Arthritis Care & Research

Aims and Scope

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

Volume 77 No. 6 June 2025

Arthritis Care & Research

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

Arthritis Care & Research

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

VOLUME 77 • JUNE 2025 • NO. 6

Cover image: The image on the cover (Kim et al; pages 689–699) shows a model of social determinants of health and risk for systemic lupus erythematosus.

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 689–699
DOI 10.1002/acr.25497

American College of Rheumatology
Empowering Rheumatology Professionals

REVIEW ARTICLE

Environmental Risk Factors for Systemic Lupus Erythematosus Through the Lens of Social Determinants of Health

Youngmin Kim,[1] Jacob J. E. Koopman,[1] May Choi,[2] Candace H. Feldman,[1] and Karen H. Costenbader[1]

Systemic lupus erythematosus (SLE) is a serious multisystem autoimmune disease, marked by alarming sociodemographic inequities. In the United States and around the world, social disadvantage is strongly tied to higher prevalence, more severe disease, and poorer outcomes. A growing list of environmental exposures that contribute to the risk and incidence of SLE have been investigated, and many are now established. However, these environmental exposures—including exposure to air pollution and other contaminants, lifestyle and behavioral factors, and psychologic stress and distress—are not evenly distributed in any population. Individuals of lower socioeconomic status and historically minoritized groups suffer from an imbalanced burden of adverse environmental exposures. In research, clinical practice, and policy making, the strong association of social determinants of health (SDoH) with these exposures has not been given adequate spotlight. In this narrative review, we examine known associations between environmental exposures and SLE risk through the lens of SDoH, laying the foundation for future research and policies to target the environmental risk factors for SLE with awareness of the populations disproportionately affected and the contributing SDoH.

Introduction

Systemic lupus erythematosus (SLE) is a complex chronic autoimmune rheumatic disease with highly heterogeneous manifestations that is challenging to predict, diagnose, and treat. Among the rheumatic diseases, SLE is also noted to have the most striking inequities; lower socioeconomic and historically marginalized racial and ethnic groups have a higher prevalence of and more severe disease, as well as worse outcomes, particularly in the United States.[1] Similar to many chronic diseases, SLE develops over time, stemming from underlying genetic susceptibility interacting with environmental exposures over the life course.[2] Not only are genetic components contributing to the development of SLE increasingly well documented, as more than 150 of such genes have now been identified, but the social and environmental exposures associated with the risk of SLE are increasingly well established, as well.

Several recent reviews have discussed research findings concerning environmental, lifestyle, and behavioral factors associated with the risk of developing SLE.[2–4] Here, we examine some of the more established environmental exposures associated with SLE risk through the lens of social determinants of health (SDoH) (Figure 1). For the purpose of this review, we refer to the nongenetic, thus extrinsic, factors as environmental factors or exposures related to SLE. Emerging evidence points not only to how these exposures are strongly associated with future risk of developing SLE, but also to how they are strongly influenced by socioeconomic status (SES) and historic privilege and SDoH. As such, we also consider which of these environmental exposures might be potentially modifiable or reversible and how we might begin to approach the daunting task of decreasing these exposures among the most highly exposed and affected. For each selected SLE-related risk factor, we first discuss the evidence relating it to the risk of developing SLE, and then assess the evidence pertaining to how, upstream, this exposure is socially determined.

Social and environmental exposures are recognized by policymakers across the world to contribute to inequities in health

Supported by the NIH (grant K24-AR-066109).
[1]Youngmin Kim, MD, Jacob J. E. Koopman, MD, PhD, Candace H. Feldman, MD, MPH, ScD, Karen H. Costenbader, MD, MPH: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [2]May Choi, MD, MPH: University of Calgary, Calgary, Alberta, Canada.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25497.
Address correspondence via email to Karen H. Costenbader, MD, MPH, at kcostenbader@bwh.harvard.edu.
Submitted for publication December 10, 2024; accepted December 11, 2024.

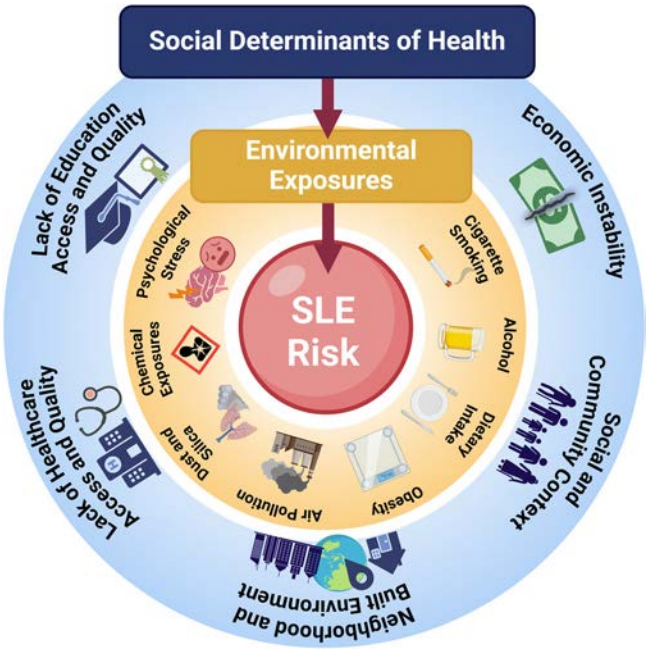

Figure 1. Social determinants of health determine exposure to many environmental factors, which, in turn, affect risk of systemic lupus erythematosus (SLE).

and quality of life. The US Department of Health and Human Services (HHS) has adopted the Healthy People 2030 initiative, encompassing objectives to promote health and well-being and to prevent disease, with proposed interventions and efforts that span physical, mental, and social health dimensions.[5] In particular, the HHS introduced five categories of SDoH: economic stability, education access and quality, health care access and quality, neighborhood and built environment, and social and community context to be specifically addressed and improved.[5] Furthermore, the Centers for Disease Control and Prevention National Center for Chronic Disease Prevention and Health Promotion also shed light on SDoH that may result in poor health, such as poverty, unsafe or unhealthy environments, poor housing, food insecurity, and lack of access to quality education.[6] Here, we consider the following well-studied risk factors for developing SLE and how they are themselves socially determined: exposure to air pollution, toxic compounds and chemicals, and the effects of climate change; lifestyle and behavioral factors, such as smoking, diet, and alcohol intake; and psychologic stress and distress.

Air pollution, climate change, and chemical and airborne exposures related to SLE risk and how they are influenced by SDoH

Air pollution and climate change. Several international studies have investigated associations between exposure to air pollution and risk of incident SLE. A population-based case-control study using the Taiwanese National Health Insurance Research Database reported modest potential effects of estimated long-term residential exposure averaged per year to carbon monoxide (CO), nitrogen dioxide (NO_2), and sulfur dioxide (SO_2) on SLE development.[7] Another Taiwanese study also implicated long-term exposure to NO_2, carbon dioxide (CO_2), and particular matter of 2.5 microns in diameter ($PM_{2.5}$) in increasing the risk of SLE.[8] A study from Chile[9] reported clearly elevated relative risk estimates per interquartile range–increase in single pollutant exposure: 1.3 for SO_2, 1.6 for CO, and 1.4 for $PM_{2.5}$. A recent study of air pollution exposure and risk of SLE among those with and without SLE genetic susceptibility using the UK Biobank reported that participants with both high SLE genetic risk and high air pollution exposure had the highest risk of SLE (vs those with low genetic risk and low air pollution exposure): adjusted hazard ratios (HRs) were 4.2 for $PM_{2.5}$, 5.3 for PM_{10}, 5.6 for NO_2 5.6, and 4.8 for nitrogen oxides. Moreover, a significant multiplicative interaction between genetic risk and NO_2 exposure was found.[10] However, two other studies have not shown significant associations between exposure to air pollution and risk of SLE, perhaps because of high collinearity with other potential risk factors or limitations in the timing of data collection and assessment of SLE onset.[11,12]

Air pollution exposure is well known to affect people disproportionately according to several measures of SES. For example, the distribution of carcinogenic air pollution emissions was recently studied by US Census tract area-level using 2018 regulatory data from the US Environmental Protection Agency, and it was reported that air pollution exposure was 51% higher in populations experiencing poverty or with lower educational status, regardless of race and ethnicity.[13] A separate study of $PM_{2.5}$ air pollution distribution across the United States reported that areas with lower mean household incomes, lower mean educational levels, and higher proportions of Black/African American residents (the last item likely due to structural racism and historic redlining) had higher air $PM_{2.5}$ concentrations.[14] These findings were reinforced in a systematic review in European countries, as well, in which a strong association between area deprivation or lower income and higher levels of particulate matter and nitrogenous oxides in the air was seen.[15] However, the relationship between population SES and air pollution exposure varies according to the industrial and economic state of regions. For instance, a Chinese study of the China Health and Retirement Longitudinal Study (CHARLS) data found higher levels of NO_2 or $PM_{2.5}$ in areas of high SES, where SES was measured with indicators of education, income, and occupation.[16] The authors suggested that given China's recent trend to urbanization and industrialization, regions of development and prosperity were prone to increased exposure to air pollution.

Rapid climate change is exacerbating adverse health exposures to air and water pollution, as well as extreme heat events. In a white paper published in 2024, Dellaripa et al[17] examined the relationship between accelerating climate change and subsequent effects on rheumatic disease manifestations and risk. Several examples of worsening air pollution due to climate change were

provided, including altered wind patterns resulting in a wider influence of high respirable PM. Variations in temperature or precipitation, as well as wildfires, contribute to amplified ozone and PM exposure.[17] The many adverse effects of climate change are disproportionately impacting those of lower SES in society, who lack resources to move, rebuild, and otherwise protect themselves.

Crystalline silica. Exposure to respirable crystalline silica dust is strongly associated with an increased risk of several autoimmune diseases, including SLE, rheumatoid arthritis, scleroderma, and antineutrophil cytoplasmic antibody-associated vasculitis. On inhalation, silica dust triggers an inflammatory response via activation of inflammatory cytokines and increased oxidative stress.[18] Exposure to respirable silica dust can result from mining, sandblasting, quarrying, glassblowing, ceramics, or construction work; therefore, occupational factors significantly contribute to differential exposures to airborne crystalline silica across populations. Several studies have elucidated the association between silica dust exposure and increased SLE risk. SLE risk, up to 10 times that in the general population, was seen from rural agricultural exposures, such as working in the fields in North and South Carolina, as well as in urban settings, such as the Roxbury Lupus Study, in which exposures were through construction work, sandblasting, and specific occupations.[2,18–22] A meta-analysis of four case-control studies and two cohort studies estimated the summary risk associated with silica exposure to be 3.5 times higher and approximately two times higher after excluding patients with silicosis who were presumed to have inhaled massive quantities of silica (in particular miners).[19]

Inequities in the burden of occupational exposures in the United States are well documented. This includes the so-called dusty trades, and lower-paid industrial and manufacturing and factory floor jobs (vs managerial-level positions) often come with higher levels of chronic respirable exposures.[23] Occupations involving manual labor carry higher risks of silica inhalation than do nonmanual work.[24] A study of the National Lung Screening Trial cohort reported that occupational exposure to respirable substances including silica and asbestos was higher among African American/Black workers than White workers.[22] Meanwhile, a study comparing the proportions of workers exposed to crystalline silica across countries revealed much higher levels of exposure in South American nations compared to European countries, suggesting higher rates of air pollution, soil silica, as well as poor ventilation, or hazardous technology use.[25]

Chemical exposures. The Women's Health Initiative Observational Study cohort found that personally applying insecticides in residences or workplaces was associated with increased risk of developing rheumatoid arthritis or SLE (HR 1.6, 95% confidence interval 1.2 to 2.1), with risks depending on the amount of exposure, without much effect of adjustment for race and ethnicity, education, occupation, region, smoking, or obesity.[26] In a study of predominantly African American women, an association between pesticide exposures in urban settings and the risk of SLE was found (odds ratio 2.2).[27] In a study investigating possible ties between regional contamination and increased prevalence of SLE in Arizona, elevated levels of 1,1-dicholoro-2,2-bis-ethylene and organophosphate metabolites were present, along with a higher prevalence of SLE in residents; yet, a statistically significant association was not found.[28] Pesticide exposure is correlated with occupation, and agricultural workers and their families often have high levels of exposure.[29] Pesticide exposures also disproportionately affect historically marginalized populations in the United States.[29–31] Usage of both prohibited and nonprohibited pesticides were reported to be higher in low-income housing than in higher-income housing, as well.[32]

Several heavy metals have also been implicated as potentially related to SLE development. A cross-sectional study reported potentially higher prevalence of rheumatic diseases including SLE, as well as other neurologic, respiratory, and cardiovascular diseases, in populations with greater exposure to mercury.[33] In particular, mercury exposure has been associated with elevated antinuclear antibodies (ANAs) in the blood in past studies.[2,34] Residential proximity to uranium plants and consumption of water contaminated with uranium have been associated with SLE risk, hinting yet another heavy metal that may trigger SLE development.[35,36] Exposure to either metal substance is highly linked with an individual's residential and occupational status. A study on the soil in the Southeastern United States indicated elevated levels of heavy metal contamination in low-income or minority-populated areas, again likely due to historic redlining and environmental racism.[37]

Per- and polyfluoroalkyl substances (PFAS), chemical substances widely used in a variety of commercial and industrial products, have increasingly been given attention because of their ubiquity and long half-life. PFAS uptake via contaminated food and water have been associated with adverse health effects, with changes in systemic inflammation and immune function suggested as possible mechanisms.[38] A case-control study in China observed a relationship between serum PFAS concentrations and SLE risk; study participants with higher concentrations for five different PFAS consistently demonstrated higher odds ratios for SLE.[39] Meanwhile, in a study on Gullah African Americans in South Carolina, PFAS levels were associated with higher Social Vulnerability Index among those with or without SLE, and there was differential exposure to PFAS by social status.[40] The relative paucity of evidence regarding PFAS in relation to SLE risk points to a need for further studies in this area.

Lifestyle and behavioral factors related to SLE and how they are socially determined

Cigarette smoking. Several past studies have reported positive associations between cigarette smoking and increased

risk of developing SLE, especially in the short term, and for the risk of anti–double-stranded DNA (anti-dsDNA)–seropositive SLE. A study of women enrolled in the Black Women's Health Study (BWHS) cohort revealed a higher risk of SLE in current smokers than never smokers,[41] and an association between anti-dsDNA–positive SLE and current smoking was found in the Nurses' Health Study (NHS) cohorts.[42] Although a study from the NHS cohorts exhibited no association between exposure via maternal or paternal smoking and SLE risk among women as adults, a case-control study in Brazil reported a higher risk of childhood-onset SLE with fetal smoking and exposure to passive smoking during pregnancy, and a case-control study in Egypt demonstrated a significant association between passive smoking and SLE risk.[43–45] A meta-analysis of 10 case-control studies and 2 cohort studies found that current smoking associated with a 1.5-fold higher risk than never smoking, confirming similar previous meta-analysis results.[46]

Cigarette smoking is strongly socially determined and associated with several measures of SES around the world. In a cross-sectional study of 353,555 participants from 2011 to 2022 in the National Health Interview Surveys, 22% and 27% smoking prevalence were observed in study participants aged 25 to 39 years and 40 to 64 years, respectively, with education levels of less than high school, whereas this was 4% and 5%, respectively, for the same age groups among study participants with a college education.[47] Additionally, although smoking prevalence has been decreasing in the United States over those years, it remained highest among those who were living at <200% of the Federal Poverty Limit income for all age groups (Figure 2).

