitions for a common standard for 2D speckle tracking

echocardiography: consensus document of the EACVI/ASE/Industry Task Force to standardize deformation imaging. Eur Heart J Cardiovasc Imaging 2015;16:1–11.
30. Midtbø H, Semb AG, Matre K, Kvien TK, Gerdts E. Disease activity is associated with reduced left ventricular systolic myocardial function in patients with rheumatoid arthritis. Ann Rheum Dis 2017;76:371–6.
31. Ikonomidis I, Tzortzis S, Lekakis J, Paraskevaidis I, Andreadou I, Nikolaou M, et al. Lowering interleukin-1 activity with anakinra improves myocardial deformation in rheumatoid arthritis. Heart 2009;95:1502–7.
32. Ustun N, Kurt M, Nacar AB, Karateke HP, Guler H, Turhanoglu AD. Left ventricular systolic dysfunction in patients with ankylosing spondylitis without clinically overt cardiovascular disease by speckle tracking echocardiography. Rheumatol Int 2015;35:607–11.
33. Ambale Venkatesh B, Volpe GJ, Donekal S, Mewton N, Liu CY, Shea S, et al. Association of longitudinal changes in left ventricular structure and function with myocardial fibrosis: the Multi-Ethnic Study of Atherosclerosis study. Hypertension 2014;64:508–15.
34. Kramer J, Niemann M, Liu D, Hu K, Machann W, Beer M, et al. Two-dimensional speckle tracking as a non-invasive tool for identification of myocardial fibrosis in Fabry disease. Eur Heart J 2013;34:1587–96.
35. Biesbroek PS, Heslinga SC, Konings TC, van der Horst-Bruinsma IE, Hofman MB, van de Ven PM, et al. Insights into cardiac involvement in ankylosing spondylitis from cardiovascular magnetic resonance. Heart 2017;103:745–52.
36. Ntusi NA, Piechnik SK, Francis JM, Ferreira VM, Matthews PM, Robson MD, et al. Diffuse myocardial fibrosis and inflammation in rheumatoid arthritis: insights from CMR T1 mapping. JACC Cardiovasc Imaging 2015;8:526–36.
37. Kobayashi H, Kobayashi Y, Yokoe I, Akashi Y, Takei M, Giles JT. Magnetic resonance imaging–detected myocardial inflammation and fibrosis in rheumatoid arthritis: associations with disease characteristics and N-terminal pro-brain natriuretic peptide levels. Arthritis Care Res (Hoboken) 2017;69:1304–11.
38. Galderisi M, Henein MY, D'Hooge J, Sicari R, Badano LP, Zamorano JL, et al. Recommendations of the European Association of Echocardiography: how to use echo-Doppler in clinical trials: different modalities for different purposes. Eur J Echocardiogr 2011;12:339–53.

DOI 10.1002/acr.24041

Erratum

In the article by Birnbaum et al in the July 2019 issue of *Arthritis Care & Research* (pages 936–948) an incorrect NIH grant number was listed in the first sentence of the footnotes. The correct sentence should have been: The Johns Hopkins Rheumatic Diseases Research Core Center, where the assays were performed receives support from the NIH (grant P30-AR-070254).

We regret the error.

echocardiography: consensus document of the EACVI/ASE/Industry Task Force to standardize deformation imaging. Eur Heart J Cardiovasc Imaging 2015;16:1–11.
30. Midtbø H, Semb AG, Matre K, Kvien TK, Gerdts E. Disease activity is associated with reduced left ventricular systolic myocardial function in patients with rheumatoid arthritis. Ann Rheum Dis 2017;76:371–6.
31. Ikonomidis I, Tzortzis S, Lekakis J, Paraskevaidis I, Andreadou I, Nikolaou M, et al. Lowering interleukin-1 activity with anakinra improves myocardial deformation in rheumatoid arthritis. Heart 2009;95:1502–7.
32. Ustun N, Kurt M, Nacar AB, Karateke HP, Guler H, Turhanoglu AD. Left ventricular systolic dysfunction in patients with ankylosing spondylitis without clinically overt cardiovascular disease by speckle tracking echocardiography. Rheumatol Int 2015;35:607–11.
33. Ambale Venkatesh B, Volpe GJ, Donekal S, Mewton N, Liu CY, Shea S, et al. Association of longitudinal changes in left ventricular structure and function with myocardial fibrosis: the Multi-Ethnic Study of Atherosclerosis study. Hypertension 2014;64:508–15.
34. Kramer J, Niemann M, Liu D, Hu K, Machann W, Beer M, et al. Two-dimensional speckle tracking as a non-invasive tool for identification of myocardial fibrosis in Fabry disease. Eur Heart J 2013;34:1587–96.
35. Biesbroek PS, Heslinga SC, Konings TC, van der Horst-Bruinsma IE, Hofman MB, van de Ven PM, et al. Insights into cardiac involvement in ankylosing spondylitis from cardiovascular magnetic resonance. Heart 2017;103:745–52.
36. Ntusi NA, Piechnik SK, Francis JM, Ferreira VM, Matthews PM, Robson MD, et al. Diffuse myocardial fibrosis and inflammation in rheumatoid arthritis: insights from CMR T1 mapping. JACC Cardiovasc Imaging 2015;8:526–36.
37. Kobayashi H, Kobayashi Y, Yokoe I, Akashi Y, Takei M, Giles JT. Magnetic resonance imaging–detected myocardial inflammation and fibrosis in rheumatoid arthritis: associations with disease characteristics and N-terminal pro-brain natriuretic peptide levels. Arthritis Care Res (Hoboken) 2017;69:1304–11.
38. Galderisi M, Henein MY, D'Hooge J, Sicari R, Badano LP, Zamorano JL, et al. Recommendations of the European Association of Echocardiography: how to use echo-Doppler in clinical trials: different modalities for different purposes. Eur J Echocardiogr 2011;12:339–53.

DOI 10.1002/acr.24041

Erratum

In the article by Birnbaum et al in the July 2019 issue of *Arthritis Care & Research* (pages 936–948) an incorrect NIH grant number was listed in the first sentence of the footnotes. The correct sentence should have been: The Johns Hopkins Rheumatic Diseases Research Core Center, where the assays were performed receives support from the NIH (grant P30-AR-070254).

We regret the error.