Similar relationships between SES and cigarette smoking have been observed in other countries. In the CHARLS study in China, lower educational status was also associated with increased prevalence of smoking.[48] A study on the International Tobacco Control Project Four Country Survey, a longitudinal cohort survey of Australian, Canadian, UK, and US smokers, demonstrated that smokers with lower educational levels or lower income possessed more smoking friends and were more likely to gain more smoking companions.[49] It should also be noted that the prevalence of cigarette smoking has declined in most parts of the world in the past few decades.[50]

Alcohol intake. Several studies have reported that alcohol intake—even at low–moderate levels, such as one half of a drink per day on average, compared to no intake—is associated with a lower risk of SLE.[41,51] These findings have been attributed to potential anti-inflammatory effects of moderate alcohol intake. The two largest cohort studies, conducted in the NHS and the BWHS, found similarly decreased risks of SLE among those consuming alcohol in moderation (NHS: HR 0.6, 95% confidence interval 0.4–0.9; BWHS: HR 0.7, 95% confidence interval 0.5–1.1); women who consumed alcohol also had higher incomes and educational levels in both cohorts than the nondrinkers, and the associations held after adjustment for race, ethnicity, income, and education.[41,52] The associations were strongest for the consumption of wine, but not beer or liquor. A meta-analysis of seven case-control studies and three cohort studies reported a decreased risk of SLE with moderate (one to five drinks per week) alcohol consumption as compared with no consumption.[51]

Alcohol intake is socially determined, and downstream of SDoH, but a complex relationship between alcohol intake and SES in American society has been observed in several studies.[53] Although higher and more frequent alcohol consumption has been found among those of higher SES, as measured by income and educational level, those of lower SES have been found to have higher risks of alcohol-dependency disorders, possibly due to the adverse effects of these disorders on social function and income.[54] According to a study employing data from Gender, Alcohol, and Culture: an International Study, higher educational levels were associated with increased overall drinking, yet lower education levels were associated with risky single-occasion drinking.[55] These relationships may not be similar worldwide, as a systematic review on alcohol consumption in countries of low and lower-middle income reported that the overall prevalence of alcohol abuse was higher in low SES populations in Southeast Asian regions, in which SES was assessed by a variety of measures, including income and education.[56]

Dietary intake. Given the strong potential for recall bias in retrospective studies of diet, it has only been possible to study dietary intake and the risk of later developing SLE in large cohorts of people observed for many years before the onset of disease. Several dietary patterns and indices, including the Alternative Healthy Eating Index, the Mediterranean diet, and the Dietary intervention to Stop Hypertension diet, were examined in relation to SLE risk in the NHS, without strong findings.[57,58] Studies in the NHS cohorts did not find significant associations between intake of neither vitamin D in the diet nor through supplements, either in childhood or adulthood, and later risk of developing SLE.[59,60] However, more recently, high intake of ultraprocessed food (UPF), which includes ready-to-eat and ready-to-heat processed meals, much of which is devoid of nutritional value, in particular sugar-sweetened/artificially sweetened beverages, was associated with increased SLE risk (56% higher risk among those in the highest vs lowest tertile of intake), after adjustment for income, smoking, obesity, and alcohol intake.[61] Furthermore, in the BWHS, intakes of certain fats, including monounsaturated fatty acids, saturated fats, and trans fatty acids, were associated with significantly lower risks of developing SLE, whereas a dietary pattern characterized by high carbohydrates, fruit, and sugar-sweetened drink intake was associated with an 88% increased SLE risk.[62]

Dietary intake (UPF intake in particular), downstream of the SDoH economic instability and food insecurity, is also strongly socially patterned in the US data from the National Health

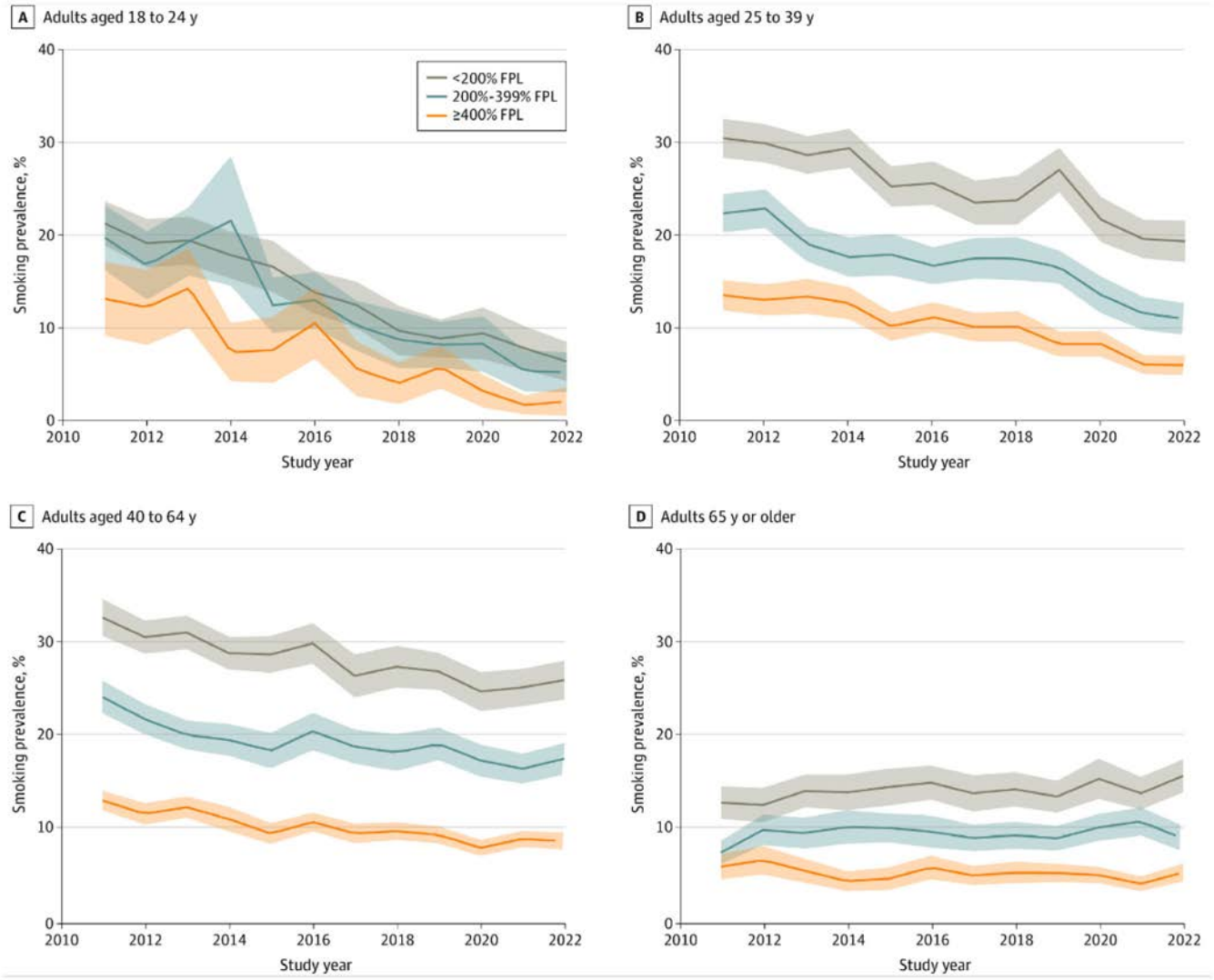

Figure 2. US annual smoking prevalence from 2011 to 2022 by age and family income. Data are from the National Health Interview Surveys. Shading denotes the 95% confidence interval. FPL, Federal Poverty Level.[47] Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25497/abstract.

Assessment Examination Survey (NHANES) show that UPF consumption increased in the entire US population[63] between 2001 and 2018. UPF intake is now, alarmingly, more than 50% of US adult daily caloric intake, with a strong gradient such that those of lower educational levels and lower incomes have the highest intake, representing >60% of daily calories for many (Figure 3). This is likely due to the industrial mass production of UPF that has a long shelf-life, making it an economical and overly available dietary option in the United States, a problem that is spreading worldwide.[64] However, there is still variation in the relationship between SES and UPF intake across countries, as discussed in a recent review.[65] Although lower education level was associated with higher UPF intake in Australia, Canada, and the United States, higher income was associated with higher UPF consumption in Chile and Brazil.[65] Higher SES, defined by occupation, social class, and deprivation indices, was associated with higher UPF consumption in Colombia and Mexico, where the contrary was observed in Australia and the United Kingdom.[65] As the spread of industrially produced UPF throughout the world is now a global problem, these relationships may evolve as poorer nations with fresh food shortages are being shipped UPF of low nutritional value.

Obesity. The relationship between obesity, another socially determined factor, and the risk of developing SLE is not well established, although there have been signals of a positive association in past cohort studies. Among 116,430 women observed in the younger of the two NHS cohorts, Nurses' Health Study II (NHSII) (aged 25–35 years at enrollment), obesity was significantly associated with an 80% higher risk of SLE, whereas no strong association between body mass index over time and SLE risk was observed in the NHS cohort of older women (aged 35–50

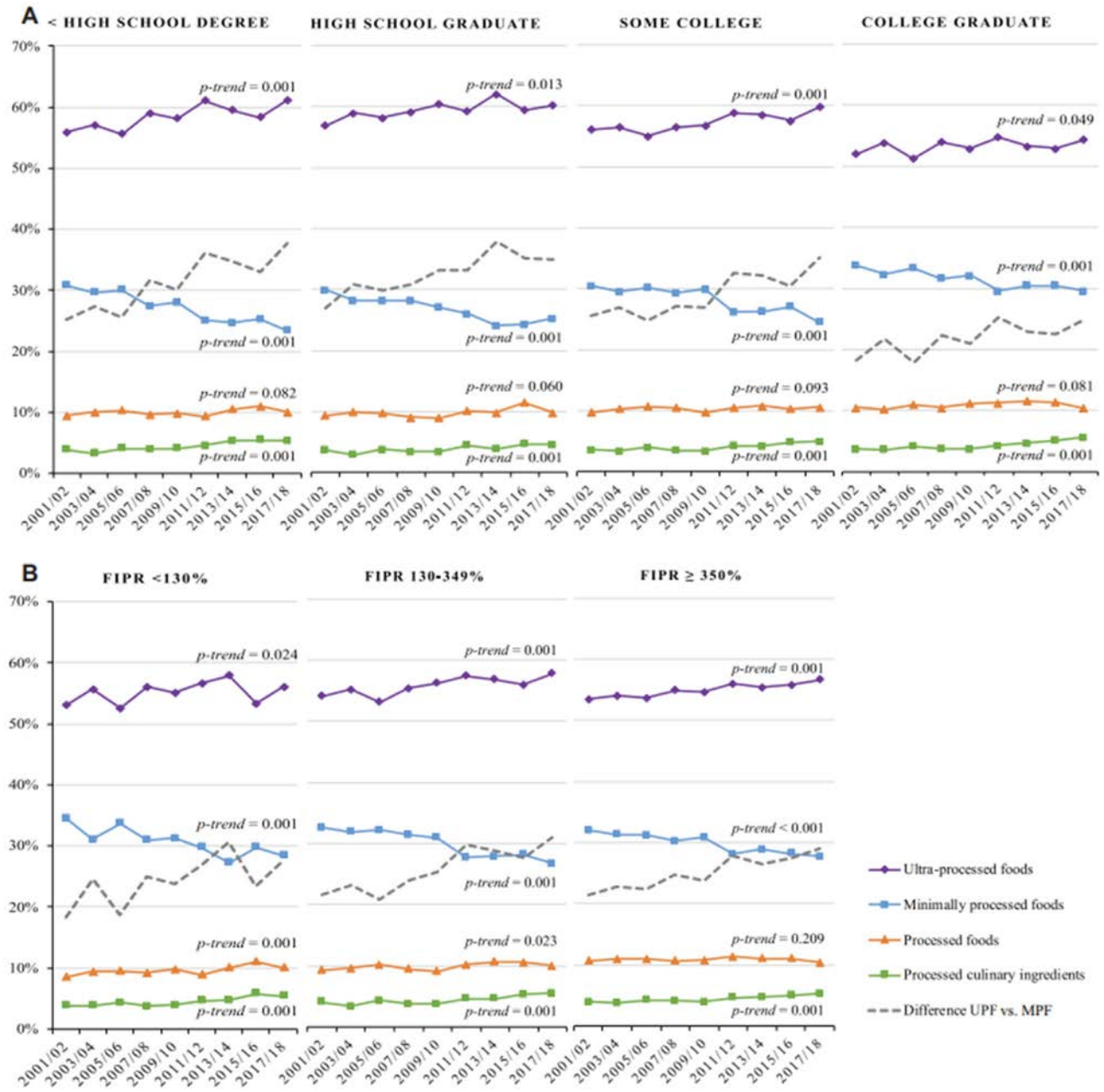

Figure 3. Consumption trends of different food groups stratified by (A) education and (B) income level. Consumption was defined as percentage of total energy intake. Income level was defined as the ratio of family income to poverty.[63] FIPR, family income to poverty ratio; MPF, minimally processed foods; UPF, ultraprocessed foods. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25497/abstract.

years at enrollment).[66] Although no relationship between adult obesity and SLE risk was seen in the BWHS cohort, obesity at age 18 (vs normal body mass index at that age) was strongly associated with more than doubled risk of future SLE.[67]

Obesity is strongly inversely correlated with SES in the United States and higher-income countries and related to multiple SDoH, such as economic instability, neighborhood walkability and safety, access to education and health care, and the social and community context. For example, a systematic review and meta-analysis of 14 studies between 1990 and 2015 demonstrated an association between low SES through life course and obesity especially for women, in which included studied incorporated various measures of SES, such as income, occupational status (manual vs nonmanual), educational level, and housing type.[68] Using data from 11 Organization for Economic Cooperation and Development nations, it was reported that in France and Sweden, obesity was 3.2 and 2.8 times more prevalent in male populations of low than high educational level, respectively, and among both Spanish and Korean women, even larger differences were shown (18 and 17 times, respectively).[69]

The same is not necessarily true in low- and middle-income countries, in which positive correlations between increasing SES measures, including social class, income, and education, and increasing risks of obesity, have been observed.[70]

Psychologic stress and distress as mediators of SDoH

The experience of psychologic stress is a response to perceived challenges, which triggers a series of physiologic and psychologic reactions that can negatively impact the immune system.[71] In the context of SLE, both depression and posttraumatic stress disorder (PTSD)—both forms of psychologic responses to emotional or physical stressors—have been associated with an increased risk of developing SLE. In a study on the NHSII cohort, exposures to past traumatic events and PTSD symptoms were both associated with higher SLE risk.[72] A separate study in this cohort reported an almost threefold higher risk of SLE among the women who self-reported the highest level of previous child abuse, compared to those who reported none, in adjusted models.[73] Furthermore, in a study in the BWHS cohort, childhood abuse, both physical and sexual, was strongly associated with an increase in SLE incidence.[74] In a large case-control study of US Medicaid recipients, the odds ratio for developing SLE among patients with PTSD was still double, after adjustment for race and ethnicity, area-level income, region, smoking, and obesity. When stratified by area-level income, the odds ratios were 1.5 for those in the lower income areas and 2.2 for those living in higher-income areas.[75] The Women's Health Initiative Observational Study cohort also reported having three or more stressful life events in the past year was associated with a 70% increased risk for the development of rheumatoid arthritis or SLE in the subsequent three years.[76] Similar results were found in other countries, such as Taiwan, where history of PTSD was associated with higher risk of developing several autoimmune diseases, including SLE.[77] Depression is also a risk factor for SLE. In a study on the NHS cohorts, depression was associated with elevated risk (HR 2.5) of later developing SLE, after adjustment for body mass index, smoking, oral contraception, and postmenopausal hormone use.[78] Three different depression indicators (clinical diagnosis, antidepressant usage, and depressed mood indicated by Mental Health Inventory-5 scores) yielded associations with increased SLE risk when analyzed separately, as well.

SDoH relates to the severity and attributes of psychologic responses to stressful events. A study of patients who had experienced trauma in Georgia revealed a strong relationship between neighborhood poverty and the severity of subsequent PTSD, whereas a Survey of the Health of Wisconsin study found an association between low SES, as measured by education and income, and higher risk of depression.[79,80] International studies have also revealed similar associations; an Australian study showed that family poverty was a significant predictor of adolescent and young adult depression and anxiety, whereas a prospective study of a German cohort proposed that stressful life events were less likely to yield mental health problems among children whose parents had higher educational attainment.[81,82] A Chinese cohort study revealed that study participants with lower SES, as measured by family income, employment status, and education, exhibited greater risk of depression.[83] The relationship between SDoH and PTSD symptoms in response to the COVID-19 outbreak extensively reported, as the pandemic was a major global source of stress. In one study on the population of Hubei Province, China, where the outbreak initiated, lower perceived social status, in terms of income, education, and employment, was associated with 30% increase in PTSD symptoms, measured by the Impact of Event Scale-Revised.[84] Another study that used an online survey on study participants from 152 countries demonstrated that individuals with no or only primary education exhibited increased higher odds of developing PTSD symptoms.[85]

Discussion

As we increasingly recognize and combat food unsafety, air and water pollution, climate change, global epidemics, and unprecedented levels of mental health stress, it is essential to realize that these threats to human health disproportionately affect lower socioeconomic and historically minoritized groups in our society. These are main causes of observed disparities in disease incidence and outcomes by many measures of SES for many health conditions, including complex autoimmune conditions such as SLE. By some measures, the incidence of autoimmunity in the United States and worldwide may be increasing.[86] In the NHANES national survey data, the prevalence of positive ANAs, often a sign of an initial break in immune tolerance before the onset of many autoimmune diseases, has increased[87] from 11% in 1988 to 1991 to 16% in 2011 to 2012. However, it is not entirely clear that the incidence of SLE is increasing, and this may be due to the rise in many exposures, such as air pollution, diet changes, and stress, whereas others, such as cigarette smoking and alcohol intake, have been declining. There have also been temporal changes in the use and accuracy of autoantibody testing, aging of the population and unstable population denominators due to immigration, and wide variation in the timeliness and accuracy of diagnosis that have contributed to difficulty assessing SLE incidence over time.

The mechanistic biologic pathways between SDoH exposures, the risk factors we have enumerated, and the development of SLE are not all clear and beyond the scope of this review. Systemic inflammation, hypothalamic pituitary axis dysregulation, epigenetic changes, oxidative stress, and priming of innate immune pathways leading to production of interferon, a key proinflammatory cytokine in SLE, are all possible biologic mechanisms.[71] These mechanisms may be shared, or unique to certain

exposures, and may be interactive and synergistic with those of other exposures and genetic susceptibility genes.

We have discussed several lifestyle, behavioral, and psychologic risk factors for SLE, including dietary and alcohol intake, cigarette smoking, obesity, depression, and PTSD, as environmental exposures, to underscore the great extent to which these are social contextually, not genetically, determined, but we acknowledge that each of these results from a complex and nuanced interaction between genetics and the environment, including SDoH. Additionally, our discussion of environmental risk factors and their relationship to SDoH does not explain the female predominance of SLE. We posit that being born with two X chromosomes puts an individual at an increased risk of SLE and increased susceptibility to the numerous environmental exposures that may accelerate or trigger the development of SLE. Men without a second X chromosome, in particular, may be more resistant to the effects of the above exposures in terms of their risk of SLE, in the absence of a high non-X chromosome genetic susceptibility burden. Women may also be more likely to have some of these SLE-related exposures, such as depression and PTSD, in addition to potentially being more susceptible to them. Furthermore, throughout this study we have used the term "environmental" to collectively define any factor that is extrinsic yet nongenetic, and downstream to SDoH; yet, in some research contexts, such a classification may oversimplify more nuanced interactions between nongenetically predetermined factors. Therefore, depending on the scope and focus of the research, future studies may benefit from different levels of flexibility in defining such terms.

We have focused our discussion primarily on the SDoH and have not concentrated on their strong historic relationships with the social constructs of race and ethnicity in the United States. The relationship between the two is complex, particularly in relation to environmental exposures and disease development, differs by exposure, and is by no means constant over time and region, both within the United States and around the world. In the United States, historically, however, those of minoritized groups have been subjugated to lower SES by every measure and experienced a hugely disproportionate burden of many of these adverse SDoH. Experiencing racism or discrimination is an ongoing source of stress in many people's lives. Although it has been shown to be related to worse outcomes among those living with SLE, it has not been well studied in relationship to the development of SLE to date.[88] Similarly, the experience of enduring sexism may be a source of stress and even trauma that has not been studied in relation to SLE risk per se, but that is likely also tied to or exacerbated by adverse SDoH.

Conclusions

Having recognized that many of the growing number of environmental exposures linked to SLE risk are strongly socially determined, potentially explaining a large part of the sociodemographic inequities observed in this complex chronic disease, what do we need to do? Many groups are beginning to address these large and seemingly overwhelming problems.[17] In an effort to reduce the burden of disease and work toward health equity, large multimodal interventions to educate and combat poverty, stress, climate change, and unhealthy lifestyle are urgently needed for those experiencing the often synergistic effects of these adverse SDoH. Moreover, they would benefit those at risk for SLE and for other health conditions, as well.[89] We are now working to assemble a nationwide cohort of young women with family histories of SLE, to identify those with highest exposure to these SLE-related environmental risk factors, with the goal of starting to educate, improve access to health care, and mitigate other exposures that are linked to SDoH. Just as it is rapidly becoming possible to identify those at highest risk for SLE using our knowledge of genetic and environmental risk factors, interventions to address the many socially determined factors that increase this risk may also become a possibility.[3,90]

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Costenbader confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Pryor KP, Barbhaiya M, Costenbader KH, et al. Disparities in Lupus and lupus nephritis care and outcomes among US Medicaid beneficiaries. Rheum Dis Clin North Am 2021;47(1):41–53.
2. Woo JMP, Parks CG, Jacobsen S, et al. The role of environmental exposures and gene–environment interactions in the etiology of systemic lupus erythematous. J Intern Med 2022;291(6):755–778.
3. Choi MY, Hahn J, Malspeis S, et al. Association of a combination of healthy lifestyle behaviors with reduced risk of incident systemic lupus erythematosus. Arthritis Rheumatol 2022;74(2):274–283.
4. Choi MY, Costenbader KH, Fritzler MJ. Environment and systemic autoimmune rheumatic diseases: an overview and future directions. Front Immunol 2024;15.
5. Office of Disease Prevention and Health Promotion. Healthy People 2030. Accessed October 9, 2024. https://health.gov/healthypeople
6. Centers for Disease Control and Prevention. Addressing Social Determinants of Health and Chronic Diseases. 2024. Accessed November 7, 2024. https://www.cdc.gov/health-equity-chronic-disease/social-determinants-of-health-and-chronic-disease/index.html#:~:text=Key%20points%201%20Social%20determinants%20of%20health%20%28SDOH%29,works%20to%20achieve%20health%20equity%20by%20addressing%20SDOH.
7. Tang KT, Tsuang BJ, Ku KC, et al. Relationship between exposure to air pollutants and development of systemic autoimmune rheumatic

diseases: a nationwide population-based case-control study. Ann Rheum Dis 2019;78(9):1288–1291.
8. Jung CR, Chung WT, Chen WT, et al. Long-term exposure to traffic-related air pollution and systemic lupus erythematosus in Taiwan: a cohort study. Sci Total Environ 2019;668:342–349.
9. Cakmak S, Blanco-Vidal C, Lukina AO, et al. The association between air pollution and hospitalization for patients with systemic lupus erythematosus in Chile: a daily time series analysis. Environ Res 2021; 192:110469.
10. Xing M, Ma Y, Cui F, et al. Air pollution, genetic susceptibility, and risk of incident systemic lupus erythematosus: a prospective cohort study. Arthritis Rheumatol 2024;76(10):1530–1537.
11. Adami G, Pontalti M, Cattani G, et al. Association between long-term exposure to air pollution and immune-mediated diseases: a population-based cohort study. RMD Open 2022;8(1):e002055.
12. Park JS, Choi S, Kim K, et al. Association of particulate matter with autoimmune rheumatic diseases among adults in South Korea. Rheumatology (Oxford) 2021;60(11):5117–5126.
13. Madrigal JM, Flory A, Fisher JA, et al. Sociodemographic inequities in the burden of carcinogenic industrial air emissions in the United States. J Natl Cancer Inst 2024;116(5):737–744.
14. Boing AF, deSouza P, Boing AC, et al. Air Pollution, socioeconomic status, and age-specific mortality risk in the United States. JAMA Netw Open 2022;5(5):e2213540.
15. Fairburn J, Schüle SA, Dreger S, et al. Social inequalities in exposure to ambient air pollution: a systematic review in the WHO European region. Int J Environ Res Public Health 2019;16(17):3127.
16. Wang Y, Wang Y, Xu H, et al. Ambient air pollution and socioeconomic status in China. Environ Health Perspect 2022;130(6):67001.
17. Dellaripa PF, Sung LH, Bain PA, et al; American College of Rheumatology Committee on Research. American College of Rheumatology White Paper: the effects of climate change on rheumatic conditions-an evolving landscape and a path forward. Arthritis Rheumatol 2024; 76(10):1459–1466.
18. Pollard KM. Silica, silicosis, and autoimmunity. Front Immunol 2016;7.
19. Morotti A, Sollaku I, Catalani S, et al. Systematic review and meta-analysis of epidemiological studies on the association of occupational exposure to free crystalline silica and systemic lupus erythematosus. Rheumatology (Oxford) 2021;60(1):81–91.
20. Finckh A, Cooper GS, Chibnik LB, et al. Occupational silica and solvent exposures and risk of systemic lupus erythematosus in urban women. Arthritis Rheum 2006;54(11):3648–3654.
21. Cooper GS, Wither J, Bernatsky S, et al; CaNIOS GenES Investigators. Occupational and environmental exposures and risk of systemic lupus erythematosus: silica, sunlight, solvents. Rheumatology (Oxford) 2010;49(11):2172–2180.
22. Juon H-S, Hong A, Pimpinelli M, et al. Racial disparities in occupational risks and lung cancer incidence: analysis of the National Lung Screening Trial. Prev Med 2021;143:106355.
23. Quinn MM, Sembajwe G, Stoddard AM, et al. Social disparities in the burden of occupational exposures: results of a cross-sectional study. Am J Ind Med 2007;50(12):861–875.
24. Fluck B, Mahdjoubi L, Fluck D, et al. Disparity in the risk of exposure to respirable crystalline silica dust among non-manual and manual employees in the construction industry. Saf Extrem Environ 2021; 3(2):125–132.
25. Ribeiro F, Bernales B, Alcaino J, et al. P1-303 inequalities in silica exposure: a study using job exposure matrix. J Epidemiol Community Health 2011;65(suppl 1):A150–A151.
26. Parks CG, Walitt BT, Pettinger M, et al. Insecticide use and risk of rheumatoid arthritis and systemic lupus erythematosus in the Women's Health Initiative Observational Study. Arthritis Care Res (Hoboken) 2011;63(2):184–194.
27. Williams JN, Chang SC, Sinnette C, et al. Pesticide exposure and risk of systemic lupus erythematosus in an urban population of predominantly African-American women. Lupus 2018;27(13):2129–2134.
28. Balluz L, Philen R, Ortega L, et al. Investigation of systemic lupus erythematosus in Nogales, Arizona. Am J Epidemiol 2001;154(11): 1029–1036.
29. Donley N, Bullard RD, Economos J, et al. Pesticides and environmental injustice in the USA: root causes, current regulatory reinforcement and a path forward. BMC Public Health 2022;22(1):708.
30. Howarth MV, Eiser AR. Environmentally mediated health disparities. Am J Med 2023;136(6):518–522.
31. Temkin AM, Uche UI, Evans S, et al. Racial and social disparities in Ventura County, California related to agricultural pesticide applications and toxicity. Sci Total Environ 2022;853:158399.
32. Vaezafshar S, Siegel JA, Jantunen L, et al. Widespread occurrence of pesticides in low-income housing. J Expo Sci Environ Epidemiol 2024;34(4):735–744.
33. Dahlgren J, Takhar H, Anderson-Mahoney P, et al. Cluster of systemic lupus erythematosus (SLE) associated with an oil field waste site: a cross sectional study. Environ Health 2007;6(1):8.
34. Gardner RM, Nyland JF, Silbergeld EK. Differential immunotoxic effects of inorganic and organic mercury species in vitro. Toxicol Lett 2010;198(2):182–190.
35. Lu-Fritts PY, Kottyan LC, James JA, et al. Association of systemic lupus erythematosus with uranium exposure in a community living near a uranium-processing plant: a nested case-control study. Arthritis Rheumatol 2014;66(11):3105–3112.
36. Erdei E, Shuey C, Pacheco B, et al. Elevated autoimmunity in residents living near abandoned uranium mine sites on the Navajo Nation. J Autoimmun 2019;99:15–23.
37. Jones DH, Yu X, Guo Q, et al. Racial disparities in the heavy metal contamination of urban soil in the Southeastern United States. Int J Environ Res Public Health 2022;19(3):1105.
38. Barton KE, Zell-Baran LM, DeWitt JC, et al. Cross-sectional associations between serum PFASs and inflammatory biomarkers in a population exposed to AFFF-contaminated drinking water. Int J Hyg Environ Health 2022;240:113905.
39. He Y, Qu C, Tian J, et al. Association of perfluoroalkyl and polyfluoroalkyl substances (PFASs) exposures and the risk of systemic lupus erythematosus: a case–control study in China. Environm Health 2023;22(1):78.
40. Vara EL, Wilson DA, Pearce JL, et al. 1103 perfluoroalkyl substances and community vulnerability: associations with lupus-related autoantibodies and disease. Lupus Sci Med 2021;8.
41. Cozier YC, Barbhaiya M, Castro-Webb N, et al. Relationship of cigarette smoking and alcohol consumption to incidence of systemic lupus erythematosus in a prospective cohort study of black women. Arthritis Care Res (Hoboken) 2019;71(5):671–677.
42. Barbhaiya M, Tedeschi SK, Lu B, et al. Cigarette smoking and the risk of systemic lupus erythematosus, overall and by anti-double stranded DNA antibody subtype, in the Nurses' Health Study cohorts. Ann Rheum Dis 2018;77(2):196–202.
43. Simard J, Costenbader K, Liang M, et al. Exposure to maternal smoking and incident SLE in a prospective cohort study. Lupus 2009;18(5): 431–435.
44. Conde PG, Farhat LC, Braga ALF, et al. Are prematurity and environmental factors determinants for developing childhood-onset systemic lupus erythematosus? Mod Rheumatol 2018;28(1):156–160.
45. Refai RH, Hussein MF, Abdou MH, et al. Environmental risk factors of systemic lupus erythematosus: a case–control study. Sci Rep 2023; 13(1):10219.

46. Costenbader KH, Kim DJ, Peerzada J, et al. Cigarette smoking and the risk of systemic lupus erythematosus: a meta-analysis. Arthritis Rheum 2004;50(3):849–857.
47. Meza R, Cao P, Jeon J, et al. Trends in US adult smoking prevalence, 2011 to 2022. JAMA Health Forum 2023;4(12):e234213.
48. Wang Q, Shen JJ, Sotero M, et al. Income, occupation and education: are they related to smoking behaviors in China? PLoS One 2018;13(2):e0192571.
49. Hitchman SC, Fong GT, Zanna MP, et al. Socioeconomic status and smokers' number of smoking friends: findings from the International Tobacco Control (ITC) four country survey. Drug Alcohol Depend 2014;143:158–166.
50. World Health Organization. WHO global report on trends in prevalence of tobacco use 2000–2030. 2024. https://www.who.int/publications/i/item/9789240088283
51. Wang J, Liu J, Pan L, et al. Association between alcohol intake and the risk of systemic lupus erythematosus: a systematic review and meta-analysis. Lupus 2021;30(5):725–733.
52. Barbhaiya M, Lu B, Sparks JA, et al. Influence of alcohol consumption on the risk of systemic lupus erythematosus among women in the Nurses' Health Study cohorts. Arthritis Care Res (Hoboken) 2017;69(3):384–392.
53. Probst C, Kilian C, Sanchez S, et al. The role of alcohol use and drinking patterns in socioeconomic inequalities in mortality: a systematic review. Lancet Public Health 2020;5(6):e324–e332.
54. Collins SE. Associations between socioeconomic factors and alcohol outcomes. Alcohol Res 2016;38(1):83–94.
55. Grittner U, Kuntsche S, Gmel G, et al. Alcohol consumption and social inequality at the individual and country levels--results from an international study. Eur J Public Health 2013;23(2):332–339.
56. Allen LN, Townsend N, Williams J, et al. Socioeconomic status and alcohol use in low- and lower-middle income countries: a systematic review. Alcohol 2018;70:23–31.
57. Barbhaiya M, Tedeschi S, Sparks JA, et al. Association of dietary quality with risk of incident systemic lupus erythematosus in the Nurses' Health Study and Nurses' Health Study II. Arthritis Care Res (Hoboken) 2021;73(9):1250–1258.
58. Tedeschi SK, Barbhaiya M, Sparks JA, et al. Dietary patterns and risk of systemic lupus erythematosus in women. Lupus 2020;29(1):67–73.
59. Costenbader KH, Feskanich D, Holmes M, et al. Vitamin D intake and risks of systemic lupus erythematosus and rheumatoid arthritis in women. Ann Rheum Dis 2008;67(4):530–535.
60. Hiraki LT, Munger KL, Costenbader KH, et al. Dietary intake of vitamin D during adolescence and risk of adult-onset systemic lupus erythematosus and rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64(12):1829–1836.
61. Rossato S, Oakes EG, Barbhaiya M, et al. Ultraprocessed food intake and risk of systemic lupus erythematosus among women observed in the Nurses' Health Study cohorts. Arthritis Care Res (Hoboken) 2024;77(1):50–60.
62. Castro-Webb N, Cozier YC, Barbhaiya M, et al. Association of macronutrients and dietary patterns with risk of systemic lupus erythematosus in the Black Women's Health Study. Am J Clin Nutr 2021;114(4):1486–1494.
63. Juul F, Parekh N, Martinez-Steele E, et al. Ultra-processed food consumption among US adults from 2001 to 2018. Am J Clin Nutr 2022;115(1):211–221.
64. Dicken SJ, Batterham RL. Ultra-processed food: a global problem requiring a global solution. Lancet Diabetes Endocrinol 2022;10(10):691–694.
65. Dicken SJ, Qamar S, Batterham RL. Who consumes ultra-processed food? A systematic review of sociodemographic determinants of ultra-processed food consumption from nationally representative samples. Nutr Res Rev 2024;37(2):416–456.
66. Tedeschi SK, Barbhaiya M, Malspeis S, et al. Obesity and the risk of systemic lupus erythematosus among women in the Nurses' Health Studies. Semin Arthritis Rheum 2017;47(3):376–383.
67. Cozier YC, Barbhaiya M, Castro-Webb N, et al. A prospective study of obesity and risk of systemic lupus erythematosus (SLE) among black women. Semin Arthritis Rheum 2019;48(6):1030–1034.
68. Newton S, Braithwaite D, Akinyemiju TF. Socio-economic status over the life course and obesity: systematic review and meta-analysis. PLoS One 2017;12(5):e0177151.
69. Devaux M, Sassi F. Social inequalities in obesity and overweight in 11 OECD countries. Eur J Public Health 2013;23(3):464–469.
70. Dinsa GD, Goryakin Y, Fumagalli E, et al. Obesity and socioeconomic status in developing countries: a systematic review. Obes Rev 2012;13(11):1067–1079.
71. Goldschen L, Ellrodt J, Amonoo HL, et al. The link between posttraumatic stress disorder and systemic lupus erythematosus. Brain Behav Immun 2023;108:292–301.
72. Roberts AL, Malspeis S, Kubzansky LD, et al. Association of trauma and posttraumatic stress disorder with incident systemic lupus erythematosus in a longitudinal cohort of women. Arthritis Rheumatol 2017;69(11):2162–2169.
73. Feldman CH, Malspeis S, Leatherwood C, et al. Association of childhood abuse with incident systemic lupus erythematosus in adulthood in a longitudinal cohort of women. J Rheumatol 2019;46(12):1589–1596.
74. Cozier YC, Barbhaiya M, Castro-Webb N, et al. Association of child abuse and systemic lupus erythematosus in black women during adulthood. Arthritis Care Res (Hoboken) 2021;73(6):833–840.
75. Case SM, Feldman CH, Guan H, et al. Posttraumatic stress disorder and risk of systemic lupus erythematosus among Medicaid recipients. Arthritis Care Res (Hoboken) 2023;75(1):174–179.
76. Parks CG, Pettinger M, de Roos AJ, et al. Life events, caregiving, and risk of autoimmune rheumatic diseases in the Women's Health Initiative observational study. Arthritis Care Res (Hoboken) 2023;75(12):2519–2528.
77. Hsu TW, Bai YM, Tsai SJ, et al. Risk of autoimmune diseases after post-traumatic stress disorder: a nationwide cohort study. Eur Arch Psychiatry Clin Neurosci 2024;274(3):487–495.
78. Roberts AL, Kubzansky LD, Malspeis S, et al. Association of depression with risk of incident systemic lupus erythematosus in women assessed across 2 decades. JAMA Psychiatry 2018;75(12):1225–1233.
79. Ravi M, Powers A, Rothbaum BO, et al. Neighborhood poverty prospectively predicts PTSD symptoms six-months following trauma exposure. Ment Health Sci 2023;1(4):213–221.
80. Park EY, Oliver TR, Peppard PE, et al. Sense of community and mental health: a cross-sectional analysis from a household survey in Wisconsin. Fam Med Community Health 2023;11(2):e001971.
81. Najman JM, Hayatbakhsh MR, Clavarino A, et al. Family poverty over the early life course and recurrent adolescent and young adult anxiety and depression: a longitudinal study. Am J Public Health 2010;100(9):1719–1723.
82. Reiss F, Meyrose A-K, Otto C, et al. Socioeconomic status, stressful life situations and mental health problems in children and adolescents: results of the German BELLA cohort-study. PLoS One 2019;14(3):e0213700.
83. Wang Y, Liu M, Yang F, et al. The associations of socioeconomic status, social activities, and loneliness with depressive symptoms in adults aged 50 years and older across 24 countries: findings from five prospective cohort studies. Lancet Healthy Longev 2024;5(9):100618.

84. Wang Y, Xu S, Chen Y, et al. A decline in perceived social status leads to post-traumatic stress disorder symptoms in adults half a year after the outbreak of the COVID-19 pandemic: consideration of the mediation effect of perceived vulnerability to disease. Fronti Psychiatry 2023;14.

85. Folayan MO, Zuñiga RAA, Ellakany P, et al. Socio-economic factors associated with post-traumatic stress symptoms among adolescents and young people during the first wave of the COVID-19 pandemic. Sci Rep 2024;14(1):2276.

86. Miller FW. The increasing prevalence of autoimmunity and autoimmune diseases: an urgent call to action for improved understanding, diagnosis, treatment, and prevention. Curr Opin Immunol 2023;80: 102266.

87. Dinse GE, Parks CG, Weinberg CR, et al. Increasing prevalence of antinuclear antibodies in the United States. Arthritis Rheumatol 2022;74(12):2032–2041.

88. Williams JN, Drenkard C, Lim SS. The impact of social determinants of health on the presentation, management and outcomes of systemic lupus erythematosus. Rheumatology (Oxford) 2023;62(Suppl 1):i10–i14.

89. Buie J, McMillan E, Kirby J, et al. Disparities in lupus and the role of social determinants of health: current state of knowledge and directions for future research. ACR Open Rheumatol 2023;5(9):454–464.

90. Cui J, Malspeis S, Choi MY, et al. Risk prediction models for incident systemic lupus erythematosus among women in the Nurses' health study cohorts using genetics, family history, and lifestyle and environmental factors. Semin Arthritis Rheum 2023;58:152143.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 700–709
DOI 10.1002/acr.25479

Measuring Frailty in Systemic Lupus Erythematosus

Patricia Katz,[1] Maria Dall'Era,[1] Laura Plantinga,[1] Kamil E. Barbour,[2] Kurt J. Greenlund,[2] and Jinoos Yazdany[1]

Objective. Recent research has explored frailty in systemic lupus erythematosus using multiple measures. We examined the agreement among frailty measures and the association of each with cross-sectional and longitudinal health outcomes.

Methods. We used data from the California Lupus Epidemiology Study to examine the following measures of frailty: Systemic Lupus International Collaborating Clinics (SLICC) Frailty Index (SLICC-FI), Short Physical Performance Battery (SPPB), and Fatigue, Resistance, Ambulation, Illness, and Loss of Weight (FRAIL) scale questionnaire. Patient-Reported Outcomes Measurement Information System Physical Function 10a (PF) was tested as a proxy measure of frailty. Agreement between frailty classifications by each measure was assessed. Cross-sectional associations of frailty classifications with hospitalization, valued life activities disability, cognitive impairment, six-minute walk test distance, self-reported disease damage, fatigue, and depressive symptoms were assessed with logistic and linear regression analyses. Associations with hospitalization, disease damage increase, and disability increase over the subsequent three years were assessed by Cox proportional hazards analyses.

Results. Percentages of participants identified as frail varied among the measures, from 10.8% to 45.9%. Agreement among classifications ranged from slight to substantial (k from 0.17 to 0.63). Most of the frailty measures were associated with both cross-sectional and longitudinal health outcomes, with the notable exception of the SPPB. SLICC-FI had the most consistent association with outcomes, followed by FRAIL and PF.

Conclusion. Multiple measures of frailty appear to identify the risk of poor health outcomes. The intended use, as well as the simplicity and practicality of implementing the measure, may be the most important considerations in choosing a frailty measure.

INTRODUCTION

Frailty has been conceptualized as an accumulation of deficits across multiple physiologic systems. These deficits result in a reduction of the body's physiologic reserves and a generalized vulnerability to stressors, which makes individuals more susceptible to poor health outcomes.[1,2] Two major conceptual approaches to measuring frailty have been developed. One, the phenotypic approach, focuses on physical frailty and uses the presence of specific criteria to define frailty (ie, deficits in specific areas). The most common of these is the Fried Frailty Phenotype (FFP), for which five criteria are evaluated: (1) low weight or weight loss, (2) weakness (grip strength), (3) slowness (gait speed), (4) exhaustion, (self-reported fatigue), and (5) inactivity (self-reported).[1] Individuals must meet three of the five criteria to be considered frail. Individuals meeting one or two criteria are considered pre-frail. The Short Physical Performance Battery (SPPB), developed by Guralnik et al for use in geriatric settings, has also been used to define frailty.[3] The other deficit-accumulation approach conceptualizes frailty more broadly and incorporates elements of physical frailty, as well as comorbid conditions, laboratory measurements, and social factors. The most common of these may be the Rockwood Frailty Index, which uses up to 71 items to construct a frailty score.[4] The brief Fatigue, Resistance, Ambulation, Illness, and Loss of Weight (FRAIL) scale questionnaire, a self-report measure incorporating elements of both approaches, has also been introduced.[5]

The California Lupus Epidemiology Study (CLUES) was supported by the Centers for Disease Control and Prevention (CDC) of the US Department of Health and Human Services (HHS) (grant U01-DP06701) as part of a financial assistance award totaling $4,400,000, with 100% funded by CDC/HHS.

[1]Patricia Katz, PhD, Maria Dall'Era, MD, Laura Plantinga, PhD, Jinoos Yazdany, MD, MPH: University of California San Francisco, San Francisco, California; [2]Kamil E. Barbour, PhD, MS, MPH, Kurt J. Greenlund, PhD: Centers for Disease Control and Prevention, Atlanta, Georgia.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25479).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25479.

Address correspondence via email to Patricia Katz, PhD, at Patti.katz@ucsf.edu.

Submitted for publication June 20, 2024; accepted in revised form November 12, 2024.

SIGNIFICANCE & INNOVATIONS

- Frailty has recently begun to be studied as a risk factor for poor health outcomes in systemic lupus erythematosus. Multiple measures of frailty have been used.
- We found that the percentages of individuals classified as frail varies widely depending on the measure used. Agreement of frailty classifications varied considerably.
- Although most of the frailty measures were predictive of poor health outcomes, some may be more accessible to use. The choice of measures may depend on the purpose of measurement and ease of use.

Frailty was introduced as a relevant concept in systemic lupus erythematosus (SLE) in 2017.[6] Using the FFP, 20% of a sample of women with SLE were classified as frail and 50% as pre-frail. Frail women had worse physical functioning, were more likely to have cognitive impairment, and were more likely to experience declines in functioning and the onset of cognitive impairment. Odds of death for frail women were also elevated (adjusted odds ratio [OR] adjusting for age, lupus duration, and baseline disease damage 5.9 [0.6–57.1; mean ± SD follow-up time 7.2 ± 1.1 years]). In that SLE sample, with a mean age of 48, the rate of frailty was twice as high as that seen in studies of adults more than two decades older. A similar prevalence of frailty among women with SLE of similar age was seen in a study by Lieber et al,[7] also using the FFP. The study by Lieber et al incorporated the FRAIL scale as well. By the FRAIL scale, 27% of the sample individuals were classified as frail.

In 2020, the Systemic Lupus International Collaborating Clinics (SLICC) group published an SLE-specific Frailty Index (SLICC-FI) based on the Rockwood-type model of frailty.[8] The SLICC-FI has 48 items, the majority of which are derived from elements of the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI), the SLICC Damage Index (SDI), and the Short Form-36 Physical Function subscale, in addition to specific comorbid conditions. The SLICC-FI has been shown to predict damage accrual, hospitalization, and death and to be associated with physical performance measures.[9–13] In spite of its good psychometric performance, the SLICC-FI can be cumbersome to calculate. Only one study has compared the SLICC-FI with other measures of frailty, showing moderate agreement (kappa = 0.41) between frailty defined by the SLICC-FI and by the FFP.[14] It is not known how the SLICC-FI corresponds with other measures of frailty.

We aimed to examine the correspondence among measures of frailty in SLE and the association of each measure of frailty with cross-sectional and longitudinal health outcomes. Because of the strong emphasis on physical functioning in the FFP and SPPB and the inclusion of physical functioning in the SLICC-FI and FRAIL scale, we also chose to examine a measure of physical functioning as a potential proxy measure of frailty that might be more readily available in research and clinical settings.

METHODS

Study sample. Subjects were participants in the California Lupus Epidemiology Study (CLUES), a multiracial/ethnic cohort of individuals with physician-confirmed SLE. Some participants (n = 171) were recruited from the California Lupus Surveillance Project (CLSP), a population-based cohort of individuals with SLE living in San Francisco County from 2007 to 2009.[15] Additional participants (n = 260) residing in the nine counties in the San Francisco Bay Area geographic region were recruited through local academic and community rheumatology clinics and through existing local research cohorts. No substantive differences existed between the two recruitment groups in distribution of sociodemographic or clinical characteristics. In addition to residence in the San Francisco Bay Area, other inclusion criteria were a confirmed SLE diagnosis; oral language fluency in English, Spanish, Cantonese, or Mandarin; age ≥18 years; and ability to provide informed consent.

Baseline study procedures involved an in-person research clinic visit, which included collection and review of medical records before the visit; a history and physical examination conducted by a physician specializing in SLE; collection of biospecimens for clinical and research purposes; and completion of a structured interview administered by an experienced research assistant. All SLE diagnoses were confirmed by study physicians according to any of the following definitions: (a) the patient met ≥4 of the 11 American College of Rheumatology (ACR) revised criteria for the classification SLE as defined in 1982 and updated in 1997,[16,17] (b) the patient met ≥3 of the 11 ACR criteria plus a documented rheumatologist's diagnosis of SLE, or (c) the patient had a confirmed diagnosis of lupus nephritis. These case definitions were used in SLE surveillance studies supported by the Centers for Disease Control and Prevention, with the recognition that all historical records may not have been accessible when surveillance activities were undertaken.[15,18–20]

CLUES specifically aimed to include a diverse patient sample, with representation from the four largest US racial/ethnic groups. Study visits and interviews were conducted in English, Spanish, Mandarin, or Cantonese. All study procedures were reviewed and approved by the University of California San Francisco Institutional Review Board, and all participants provided consent.

A subgroup of participants were unable to attend the baseline in-person visit (n = 37 [22%] from CLSP and n = 62 [24%] from additional Bay Area recruits). For these individuals, medical records were collected, and the same structured interview was administered by telephone.

Subsequent to baseline, follow-up interviews have been conducted annually, in addition to a follow-up in-person visit at year 3 for a subset. Data for the primary analyses were drawn from the year 3 interviews (n = 343). Data for follow-up analyses were also collected from years 1, 4, 5, and 6. Retention for annual follow-up interviews through year 6 has averaged 86%.

Variables. *Frailty.* The SLICC-FI requires completion of the SLEDAI and the SDI, as well as information about comorbid conditions and patient-reported measures of physical functioning, fatigue, and pain (Table 1 and Supplementary Table 1). Much of the content of the SLICC-FI is derived from the specific items of the SLEDAI and the SDI. Each of the 48 items has been assigned a weighting. Item scores are summed and converted to a 0 to 1 scale. Scores can be categorized as robust (score 0–0.03), less fit (>0.03–0.10), least fit (>0.10–0.21), or frail (>0.21). We calculated SLICC-FI scores according to the published protocol.[8]

The FRAIL scale is a five-item, self-report questionnaire (Table 1) with scores ranging from zero to five. In the one study that used the FRAIL scale for SLE, frailty was associated with greater disease damage, higher levels of markers of inflammation, greater disability, and worse scores on a range of patient-reported outcome measures.[7] Correspondence between frailty determinations by the FRAIL scale and the FFP were examined in that study and were found to be moderate (kappa = 0.46; $P = 0.0004$). The original recommended cut-point to define frailty is ≥3, although lower cut-points have been explored and found to have more favorable accuracy.[21,22] We examined cut-points of both ≥3 and ≥2 to define frailty.

In a subset of the CLUES cohort who returned for an in-person visit at year 3, the SPPB was performed.[3] The SPPB has three components: a test of standing balance, timed chair stands, and a timed four-meter walk. Each component is scored from 0 to 4 for a total score of 0 to 12. Higher scores reflect better status. Scores of ≥10 have been categorized as robust, 8 or 9 as pre-frail, and ≤7 as frail.

We examined the 10-item Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function 10a (PF) scale[23] as a potential proxy measure of frailty that may be more readily available in both research and clinical settings.[24] PROMIS PF scores were converted to T-scores, with population mean 50 and SD 10, according to PROMIS scoring protocols. We examined a T-score one SD below the population mean (≤40), representing a moderate level of functional impairment,[25] as a possible indicator of frailty. PROMIS PF scores ≤40 were found to differentiate between functional levels that were acceptable versus unacceptable to adults.[26] Further, in another study, mean PF scores for individuals classified as frail by both the FRAIL scale and the FFP were <40 and were approximately one SD lower than scores for individuals who were not frail.[27]

Outcomes. We examined associations with both cross-sectional and longitudinal outcomes expected to reflect frailty. Cross-sectional outcomes were hospitalizations, disability, six-minute walk test, cognitive impairment, self-reported disease

Table 1. Components of the FRAIL scale, SLICC Frailty Index, and SPPB*

	SLICC Frailty Index	FRAIL scale	SPPB
General	Based primarily on SDI, SLEDAI, and SF-36 subscales	Questionnaire	Performance based
Components	• Forty-eight health "deficits" identified • Fourteen related to organ damage • Fourteen reflect active inflammation • Six reflect comorbid conditions • Fourteen related to function, mobility, health attitude, and mental health	• Fatigue • Difficulty walking up 10 steps • Difficulty walking several hundred yards • Presence of ≥5 of 11 illnesses • Weight loss	• Balance • Time to complete five chair stands • Time to complete four-meter walk
Scoring	Each deficit is scored from 0 to 1. The total number of deficits is summed and divided by 48.	Each component can receive 1 point. Points are summed for total score ranging from 0 to 5.	Each component is scored 0–4, for a total score range of 0–12, with higher scores reflecting better status.
Frailty definitions	• Robust, 0–0.03 • Less fit, >0.03–0.10 • Least fit, >0.10–0.21 • Frail, >0.21	• Robust, 0 • Prefrail, 1 or 2 • Frail, ≥3	• Robust, ≥10 • Pre-frail, 8 or 9 • Frail, ≤7

* FRAIL, Fatigue, Resistance, Ambulation, Illness, and Loss of Weight; SDI, Systemic Lupus International Coordinating Clinics Damage Index; SF-36, Short Form-36 Physical Function subscale; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; SLICC, Systemic Lupus International Collaborating Clinics; SPPB, Short Physical Performance Battery.

damage, fatigue, and depressive symptoms assessed in the same year. Longitudinal outcomes were hospitalization, functional decline, and increase in disease damage over the subsequent three years.

Hospitalizations were self-reported at each interview. The year 3 report of any versus no hospitalization in the prior year was used in cross-sectional analyses. For longitudinal analyses, we used the report of any versus no hospitalization at the subsequent year 4, year 5, or year 6 interviews.

Disability was measured with the shortened version of the Valued Life Activities (S-VLA) disability and accommodations scale.[28] The S-VLA scale assesses difficulty in specific activity domains, ranging from self-care to social and recreational activities. It has 14 items, each rated from 0 (no difficulty) to 3 (unable to perform). Scores are calculated as the mean of all items the respondent rates. No minimally important difference has been defined for the S-VLA scale, so we used one-half SD of the year 3 score (0.3) to estimate the minimally important difference.[29] Individuals whose scores worsened from the year 3 score by ≥0.3 at any of the follow-up periods were classified as having an increase in disability.

The six-minute walk test (6MWT) was also examined as a measure of functioning. The 6MWT was implemented after year 3 in-person visits had already begun, so it was available for only the subset of our sample whose in-person visits were later in the year 3 data collection period (n = 122). For the 6MWT, individuals walk at a normal speed between 2 markers delineating a 30-meter course for 6 minutes, with rest breaks as needed.[30] Scores are the total distance walked in the six-minute period.

Cognitive impairment was assessed using the Controlled Word Association Test (COWAT).[31] Responses on the COWAT were scored and transformed to z-scores, adjusted for age and education. Individuals whose z-scores were −1.5 or lower were classified as having impairment.

To assess disease damage, we used the Brief Index of Lupus Damage (BILD), which has been validated as a patient-reported proxy for the SDI.[32,33] An increase of ≥2 was used to define an increase in BILD for the longitudinal analyses. BILD scores ≥2 have previously been linked to poor health outcomes (eg, hospitalization, death).[33] The BILD was only administered at year 4 during the follow-up period.

Fatigue was measured with the four-item PROMIS Fatigue short form, transformed to T-scores with mean of 50 and SD of 10.[23] Depressive symptoms were measured with the Patient Health Questionnaire (PHQ)-8.[34] The PHQ was developed based on diagnosis criteria for depression. Scores range from 0 to 24.

Other variables. The SLEDAI[35] and SDI[36] were completed by physicians for participants who completed an in-person research visit at year 3. Sociodemographic characteristics, medication use, and comorbid conditions were self-reported by CLUES participants in the year 3 structured interviews. Age and disease duration were calculated at the time of year 3 interviews based on baseline information.

Statistical analysis. Descriptive statistics were calculated for the entire sample and for a subset of individuals who completed the in-person SPPB. We then calculated the percentage of the sample who would be classified as frail by each measure: the SLICC-FI, the FRAIL scale, SPPB, and the PROMIS PF. We then calculated the agreement between frailty determinations according to each measure by examining the percent agreement in frailty determinations and the unweighted kappa coefficients for agreement. Kappa coefficients were classified as slight (0–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and high (≥0.81).

Next, we examined cross-sectional associations between each frailty measure and the outcomes of hospitalization, S-VLA disability scale, cognitive impairment, 6MWT distance, BILD, fatigue, and depressive symptoms using logistic (hospitalization and cognitive impairment) and linear (S-VLA scale, 6MWT, BILD, PHQ, and PROMIS Fatigue) regression analyses. Two models were constructed for each outcome: (1) unadjusted and (2) adjusted for age, sex, income, and SLE duration. Finally, we examined the association of each frailty measure with subsequent hospitalizations, increase in disability, and increase in BILD longitudinally using Cox regression analyses, again constructing two models for each outcome, as defined above.

RESULTS

The sample (n = 246) had a mean ± SD age of 47.7 ± 14.0 years, was 90.2% female, and had mean ± SD SLE duration of 18.8 ± 10.5 years (Table 2). The subset who completed the SPPB (n = 166) were similar.

Prevalence of frailty. The prevalence of frailty varied considerably based on the various measures. The SLICC-FI identified 45.9% (n = 113) of the sample as frail (Table 3). In comparison, the FRAIL scale using the standard cut-point of ≥3 identified only 17.5% of the sample as frail. Using a FRAIL score of ≥2, 33.3% were classified as frail. A PROMIS PF score ≤40 identified 23.6% as frail. Among the subset completing the SPPB, 10.8% were scored as frail by SPPB. The proportion of the total sample and the SPPB subset who were defined as frail were similar for the other frailty measures. Detailed information on scoring of the SLICC-FI is shown in Supplementary Table 1. Component scores of both the FRAIL scale and the SPPB are shown in Supplementary Table 2.

Agreement among measures. Although agreement in classifications of the standard measures of frailty was relatively high, ranging from 63.2% (between SLICC-FI and SPPB) to 85.8% (FRAIL ≥3 and PROMIS PF ≤40), kappa coefficients were generally fair to moderate (Table 4). The SPPB had the lowest kappa coefficients when compared with other measures. PROMIS PF ≤40 had the most robust agreement with the other

Table 2. Characteristics of study sample*

Characteristics	Total (n = 246), mean ± SD or % (n)	Subset with SPPB scores (n = 166)
Sociodemographic		
Age, y	47.7 ± 14.0	47.0 ± 14.1
Female	90.2 (222)	89.2 (148)
Race/ethnicity[a]		
Asian	37.0 (91)	41.0 (68)
Black	10.2 (25)	8.4 (14)
Hispanic	23.2 (57)	22.4 (38)
White	29.7 (73)	27.7 (46)
Below poverty[b]	20.6 (50)	19.2 (29)
General health		
Comorbidities		
Cancer	2.0 (5)	2.4 (4)
Fibromyalgia	9.5 (23)	12.3 (20)
Diabetes	6.9 (17)	5.4 (9)
Cardiovascular disease	20.3 (50)	18.1 (30)
Lung disease	18.7 (46)	16.9 (28)
SLE-related		
Disease duration, y	18.8 ± 10.5	18.0 ± 10.3
SLEDAI	3.2 ± 3.5	3.2 ± 3.4
SDI	2.0 ± 2.1	1.9 ± 2.1
GC use	83.8 (114)	84.1 (74)
High dose GC (≥7.5 mg/day)	32.8 (44)	34.5 (30)
Immunosuppressive medication use	53.7 (132)	52.4 (87)
Outcomes, cross-sectional		
Hospitalization, cross-sectional	18.7 (46)	16.9 (28)
S-VLA scale	0.52 ± 0.60	0.48 ± 0.55
Cognitive impairment[c]	18.7 (46)	18.7 (31)
Six-minute walk test distance, meters (n = 122)	427.2 ± 95.7	427.9 ± 96.9
PROMIS Fatigue	52.1 ± 11.3	52.3 ± 11.4
PHQ-8	6.1 ± 5.1	6.1 ± 5.1
BILD	2.0 ± 2.2	1.8 ± 2.0
Outcomes, longitudinal		
Hospitalization, any during follow-up	29.3 (73)	27.7 (46)
Decline in S-VLA ≥0.3	26.9 (52)	26.4 (34)
Increase in BILD ≥2	14.2 (35)	13.3 (22)

* All data except longitudinal outcomes are from California Lupus Epidemiology Study year 3. Longitudinal outcomes occurred during years 4–6. BILD, Brief Index of Lupus Damage; GC, glucocorticoid; PHQ, Patient Health Questionnaire; PROMIS, Patient-Reported Outcomes Measurement Information System; SDI, Systemic Lupus International Coordinating Clinics Damage Index; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; SPPB, Short Physical Performance Battery; S-VLA, Short Valued Life Activities disability questionnaire.
[a] Race/ethnicity categories are mutually exclusive.
[b] Below poverty defined as <125% of the federal poverty level for household size.
[c] Defined as z-scores −1.5 or lower on the Controlled Word Association Test.

measures. Detailed examinations of the frailty classifications revealed that the SLICC-FI tended to identify individuals as frail who were not identified as frail on other scales. Results were similar among the subset of participants who completed the SPPB (Supplementary Table 3).

Table 3. Prevalence of frailty defined by SLICC Frailty Index, FRAIL scale, SPPB, and PROMIS PF ≤40*

Frailty measure	Total (n = 246)	Subset with SPPB scores (n = 166)
SLICC Frailty Index		
Total score, mean ± SD	0.20 ± 0.10	0.20 ± 0.10
Robust (0–0.03)	0	0
Less fit (>0.03–0.10)	15.9 ± 39	18.7 ± 31
Least fit (>0.10–0.21)	38.2 ± 94	38.6 ± 64
Frail (>0.21)	45.9 ± 113	42.8 ± 71
FRAIL scale		
Total score, mean ± SD	1.0 ± 1.2	1.03 ± 1.24
Not frail (0)	50.4 ± 124	50.6 ± 84
Pre-frail (1, 2)	32.1 ± 79	31.3 ± 52
Frail (≥3)	17.5 ± 43	18.1 ± 30
Frail (≥2)	33.3 ± 82	32.5 ± 54
SPPB		
Total score, mean ± SD	–	9.66 ± 1.67
Robust (≥10)	–	56.6 ± 94
Pre-frail (8, 9)	–	32.5 ± 54
Frail (≤7)	–	10.8 ± 18
PROMIS PF		
T-score ≤40, % (n)	23.6 (58)	20.5 (34)

* FRAIL, Fatigue, Resistance, Ambulation, Illnesses, and Loss of Weight; PROMIS PF, Patient-Reported Outcome Measurement Information System Physical Function 10a; SLICC, Systemic Lupus International Coordinating Clinics; SPPB, Short Physical Performance Battery.

Cross-sectional associations with outcomes. In unadjusted analysis, the SLICC-FI demonstrated significant associations with each of the cross-sectional outcomes except for cognitive impairment (Table 5). The FRAIL scale with score ≥3 and PROMIS PF with score ≤40 were significantly associated with each of the cross-sectional outcomes; the FRAIL scale with score ≥2 was associated with each outcome except hospitalization. The SPPB was associated with S-VLA score, PHQ, PROMIS Fatigue, and 6MWT distance, but not hospitalizations, BILD, or cognitive impairment. Among the subset who completed the SPPB, associations with cross-sectional outcomes were similar (Supplementary Table 4).

In analyses adjusted for age, sex, income, and SLE duration, none of the frailty measures were associated with cognitive impairment. SLICC-FI, FRAIL ≥3, and PF40 were each significantly associated with all of the other cross-sectional outcomes. FRAIL ≥2 was associated with all outcomes except hospitalization and cognition. SPPB was associated only with PHQ-8, S-VLA score, and 6MWT distance.

Table 4. Agreement of frailty classifications among measures*

	FRAIL scale (≥3)		FRAIL scale (≥2)		SPPB		PF40	
Frailty measure	Not frail	Frail	Not frail	Frail	Not frail	Frail	Not frail	Frail
SLICC-FI								
Not frail	53.7 (132)[a]	0.4 (1)	51.2 (126)[a]	2.9 (7)	54.8 (91)[a]	2.4 (4)	53.3 (131)[a]	0.8 (2)
Frail	28.9 (71)	17.1 (42)[a]	15.5 (38)	30.5 (75)[a]	34.3 (57)	8.4 (14)[a]	23.2 (57)	22.8 (56)[a]
Total % agreement	70.8		81.7		63.2		76.1	
Kappa (95% CI)	0.38 (0.29–0.48)		0.62 (0.53–0.72)		0.17 (0.06–0.28)		0.50 (0.40–0.60)	
FRAIL (≥3)								
Not frail	–	–	–	–	75.9 (126)[a]	6.0 (10)	72.4 (178)[a]	10.2 (25)
Frail	–	–	–	–	13.3 (22)	4.8 (8)[a]	4.1 (10)	13.4 (33)[a]
Total % agreement	–	–	–	–	80.7		85.8	
Kappa (95% CI)	–	–	–	–	0.23 (0.04–0.41)		0.57 (0.44–0.69)	
FRAIL (≥2)								
Not frail	–	–	–	–	63.9 (106)[a]	3.6 (6)	63.8 (151)[a]	2.9 (7)
Frail	–	–	–	–	25.3 (42)	7.2 (12)[a]	12.6 (31)	20.7 (51)[a]
Total % agreement	–	–	–	–	71.1		84.5	
Kappa (95% CI)	–	–	–	–	0.20 (0.06–0.34)		0.63 (0.52–0.73)	
SPPB								
Not frail	–	–	–	–	–	–	74.7 (124)[a]	14.5 (24)
Frail	–	–	–	–	–	–	4.8 (8)	6.0 (10)[a]
Total % agreement	–	–	–	–	–	–	80.7	
Kappa (95% CI)	–	–	–	–	–	–	0.28 (0.10–0.46)	

* Table values are % (n) unless otherwise noted. FRAIL, Fatigue, Resistance, Ambulation, Illnesses, and Loss of Weight scale; PF40, Patient-Reported Outcome Measurement Information System (PROMIS) Physical Function 10a ≤40; SLICC-FI, Systemic Lupus International Coordinating Clinics Frailty Index; SPPB, Short Physical Performance Battery; 95% CI, 95% confidence interval.
[a] Indicates agreeing classifications.

Longitudinal associations with outcomes. Seventy-three participants (29.3%) reported a hospitalization in the follow-up period (cumulative frequency: year 4, n = 30; year 5, n = 51; year 6, n = 73), and 52 participants (26.9%) had an increase in disability (cumulative frequency: year 4, n = 25; year 5, n = 44; year 6, n = 52). BILD increases from year 3 to year 4 were experienced by 35 participants (14.2%).

In unadjusted analyses, the SLICC-FI was statistically significantly associated with higher odds of follow-up hospitalizations (hazard ratio [HR] 1.9, 95% confidence interval [95% CI] 1.2–3.0), increase in disability (HR 2.7, 95% CI 1.5–4.8), and an increase in disease damage as measured by the BILD (HR 4.2, 95% CI 1.9–9.2) (Table 6). FRAIL score ≥2 was statistically significantly associated with follow-up hospitalizations (HR 2.1, 95% CI 1.3–3.3) and increase in disability (HR 2.4, 95% CI 1.4–4.2), whereas PROMIS PF ≤40 was statistically significantly associated with higher hazard of increase in disability (HR 2.0, 95% CI 1.2–3.6) and increase in disease damage (HR 3.0, 95% CI 1.6–5.9). FRAIL score ≥3 was significantly associated only with rate of hospitalization (HR 1.8, 95% CI 1.0–3.0). SPPB was not statistically significantly associated with any of the longitudinal outcomes. Among the subset of the cohort who completed the SPPB, FRAIL score ≥3 and PROMIS PF score ≤40 were also associated with hospitalization (Supplementary Table 5).

DISCUSSION

Percentages of participants identified as frail varied among the frailty measures, from 10.8% to 45.9%. Agreement among classifications ranged from slight to substantial but was generally fair to moderate. This wide range in prevalence based on differences in the method used to measure frailty is consistent with what has been noted in the general population and for rheumatoid arthritis.[37] Although most of the measures of frailty were associated with poor health outcomes among this SLE cohort—demonstrating that they measure a construct closely associated

Table 5. Cross-sectional association of frailty with health outcomes based on various definitions from study year 3*

Model characteristics	Defining frailty measure				
	SLICC-FI	FRAIL ≥3	FRAIL ≥2	SPPB	PF40
Model 1: Unadjusted					
Hospitalization					
OR (95% CI)[a]	2.6 (1.4 to 5.2)	2.2 (1.1 to 4.7)	1.7 (0.9 to 3.3)	1.5 (0.5 to 4.9)	2.6 (1.3 to 5.1)
P	0.0044	0.036	0.11	0.52	0.007
Cognitive impairment					
OR (95% CI)	1.5 (0.8 to 2.9)	2.2 (1.1 to 4.7)	2.1 (1.1 to 4.1)	2.5 (0.8 to 7.2)	2.0 (1.0 to 4.0)
P	0.21	0.036	0.022	0.10	0.050
BILD					
β (95% CI)[b]	1.7 (1.2 to 2.2)	1.0 (0.3 to 1.7)	1.4 (0.9 to 2.0)	0.6 (−0.4 to 1.6)	1.3 (0.7 to 1.9)
P	<0.0001	0.0071	<0.0001	0.22	<0.0001
PHQ-8					
β (95% CI)	5.8 (4.7 to 6.5)	7.0 (5.5 to 8.4)	6.4 (5.3 to 75)	4.0 (1.5 to 6.4)	5.9 (4.5 to 7.2)
P	<0.0001	<0.0001	<0.0001	0.0018	<0.0001
S-VLA					
β (95% CI)	0.81 (0.70 to 0.92)	0.84 (0.67 to 1.00)	0.87 (0.76 to 0.99)	0.45 (0.19 to 0.71)	1.02 (0.89 to 1.14)
P	<0.0001	<0.0001	<0.0001	0.001	<0.0001
PROMIS Fatigue					
β (95% CI)	12.4 (10.0 to 14.8)	13.3 (9.9 to 16.6)	11.6 (8.9 to 14.2)	5.9 (0.4 to 11.5)	12.6 (9.7 to 15.5)
P	<0.0001	<0.0001	<0.0001	0.036	<0.0001
6MWT distance (n = 122)					
β (95% CI)	−62.1 (−94.4 to −29.9)	−88.4 (−133.0 to −43.8)	−56.5 (−91.8 to −21.2)	−158.3 (−233.7 to −82.8)	−84.0 (−127.8 to −40.2)
P	0.0002	0.0001	0.002	<0.0001	0.0002
Model 2: Adjusted for age, sex, income, SLE duration					
Hospitalization					
OR (95% CI)[a]	2.6 (1.3 to 5.4)	2.4 (1.1 to 5.4)	1.8 (0.9 to 3.6)	2.0 (0.6 to 7.4)	2.5 (1.1 to 5.5)
P	0.0103	0.039	0.12	0.29	0.022
Cognitive impairment					
OR (95% CI)	1.1 (0.5, 2.3)	2.0 (08 to 4.8)	1.6 (0.8 to 3.5)	3.5 (1.0 to 12.6)	1.5 (0.6 to 3.3)
P	0.86	0.13	0.20	0.054	0.36
BILD					
β (95% CI)[b]	1.5 (1.0 to 2.0)	1.0 (0.3 to 1.8)	1.2 (0.7 to 1.8)	0.7 (−0.4 to 1.7)	1.2 (0.5 to 1.8)
P	<0.0001	0.0046	<0.0001	0.22	0.0006
PHQ-8					
β (95% CI)	5.9 (4.7 to 7.1)	6.9 (5.4 to 8.5)	6.7 (5.5 to 7.9)	4.2 (1.3 to 7.0)	6.0 (4.5 to 7.6)
P	<0.0001	<0.0001	<0.0001	0.0043	<0.0001
S-VLA					
β (95% CI)	0.77 (0.65 to 0.89)	0.83 (0.66 to 1.00)	0.86 (0.75 to 0.98)	0.52 (0.21 to 0.82)	0.96 (0.83 to 1.10)
P	<0.0001	<0.0001	<0.0001	0.0009	<0.0001
PROMIS Fatigue					
β (95% CI)	12.9 (10.4 to 15.5)	12.9 (9.4 to 16.5)	12.2 (9.5 to 15.0)	5.8 (−0.6 to 12.1)	13.6 (10.4 to 16.8)
P	<0.0001	<0.0001	<0.0001	0.08	<0.0001
6MWT distance (n = 122)					
β (95% CI)	−62.1 (−96.2 to −27.9)	−82.7 (−127.5 to −38.0)	−44.7 (−82.5 to −7.0)	−151.7 (−232.4 to −71.1)	−61.9 (−109.3 to −14.4)
P	0.0005	0.0004	0.021	0.0003	0.011

* Higher scores or positive β indicate worse outcomes for all measures except 6MWT distance, for which a lower score or negative β indicates a worse outcome. BILD, Brief Index of Lupus Damage; FRAIL, Fatigue, Resistance, Ambulation, Illnesses, and Loss of Weight scale; OR, odds ratio; PF40, Patient-Reported Outcome Measurement Information System (PROMIS) Physical Function 10a ≤40; PHQ, Patient Health Questionnaire; SLE, systemic lupus erythematosus; SLICC-FI, Systemic Lupus International Coordinating Clinics Frailty Index; SPPB, Short Physical Performance Battery; S-VLA, shortened Valued Life Activities disability questionnaire; 6MWT, six-minute walk test; 95% CI, 95% confidence interval.

[a] OR (95% CI) and *P*-value from logistic regression analysis.

[b] β (95% CI) and *P*-value from linear regression analysis.

Table 6. Longitudinal association of frailty with health outcomes from study years 4–6*

Model	Frailty classification				
	SLICC-FI	FRAIL ≥3	FRAIL ≥2	SPPB	PF40
Model 1: Unadjusted					
Any follow-up hospitalization					
HR (95% CI)	1.9 (1.2–3.0)	1.8 (1.0–3.0)	2.1 (1.3–3.3)	1.0 (0.4–2.6)	1.5 (0.9–2.5)
P	0.009	0.041	0.002	0.96	0.11
S-VLA decline ≥0.30					
HR (95% CI)	2.7 (1.5–4.8)	1.8 (1.0–3.5)	2.4 (1.4–4.2)	1.3 (0.4–3.6)	2.0 (1.2–3.6)
P	0.001	0.053	0.001	0.67	0.012
BILD increase ≥2					
HR (95% CI)	4.2 (1.9–9.2)	1.9 (0.9–3.9)	1.4 (0.7–2.7)	1.1 (0.3–3.8)	3.0 (1.6–5.9)
P	0.000	0.089	0.38	0.84	0.001
Model 2: Adjusted for age, sex, income, SLE duration					
Any follow-up hospitalization					
HR (95% CI)	2.0 (1.2–3.3)	1.7 (1.0–3.0)	2.2 (1.4–3.5)	0.9 (0.3–2.3)	1.5 (0.8–2.5)
P	0.008	0.06	0.003	0.75	0.17
S-VLA decline ≥0.30					
HR (95% CI)	2.4 (1.3–4.5)	1.8 (0.9–3.3)	2.0 (1.1–3.6)	1.3 (0.4–3.7)	1.7 (0.9–3.1)
P	0.004	0.08	0.015	0.67	0.094
BILD increase ≥2					
HR (95% CI)	3.2 (1.4–7.1)	1.5 (0.7–3.2)	1.1 (0.5–2.2)	0.5 (0.1–2.0)	2.1 (1.9–4.4)
P	0.006	0.32	0.80	0.36	0.035

* Table values are HR (95% CI) from Cox regression analysis. BILD, Brief Index of Lupus Damage; FRAIL, Fatigue, Resistance, Ambulation, Illnesses, and Loss of Weight scale; HR, hazard ratio; PF40, Patient-Reported Outcome Information Measurement System (PROMIS) Physical Function 10a ≤40; SLICC-FI, Systemic Lupus International Coordinating Clinics Frailty Index; SPPB, Short Physical Performance Battery; S-VLA, shortened Valued Life Activities disability questionnaire; 95% CI, 95% confidence interval.

with adverse health outcomes—the relationship between frailty and health outcomes differed according to the instrument and cut-point used to defined frailty.

The SLICC-FI identified almost half of the cohort as frail. The mean ± SD SLICC-FI score of the CLUES cohort was similar to that of the SLICC development cohort (CLUES, 0.20 ± 0.10; SLICC, 0.17 ± 0.08).[8] Similar mean SLICC-FI scores have been found in multiple cohorts, even in an incident SLE cohort.[10,12,38–40] Although the mean scores from these cohorts have been remarkably similar, the proportion of individuals classified as frail has varied, from 27.1% in the SLICC development study to 29.5% and 35.6% in two single center cohorts.[12,38] Our cohort demonstrated a frailty prevalence higher than that reported in previously published SLE cohort studies. Some of the difference may be due to differences in age—the mean age of the CLUES cohort was 48 years, whereas some of the other cohorts had mean ages of 35 to 43 years and considerably shorter duration of SLE. Greater age or longer disease duration (or both) might result in more time for SLE organ damage and comorbidities to occur. The SLICC-FI was associated with all cross-sectional outcomes except cognitive impairment in cross-sectional unadjusted analysis and with all longitudinal outcomes.

In contrast to the SLICC-FI, the SPPB identified the lowest proportion of participants as frail. SPPB scores were associated with some of the cross-sectional outcomes but had no association with longitudinal outcomes. Although the SPPB has been shown to be a strong predictor of poor outcomes in older populations and in some illness populations, it has not been used in SLE as a measure of frailty before this study. The singular focus of the SPPB on lower extremity function may limit its usability in SLE.

The FRAIL scale has been reported previously in two studies of SLE.[7,27] Both of these studies used a cut-point of ≥3 and classified 26.9% and 23.5%, respectively, of their samples as frail, compared to 17.5% in our study. When we used a cut-point of ≥2, 33.3% of our cohort was identified as frail. The FRAIL scale was associated with cross-sectional outcomes when we used both cut-points, but only scores ≥2 were associated with longitudinal outcomes. The lack of association of scores ≥3 with longitudinal outcomes may be because of the small number of individuals identified as frail using the higher cut-point. A previous study of SLE showed that the FRAIL scale was a reasonable proxy for frailty defined by the FFP.[7] We showed that it has similar agreement with the SLICC-FI.

We tested the PROMIS PF as a potential proxy measure for frailty. Using a cut-point of ≤40, a level that has been shown to differentiate between acceptable and unacceptable functioning, the PROMIS PF demonstrated high levels of agreement and moderate kappa coefficients with both the SLICC-FI and the FRAIL scale. One advantage of the PROMIS PF ≤40 is that it is already incorporated into many research settings. Because its implementation in clinical settings was recently recommended by the ACR as a quality measure for SLE[41] and it has been recommended as a measure of functional status for rheumatoid arthritis,[42] it may similarly be in place in many clinical settings. In addition, another study has

shown a strong relationship between PROMIS PF scores and frailty defined by both the FRAIL scale and the FFP.[7]

Adjusted analyses examined the association of each of the frailty measures after controlling for age, sex, income, and SLE duration. After adjustment, none of the measures was associated with cognitive impairment. Otherwise, there were no substantive differences in the results.

Among the frailty measures tested, the SLICC-FI was most consistently associated with both cross-sectional and longitudinal outcomes. However, it is also the most detailed measure and may be the most difficult to perform in routine clinical settings or even in some research settings. Additionally, the usefulness of a measure that categorizes such a high proportion of individuals as frail as a way to identify those at risk or in need of intense intervention is questionable. The SPPB's lack of associations with important outcomes suggests that it may not be the best measure of frailty for SLE. In addition, the performance-based components of the SPPB are unlikely to be feasible outside of clinical research settings. Although the FFP was not tested in this study, the need for performance-based measures likely also limits its broad usability in SLE.

Both of the questionnaire-based methods, the FRAIL scale and the PROMIS PF ≤40, appear to hold promise for use in SLE. Further work is needed to identify the best cut-point for the FRAIL scale in SLE. However, the consistency of findings between our study and the previous SLE study[7] provides evidence that the FRAIL scale can be used as a parsimonious and valid proxy for frailty as defined by both the phenotype approach (FFP) and the deficit-accumulation approach (SLICC-FI). The PROMIS PF is increasingly being used in rheumatology clinical settings and has been recommended for use in clinical settings as a quality measure for SLE. In addition to tracking function among people with SLE in general, very low scores could be used as a proxy for measuring frailty and flagged as a marker for individuals who are at particularly high risk of poor outcomes.

This study does have limitations. A longer follow-up time and a larger number of participants may have yielded stronger associations with the frailty measures. However, during the three-year follow-up period, a significant proportion of the cohort experienced a poor health outcome—29.3% were hospitalized, 26.9% had an increase in disability, and 14.2% had an increase in disease damage—and at least one frailty measure was associated with each of those outcomes. We were not able to conduct analyses to examine the risk of death conferred by each frailty measure because we did not have a sufficient number of deaths in the cohort. Disease damage was assessed only once during the follow-up period. Hospitalizations were self-reported in our cohort, which may reduce the accuracy of ascertainment. Our cohort is also a research cohort, not a clinical cohort, so the range of disease severity may have been more limited. In addition, the sample size for the analysis of the SPPB was smaller than that available for the other measures, which may have affected the results.

In summary, multiple measures of frailty appear to be associated with the risk of poor health outcomes. Both the phenotypic and the deficit-accumulation approaches to identifying frailty appear to have value, and each has advantages and disadvantages. van Onna and Boonen noted that the phenotypic approach may be more useful in evaluating frailty interventions, whereas the deficit-accumulation approach may be a more robust predictor of outcomes because it includes information from more domains and has a wider score range.[37] Although we found that both approaches functioned well as predictors of outcomes, there are significant challenges to using both the SLICC-FI and the SPPB. The SLICC-FI is complex and requires a significant amount of clinical input, whereas the SPPB requires performance-based assessments. Questionnaire-based methods, such as the FRAIL scale or the PROMIS PF, have fewer barriers to implementation and use and appear to yield similar information. Ultimately, the choice of measure may be based on the intended use, as well as the simplicity and practicality of implementing the measure.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Katz confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Fried LP, Tangen CM, Walston J, et al; Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56:M146–M156.
2. Bandeen-Roche K, Xue QL, Ferrucci L, et al. Phenotype of frailty: characterization in the Women's Health and Aging Studies. J Gerontol A Biol Sci Med Sci 2006;61:262–266.
3. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49(2):M85–M94.
4. Rockwood K. Conceptual models of frailty: accumulation of deficits. Can J Cardiol 2016;32:1046–1050.
5. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. J Nutr Health Aging 2012;16:601–608.
6. Katz PP, Andrews J, Yazdany J, et al. Is frailty a relevant concept in SLE? Lupus Sci Med 2017;4:e000186.
7. Lieber SB, Nahid M, Paget S, et al. Evaluation of a patient-reported frailty tool in women with systemic lupus erythematosus. J Rheumatol 2022;49:60–67.
8. Legge A, Kirkland S, Rockwood K, et al. Construction of a frailty index as a novel health measure in systemic lupus erythematosus. J Rheumatol 2020;47:72–81.

9. Legge A, Kirkland S, Rockwood K, et al. Evaluating the properties of a frailty index and its association with mortality risk among patients with systemic lupus erythematosus. Arthritis Rheumatol 2019;71:1297–1307.
10. Legge A, Kirkland S, Rockwood K, et al. Prediction of damage accrual in systemic lupus erythematosus using the Systemic Lupus International Collaborating Clinics Frailty Index. Arthritis Rheumatol 2020;72:658–666.
11. Legge A, Kirkland S, Rockwood K, et al. Prediction of hospitalizations in systemic lupus erythematosus using the Systemic Lupus International Collaborating Clinics Frailty Index. Arthritis Care Res (Hoboken) 2022;74:638–647.
12. Legge A, Malone A, Hanly JG. External validation of the Systemic Lupus International Collaborating Clinics Frailty Index as a predictor of adverse health outcomes in systemic lupus erythematosus. Rheumatology (Oxford) 2022;61:1919–1927.
13. Keramiotou K, Anagnostou C, Konstantonis G, et al. SLICC-Frailty Index is independently associated with impaired physical function, activities of daily living and quality of life measures. Rheumatology (Oxford) 2022;61:3808–3813.
14. Lieber SB, Nahid M, Legge A, et al. Comparison of two frailty definitions in women with systemic lupus erythematosus. Rheumatology (Oxford) 2024;63:1384–1390.
15. Dall'Era M, Cisternas MG, Snipes K, et al. The incidence and prevalence of systemic lupus erythematosus in San Francisco County, California. Arthritis Rheumatol 2017;69:1996–2005.
16. Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25(11):1271–1277.
17. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725.
18. Lim SS, Bayakly AR, Helmick CG, et al. The incidence and prevalence of systemic lupus erythematosus, 2002-2004: the Georgia Lupus Registry. Arthritis Rheumatol 2014;66:357–368.
19. Somers EC, Marder W, Cagnoli P, et al. Population-based incidence and prevalence of systemic lupus erythematosus: the Michigan lupus epidemiology and surveillance program. Arthritis Rheumatol 2014;66:369–378.
20. Ferucci ED, Johnston JM, Gaddy JR, et al. Prevalence and incidence of systemic lupus erythematosus in a population-based registry of American Indian and Alaska Native people, 2007-2009. Arthritis Rheumatol 2014;66:2494–2502.
21. Aprahamian I, Oiring de Castro Cezar N, Izbicki R, et al. Screening for frailty with the FRAIL scale: a comparison with the phenotype criteria. J Am Med Dir Assoc 2017;18:592–596.
22. Rodríguez-Laso A, Martín-Lesende I, Sinclair A, et al. Diagnostic accuracy of the FRAIL scale plus functional measures for frailty screening: a cross-sectional study. BJGP Open 2022;6:BJGPO.2021.0220.
23. PROMIS. Health Measures. Accessed September 17, 2024. https://www.healthmeasures.net/explore-measurement-systems/promis
24. Izadi Z, Kay J, Schmajuk G, et al. Evaluation of psychometric properties of the patient-reported outcomes measurement information system (PROMIS) physical function 10-item short form in systemic lupus erythematosus. Ann Rheum Dis 2019;78:1171–1172.
25. PROMIS score cut points. Health Measures. Accessed September 17, 2024. https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/promis-score-cut-points
26. Jacobson RP, Kang D, Houck J. Can Patient-Reported Outcomes Measurement Information System® (PROMIS) measures accurately enhance understanding of acceptable symptoms and functioning in primary care? J Patient Rep Outcomes 2020;4:39.
27. Lieber SB, Nahid M, Rajan M, et al. Association of baseline frailty with patient-reported outcomes in systemic lupus erythematosus at 1 year. J Frailty Aging 2023;12:247–251.
28. Katz PP, Radvanski DC, Allen D, et al. Development and validation of a short form of the Valued Life Activities Disability Questionnaire for rheumatoid arthritis. Arthritis Care Res (Hoboken) 2011;63:1664–1671.
29. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41(5):582–592.
30. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111–117.
31. Spreen O, Strauss E. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. 2nd ed. Oxford University Press; 1998.
32. Yazdany J, Trupin L, Gansky SA, et al. Brief index of lupus damage: a patient-reported measure of damage in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2011;63:1170–1177.
33. Katz P, Trupin L, Rush S, et al. Longitudinal validation of the Brief Index of Lupus Damage. Arthritis Care Res (Hoboken) 2014;66:1057–1062.
34. Kroenke K, Strine TW, Spitzer RL, et al. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009;114:163–173.
35. Buyon JP, Petri MA, Kim MY, et al. The effect of combined estrogen and progesterone hormone replacement therapy on disease activity in systemic lupus erythematosus: a randomized trial. Ann Intern Med 2005;142:953–962.
36. Gladman D, Ginzler E, Goldsmith C, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index for systemic lupus erythematosus. Arthritis Rheum 1996;39:363–369.
37. van Onna M, Boonen A. Frailty in inflammatory arthritis: a fragile construct. J Rheumatol 2023;50:463–465.
38. Lima K, Legge A, Hanly JG, et al. Association of the Systemic Lupus International Collaborating Clinics Frailty Index and damage accrual in longstanding lupus erythematosus. Arthritis Care Res (Hoboken) 2023;75:578–584.
39. Elera-Fitzcarrald C, Gamboa-Cárdenas RV, Reatégui-Sokolova C, et al. The SLICC-FI predicts damage accrual in SLE patients. Data from the Almenara Lupus Cohort. Lupus 2022;31:1666–1670.
40. Ugarte-Gil MF, Dubey J, McGwin G Jr, et al. Association of Systemic Lupus International Collaborating Clinics Frailty Index with damage in systemic lupus erythematosus patients: results from a multiethnic, multicenter US cohort of patients with lupus. Arthritis Care Res (Hoboken) 2023;75:585–589.
41. Katz PP, Barber CEH, Duarte-García A, et al. Development of the American College of Rheumatology patient-reported outcome quality measures for systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2024;76:777–787.
42. Barber CEH, Zell J, Yazdany J, et al. 2019 American College of Rheumatology recommended patient-reported functional status assessment measures in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2019;71:1531–1539.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 710–719
DOI 10.1002/acr.25495

Lived Employment Experiences of Young Adults With Childhood- and Adult-Onset Systemic Lupus Erythematosus: A Multicenter Canadian Qualitative Study

Michael A. Golding,[1] Fareha Nishat,[2] Kaitlyn A. Merrill,[1] Diane Lacaille,[3] Umut Oğuzoğlu,[1] Jennifer L. P. Protudjer,[4] Roberta L. Woodgate,[1] Jennifer N. Stinson,[2] Christine Peschken,[1] Zahi Touma,[5] and Lily Siok Hoon Lim[1]

Objective. This study examined the lived employment experiences of young adults with childhood- and adult-onset systemic lupus erythematosus (SLE).

Methods. Participants were recruited from three Canadian lupus clinics and asked to complete semistructured, qualitative video/phone interviews. Interviews were transcribed verbatim and analyzed using thematic analysis. Participants were recruited until consolidated thematic saturation.

Results. Twenty-one participants (median age: 27 years)—14 woman, 5 men, and 2 gender-nonconforming individuals—were included. Thirty-eight percent had childhood-onset SLE. Seventy-one percent of the participants were employed, 19% were looking for work, and 10% were not working and not looking for work. Qualitative analysis revealed two themes. 1) "Maintaining control internally and externally": Participants described how the ability to exercise control over their symptoms (internally) and their job (externally) allowed them to gain and maintain employment. 2) "Tough choices: Health, then work and everything else": Participants described challenges in maintaining a balance among their health, other social responsibilities, and work because of their SLE-related limitations. Within this theme, participants also offered advice on how others could best manage the conflicting demands on their time and energy, which was summarized in a subtheme called "Recommendations for others—'take care of yourself first.'"

Conclusion. When faced with the competing demands of their health (managing their SLE) and work, many young adults with SLE choose to prioritize their health, sacrificing their work or social responsibilities. Efforts aimed at promoting the employment success of young adults with SLE should inform individuals of these challenges and offer potential coping strategies.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a multisystem disease, affecting almost any organ within the body with many possible manifestations. Each patient with SLE has a personal disease with varying combinations of disease features and severity. Patients with SLE are at risk of organ damage, which could adversely impact their quality of life and life spans.[1] Although effective treatments are available, treatments are not without adverse effects.[2,3] SLE or its treatment can result in permanent alteration in organ function, ie, damage.[2,3] At any time, >50% of patients with SLE have either depression or anxiety.[4,5] Disabling fatigue, sometimes even without overt disease activity, also complicates the course and quality of life of many patients with SLE.[6,7]

Supported by the Canadian Institutes of Health Research Project (grant 168990).

[1]Michael A. Golding, MA, Kaitlyn A. Merrill, BSc, Umut Oğuzoğlu, PhD, Roberta L. Woodgate, PhD, Christine Peschken, MD, MSc, FRCPC, Lily Siok Hoon Lim, MBBS, MRCPCH, FRCPC, PhD: University of Manitoba, Winnipeg, Manitoba, Canada; [2]Fareha Nishat, MPH, Jennifer N. Stinson, PhD: The Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada; [3]Diane Lacaille, FRCPC, MSc: University of British Columbia and Arthritis Research Canada, Vancouver, British Columbia, Canada; [4]Jennifer L. P. Protudjer, PhD: University of Manitoba and George and Fay Yee Centre for Healthcare Innovation, Winnipeg, Manitoba, Canada, and Karolinska Institutet, Stockholm, Sweden; [5]Zahi Touma, MD, PhD: University of Toronto and Toronto Western Hospital, University Health Network, Toronto, Ontario, Canada.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25495).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25495.

Address correspondence via email to Lily Siok Hoon Lim, MBBS, MRCPCH, FRCPC, at llim@chrim.ca.

Submitted for publication November 15, 2024; accepted in revised form November 27, 2024.

SIGNIFICANCE & INNOVATIONS

- The current study provides the first qualitative investigation of the employment experiences of young adults with both childhood- and adult-onset systemic lupus erythematosus.
- Participants described how exercising control over both their symptoms and their job environment helped them to gain and maintain employment.
- However, even participants who were working described experiencing trouble maintaining a good balance among their health, work, and other social responsibilities because of their lupus-related symptoms.
- Given these tensions, participants stressed the importance of prioritizing one's health by practicing moderation in what they take on and by advocating for a flexible work environment.

Employment is a core social determinant of health.[8] It is important, not only for obvious financial reasons, but also as an important means to access health care or drug benefits. In addition, employment is an important means of socialization, allowing workers to build networks of social support. Becoming unemployed is associated with worse self-rated physical and mental health as well as worse quality of life.[8,9]

Prior studies in employment among adult patients with SLE showed that 46% to 51% were employed, whereas about one-third were work-disabled or otherwise not working.[10,11] Work loss was found to be higher in participants with SLE compared with controls in some studies, but not in others[11–13]; rates of re-entry into the work force were lower in patients with SLE.[11,12] Many factors have been identified to be associated with poorer employment status, such as the presence of severe disease, organ damage, cognitive dysfunction (memory issues), fatigue, and mental health issues.[13–16] There has also been some information focused on the nature of the work that patients with SLE do, eg, less autonomy and more physical exertion, that might be associated with employment outcomes.[17]

Young adulthood is a distinctive and important life stage between 18 and 30 years of age.[18] This is the time when many individuals leave their family home, finish school, establish their independence, and start their own careers and families. In social studies, failure to establish employment during young adulthood has been associated with a lifetime of financial difficulties.[9] It is conceivable that having a devastating chronic disease such as SLE could interfere with the ability of young adult patients to transition successfully through this life stage, not achieving the same expected milestones as their peers.

Most previous studies of employment in SLE have not considered the life stage of the participants.[11,12,19] This makes for reduced clarity and difficult interpretations; the employment-related options and potential for a young adult are understandably different from an individual close to retirement. A few studies have investigated the employment transitions of young adults with SLE using survey methods; however, these samples have been combined with patients with juvenile arthritis.[17,20] Qualitative studies in patients with rheumatoid arthritis assessing their lived employment experience have revealed additional aspects important to patients that were not apparent from traditional physician directed research.[21] We therefore undertook this qualitative study to examine the lived employment experience of patients with SLE during young adulthood.

METHODS

Study participants. All participants were diagnosed with SLE according to the American College of Rheumatology (ACR)/Systemic Lupus Erythematosus International Collaborating Clinics (SLICC) 1997, SLICC 2012, or ACR/The European Alliance of Associations for Rheumatology (EULAR) 2020 criteria by their rheumatologists.[22–24] Participants with either childhood-onset SLE (cSLE;<18 years) or adult-onset SLE (aSLE; ≥18 years) were eligible to participate, but they had to be within the young-adult life stage (18.0–30.0 years) at the interview. We recruited all patients who were employment eligible, including young adults who were primarily students because that is their life stage–appropriate employment. However, upon analyzing the interview data, we realized that the experiences of the students differed considerably from the participants who were no longer in school. Therefore, we analyzed the data from the students separately from those primarily working. Findings from the nonstudent sample are reported herein, and results from the student sample are to be described in a forthcoming manuscript. Participants were not eligible to participate if they 1) were unable to understand and speak English and/or 2) had severe preceding non–SLE-related physical and cognitive impairments that preclude full-time employment, even in the absence of SLE.

Participants were recruited between August 2021 and December 2022 from three adult rheumatology clinics in Canada: The University of Toronto lupus clinic (Toronto Western Hospital), the rheumatology clinic at Mount Sinai Hospital in Toronto, and the rheumatology clinics at Health Sciences Centre in Winnipeg. Eligible patients were identified before the clinic visits. Because this study was conducted during the COVID-19 pandemic, patients were approached according to their mode of visit. For in-person attendance, patients were first approached by a research assistant. For virtual (mostly phone) visits, patients were given a brief overview of the study by their rheumatologist, and those who indicated interest gave consent to be contacted by phone or email by the study personnel. The study personnel explained the study, provided written materials about the study, and obtained consent either in person or via email. This study was approved by the research ethics boards of the University of Manitoba (REB #HS23967[H2020:254]), SickKids Toronto (REB #100070980), and the Toronto Western Hospital (REB #20-5390).

Recruitment strategy. We aimed for equal recruitment of patients with aSLE and cSLE, evenly distributed through various age categories (18–20, 21–24, 25–27, and 28–30 years). In addition, the study population was purposively sampled to ensure wide representation. The factors for sampling included 1) employment status; 2) sex, with an aim to oversample males; 3) disease severity (major organ involvement [ie, central nervous system, renal or nonserositis cardiorespiratory involvement] or not); and 4) occupation types (managers; technicians/associate professionals, clerical, service/sales workers, armed forces; skilled agricultural, forestry/fishery, craft/trades, plant/machine operators, elementary occupations).[25] Participants were recruited until thematic saturation was reached.

Pre-interview questionnaire. After consent was given, patients completed a pre-interview questionnaire that collected the following information: demographics, gender roles, nature of disease, pain and fatigue, employment status, nature of work information, impacts of travel, and impacts of the pandemic. Participants had the option of completing these steps using printed forms or online via REDCap. Participants were required to complete this step before proceeding to the qualitative interview.

For nature of disease, we collected information on the type of onset (childhood or adult), history of severe disease as indicated by major organ involvement and hospitalization (ever or recent within two years), and self-reported worst SLE disease activity in the three months before the interview. Self-reported disease activity was measured using a 21-point numerical rating scale (0–10, in 0.5-point increments).[26] Pain related to SLE (within the last week) and fatigue (within the last three months) were also measured using the same numerical rating scale.

Nature of work information was collected using two measures: 1) items from the US Social Security Administration Survey of Disability and Work (SDW) and 2) the Abbreviated Job Content Questionnaire (AJCQ).[27–29] The US Social Security SDW consists of 15 items used to assess physical demands. Each item is scored as: no (1), yes sometimes (2), or yes a lot (3). The responses of all items were summed and then averaged. The scores range from 1 to 3 with higher scores representing higher physical demands.[27] The AJCQ was used to assess the contents and psychological and physical demands of work.[29] Fifteen items were scored according to four domains of skill discretion (SD), decision authority (DA), decision latitude (DL = SD + DA), psychological job demands (PJDs), and physical exertion.

Quantitative analysis. Summary statistics were calculated: median (25th–75th percentiles) for continuous variables and proportions/percentages for categorical variables. Tests were performed in Excel.

Qualitative interviews with a semistructured guide. The semistructured guide was created with input from qualitative experts and patient partners and was divided into three sections; the first and last sections were common to all participants, but the questions contained in the second section were determined by each participant's employment status. The first section served as an icebreaker, inviting participants to reflect on their experiences and perceptions of their disease and to provide broad information about their current work. The second section asked questions about their perceptions of the influence of SLE on their work status, the challenges associated with their work status, and how SLE affected them in their role. The final section asked about challenges and facilitating factors for their work status, perceptions of other possible barriers beyond SLE (eg, gender and race, as indicated by participants), their disclosure of the SLE diagnosis, and plans for work.

Participants were interviewed individually between September 1, 2021, and December 31, 2022, using video conferencing platforms (Microsoft Teams or Zoom Healthcare) or via the phone, if video conferencing was not feasible and or preferred. Three interviewers (masked for review) completed all the interviews in this study. All interviews were audio recorded.

Qualitative analysis. All interviews were transcribed verbatim by a professional transcription company. Afterwards, each transcript was analyzed by thematic analysis by two of the three interviewers (masked for review). In this approach, the analyst organizes the data by applying short descriptions (ie, codes) to relevant sections of text.[30] Once coding is complete, the analysts look for broader patterns among the codes, which are referred to as themes in the language of thematic analysis. Themes describe broad patterns in the data related to the research question and, unlike codes, which typically express one idea, contain several different aspects organized around a central concept.[30] In the present study, the three analysts worked independently but concurrently, using NVivo version 1.6.1. The analysts began the analysis by doing a preliminary coding of four interviews. Once this was complete, they met to discuss their individual coding process and reconcile differences in code descriptions. Through this meeting, the analysts developed a codebook that informed the coding of the subsequent interviews. During the coding phase, the analysts met periodically to discuss their progress and revise the codebook as necessary. When coding was complete, each analyst developed a series of preliminary themes based on their coding. The analysts then met to discuss their findings and to reach a consensus on the content and names of the themes.

After this process, one analyst performed an exploratory comparative analysis of the interviews by gender. As part of the exploratory analysis, the transcripts were organized by gender, reread, and recoded as necessary. Once this process was complete, the analyst then looked for quantitative differences in the allocation of codes among men, woman, and gender-nonconforming participants by examining both the frequency with which a particular code was used and the proportion of

participants receiving that code. The analyst also looked for more qualitative differences in the nature of the codes and themes across the three genders.

RESULTS

Study participants. Forty-three individuals consented to participate in the overall study (ie, student and nonstudent manuscripts). However, nine of the consenting individuals did not complete the qualitative interview. Of these nine participants, only two completed the pre-interview questionnaire. Given that the interview was used to determine whether participant data were included in the student sample or nonstudent sample (ie, reported herein), we were unable to determine which of the noncompleters were eligible for the current report.

In total, 21 participants were included in the final sample. More than one-third had cSLE, and over half had major organ disease and required hospitalization for management. Self-reported disease activity was low to moderate in the majority. See Table 1 for demographics and baseline clinical features of the participants.

Of the 21 participants, 90% had completed high school, and most had a postsecondary degree (67%) or at least some postsecondary training (81%). At the time of the interviews, 15 participants (71%) were employed, 4 (19%) were looking for work and 2 (10%) were not working and not looking for work. See Table 2 for a general overview of the work characteristics of the participants. Physical demands were moderate to high in the majority of participants (Figure 1). PJDs were moderate to high. Most were not in highly repetitive jobs (low to moderate SD) and had significant decision authority and latitude (low to moderate DA and DL). For a more detailed overview of participants' job demands, please see the tables contained in the supplementary materials.

Table 1. Demographics and self-reported disease histories of young adults with SLE*

Characteristics	N = 21
Median age in years (25th–75th percentile; min–max)	27 (23–29; 19–30)
Sex, n (%)	
Female	16 (76.2)
Male	5 (23.8)
Gender, n (%)	
Woman	14 (66.7)
Man	5 (23.8)
Nonbinary/nonconforming	2 (9.5)
Race, n (%)	
Asian	8 (38.1)
White	6 (28.6)
Indigenous and others[a]	7 (33.3)
Education, n (%)	
Less than high school	2 (9.5)
High school	2 (9.5)
Some college/university	3 (14.3)
College/university	10 (47.6)
Graduate degree	4 (19.0)
Onset group, n (%)	
Childhood onset (<18 years)	8 (38.1)
Major organ involvement, n (%)	
History of major organ involvement ever	12 (57.1)
Recent[b] major organ involvement	6 (28.6)
Hospitalization for SLE-related reasons, n (%)	
History of hospitalization for SLE-related reasons	11 (52.4)
Recent[b] hospitalization for SLE-related reasons	5 (38.1)
Reasons for recent hospitalization, number of events	
Flare of disease	6
Infection	3
Disease-related complications (eg, control of disease, blood pressure)	4
Median pain score (25th–75th percentile; min–max)	3 (1–6; 0–10)
Median self-rated worst SLE disease activity in the 3 mo before interview (25th–75th percentile; min–max)[c]	4 (2–6; 0–10)

* SLE, systemic lupus erythematosus.
[a] Others included multiracial individuals and those with African origins; Filipino-Canadians were included within the Asian group.
[b] Recent denotes two years before the interview.
[c] Disease activity based on a 0 to 10 (0.5-point increments) visual analog scale.

Theme 1: Maintaining control internally and externally. Above all, the ability to exercise control over one's symptoms (internally) and one's job tasks (or work autonomy, externally) was deemed by participants to be most important in attaining and maintaining employment. Although both types of control were necessary to ensure participants had the ability to maintain employment, entering the workforce was predicated on having adequate control over one's disease. Most participants reported that they had gained a considerable degree of control over lupus with medications and by understanding their own physical and mental capacities and disease triggers. This understanding did not come easy, however, because one's limits were only apparent once they were exceeded. Participants expressed that exceeding their mental and physical limits led to exacerbations of their SLE, but these experiences were important in demonstrating their limits so that they could be respected in the future.

Beyond disease management, the occupational outcomes of participants also appeared tied to their work autonomy, which is the ability to exercise control over work tasks. Although participants exercised their autonomy in different ways, it was often used to take breaks when experiencing mental or physical fatigue, move and stretch at will, and to avoid temperature extremes. With this autonomy, participants were better able to manage their symptoms in the workplace, largely by ensuring they did not exceed their limits through stress and strain (Table 3).

Remote work was also described in largely positive terms because it provided participants even greater autonomy over how they worked. For some, the flexibility afforded by remote work was used to assume more comfortable working positions, whereas others appreciated having some control over when they

Table 2. Work characteristics of young adults with SLE*

Characteristics	N = 21
Current employment state, n (%)	
Employed	15 (71.4)
Unemployed	4 (19.0)
Not working and not seeking work	2 (9.5)
Work schedule before COVID-19, n (%)[a]	
Regular office hours	4 (26.7)
Shift work, flexible hours or locations	6 (40.0)
Not applicable	5 (41.7)
Occupation groups	
Managers and professionals	7 (33.3)
Technicians/associate professionals, clerical support workers, service and sales workers, armed forces	8 (38.1)
Skilled agricultural, forestry/fishery workers, craft and trades workers, plant/machine operators/assemblers, elementary occupations	0 (0.0)
Not working	6 (28.6)
Work schedule since COVID-19, n (%)[a]	
Regular office hours	7 (33.3)
Shift work	6 (28.5)
Flexible hours or locations	2 (9.5)
Median number of hours (25th–75th percentile)[a]	40 (36.9–40)
Received accommodations at work due to lupus, n (%)[a]	2 (13.3)
Worked from home during pandemic, n (%)[a]	
Always	6 (40)
Worked from home before pandemic, n (%)[a]	
Always	1 (6.7)
Current personal income, n (%)	
<$30,000	5 (23.8)
$30,000–$59,999	7 (33.3)
≥$60,000	5 (23.8)
Not disclosed	5 (23.8)
Median self-rated physical difficulty in getting to and from work[b] (25th–75th percentile)	0 (0–1.5)
Median difficulty created by the time required to commute to work[b] (25th–75th percentile)	1 (0–2)
US Social Security Administration Survey of Disability and Work,[c] median (25th–75th percentile)	1.9 (1.8–2.4)
Abbreviated Job Content Questionnaire, median (25th–75th percentile)[d]	
SD	24 (22–25)
DA	24 (20–26)
DL	46 (41–53)
PJDs	23 (18.5–26.5)
PE	2 (1.5–4)

* DA, decision authority; DL, decision latitude; PE, physical exertion; PJD, psychological job demands; SD, skill discretion; SLE, systemic lupus erythematosus.
[a] Calculated based on 15 currently working individuals; accommodation information only available for those who were working at the time of the interview.
[b] Seven-point Likert scale. The higher the number, the greater the difficulty.
[c] Average of 15 items measuring physical demands. The higher the score, the greater the physical demands.
[d] SD: the higher the score, the more repetitive the work. DA: the higher the score, the less authority. DL: the higher the score, the less latitude. PJD: the higher the score, the fewer demands. PE: the higher the score, the fewer the physical demands.

started their day. Like participants working in person, the autonomy afforded by remote work allowed workers to better manage their intercurrent symptoms and physical and mental capacities, to continue working.

By contrast, participants who had little autonomy over their work conditions appeared to have greater difficulty staying engaged with work. Typically, participants who had less autonomy over their job were those who were limited to jobs that required less specialized training and were more physically demanding. Although lupus was not described as a direct impediment to obtaining employment, the demands of these physically exhausting jobs would inevitably prove too great, in the absence of opportunities to manage one's symptoms, forcing the individual to seek another job or stop working. This was the case among some of the participants who were not working at the time of the interviews.

Theme 2: Tough choices—health, then work, and everything else. Throughout the interviews, participants described how lupus-related capacity limitations meant that they often struggled to maintain a tenable balance between their health and other competing responsibilities (eg, work and social roles). Consequently, participants were forced to make the tough decisions to sacrifice their work, social roles, or their health to balance their competing needs. For working participants, maintaining this balance could mean working fewer hours or avoiding more cognitively or physically exhausting jobs. Even more commonly, however, participants chose to sacrifice household chores and leisure activities to prioritize rest. In some cases, prioritizing rest was described as a strategy to avoid overexertion and possibly triggering SLE symptoms. For other participants, rest was the only option because the rigors of the day left them exhausted.

For many, the ability to rest was partly facilitated by their family, friends, and romantic partners, who took on a greater share of the household duties. While few participants described this reliance on their loved ones as causing strain in their relationships, limitations in their ability to socialize were nevertheless described as a source of dissatisfaction by some. Participants who were not working experienced even greater dissatisfaction because they were forced to prioritize their health at the expense of their financial well-being and career goals. Participants who were not currently looking for work because of their lupus still held out hope that, by improving their disease control or seeking additional training, they might find a job that would allow them to strike a balance between their desire for work and their health (Table 4).

Subtheme: Recommendations for others—"take time for yourself.". When asked what advice they would give to other young adults with SLE regarding employment, many participants emphasized how practicing moderation at work and at

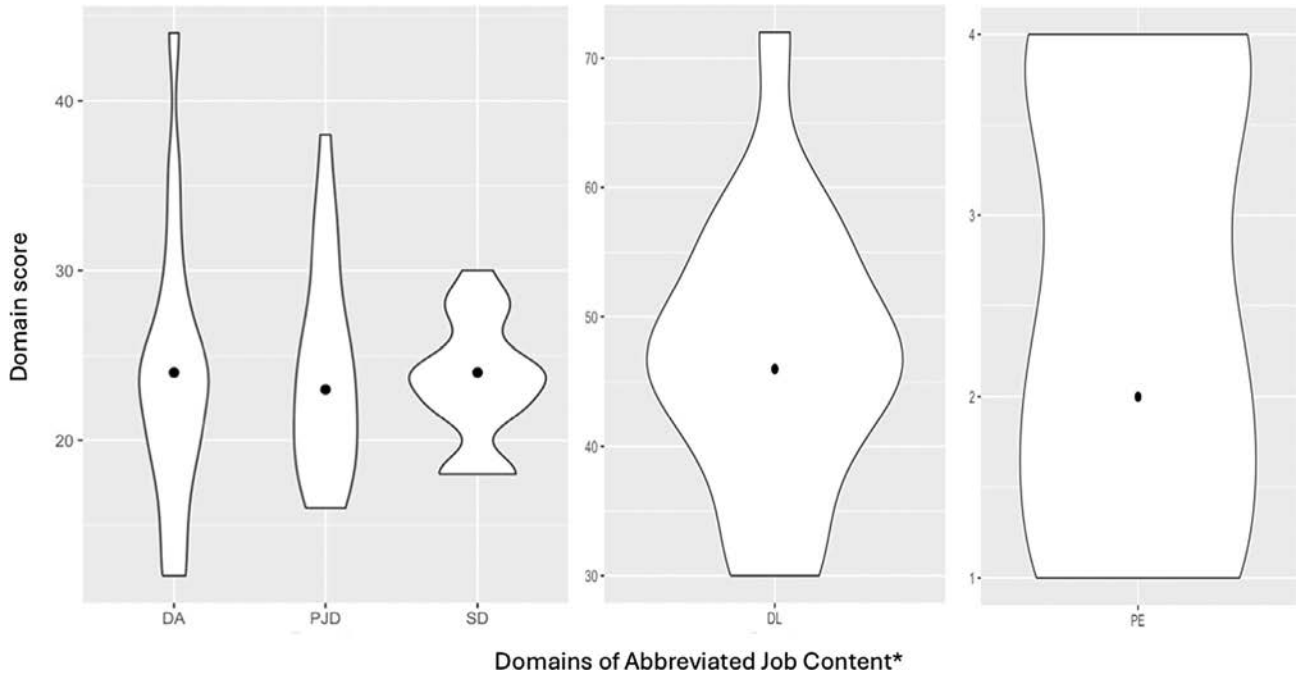

Figure 1. Psychosocial job characteristics of young adult participants with SLE as measured by the Abbreviated Job Content Questionnaire. Distribution of job domains is presented by violin plots. The *y* axis denotes domain score, and *x* axis denotes individual domains. DA: the higher the score, the less authority. PJDs: the higher the score, the less demands. SD: the higher the score, the more repetitive the work. DL: the higher the score, the less latitude. PE: the higher the score, the fewer the physical demands. Note the different *y* axis scale of DL and PE compared with SD, DA, and PJDs. The black dot denotes the median of domain score. Possible range of scores: i) 12–48 for SD, DA, and PJD; ii) 24–96 for DL; iii) 1–4 for PE. DA, decision authority; DL, decision latitude; PE, physical exertion; PJD, psychological job demands; SD, skill discretion; SLE, systemic lupus erythematosus.

home could help ensure individuals had the capacity to manage their competing demands. Participants described different ways of doing this, but most involved taking breaks when experiencing fatigue, appropriately managing workplace stress, and taking rest days when necessary. Although these practices were deemed important for maintaining a healthy balance between one's health

Table 3. Theme 1, qualitative codes, and supporting quotations of the lived employment experiences of young adults with SLE*

Theme and subthemes	Code	Representative quotes
Maintaining control internally and externally	Adaptation over time	"*Yes now I am so much better… after taking the medicine I was able to walk comfortably so I started doing Yoga*"—Woman, age 28, aSLE* (diagnosed at around age 25)
		"*[T]here were so many times that I probably put myself into completely avoidable flare-ups by just not listening to my body and not listening to my doctor*"—Woman, age 27, aSLE (diagnosed at age 21)
	The importance of flexibility	"*I am doing okay so far because it is just me. You are by yourself and you got a bunch of carts to scan and you are on your own, kind of, and you can take breaks and then do your work, you know, so I manage my day. It is not like your manager tells you can do this cart first, do that cart first. So that is kind of, you know, in my hands when to do what.*"—Woman, age 27, aSLE (diagnosed at age 24)
		"*I was like I can't do this, like my knees were always hurting, I can't really take breaks*"—Woman, age 22, aSLE (diagnosed at age 21)
	COVID-19 led to a more flexible work environment	"*I don't think I've experienced much of a hindrance [because of lupus] because everything has been very remote. However, the best part has been where I've been able to go in and out whenever it's necessary that I can take half a day off very easily if I'm overly tired*"—Woman, age 24, aSLE (diagnosed at around age 20)
	Stress and lupus	"*I have to be aware of flare-ups mainly and things that could trigger it such as like increased stress or stuff like that, so I just have to watch. I am the type of guy who gets fairly stressed out, I have noticed, my parents have noticed as well, I tend to get stressed out a lot, I guess I overthink things so I do have to watch that, but otherwise I feel like, I feel pretty normal honestly it does not really affect me*"—Man, age 24, aSLE (diagnosed at age 18)

* aSLE, adult-onset SLE; SLE, systemic lupus erythematosus.

Table 4. Theme 2, qualitative codes, and supporting quotations of the lived employment experiences of young adults with SLE*

Theme and subthemes	Code	Representative quotes
Tough choices: health, then work and everything else	Change in aspirations	"*I went and talked with other people trying to find ways to still do my profession even if it might not be what I…like instead of practicing as a general practitioner maybe I could specialize in something that would be less hands-on, it would be less strenuous for me*"—Woman, age 29, aSLE (diagnosed in third year of university)
	Limited job options	"*It [lupus] might limit you from very cut-throat environments, like start-ups, where you got to work like 80-hour weeks. Lupus patients aren't really cut out for the kind of work, not only is it a fatigue thing, but you probably can't dedicate your time as much…*"—Man, age 29, cSLE (diagnosed at age 11)
	Lupus leaves little energy for much else	"*It does end up making me very tired and sore, so I usually end up, like once I am done with work, I usually just close my laptop and lie down for about 3–4 hours and I might even just like fall asleep, but it's just not at that point it's like my body is so exhausted that it just needs to rest.*"—Woman, age 24, aSLE (diagnosed at around age 20)
	Social support—friends and family	"*Well I mean my partner is a fantastic help, he does a lot. See part of the deal when we moved to this apartment, because we did not have our own laundry unit, that he would do all the laundry, but then when we moved…he ended up also doing most of the dishes, he also does the Swiffer…So he is taking care of a lot of things and he is amazing, like he takes care of me, he you know gets the groceries, he makes sure I take my pills if I am being grumpy and trying to procrastinate*"—Gender nonconforming, age 30, aSLE (diagnosed at age 27)
	Hope, despite disability	"*…once I get my lupus, like controlled, yes I am going to try and find work. It is probably going to be physical labor work, unless I finish my schooling.*"—Man, age 30, aSLE (diagnosed at around age 22)

* SLE, systemic lupus erythematosus; aSLE, adult-onset SLE; cSLE, childhood-onset SLE.

and their job, it was apparent that implementing them could invite a sense of shame for some, as participants tended to qualify their statements about rest by underscoring the difference between rest and laziness. In justifying the need for rest, participants appeared to reframe rest as something indulgent to an investment in oneself.

However, taking additional time to manage one's symptoms at work through breaks and pacing is often dependent on the flexibility of one's job and employer. This fact was not lost on the participants, as some also stressed the importance of asking for both formal and informal accommodations when employers do not readily provide the flexibility that is necessary for managing the symptoms of lupus (Table 5).

Gender comparisons. No major differences were found in the themes across genders; however, it is interesting to note that both gender-nonconforming participants reported finding support for their health- and work-related concerns in online communities. By comparison, only 3 of the 18 cisgendered participants (17%) reported turning to online friends and communities for support.

DISCUSSION

This is the first study to focus on the lived employment experience of young adults with SLE. Although most participants wanted to continue working, this state required the coexistence

Table 5. Theme 2 subtheme, qualitative codes, and supporting quotations of the lived employment experiences of young adults with SLE*

Subtheme	Code	Representative quotes
Recommendations for others—take care of yourself first	Importance of moderation	"*You need to make sure you have work-life balance, you need to allow yourself to disconnect from work when you leave*"—Woman, age 24, cSLE (diagnosed at age 10 or 11)
	Reframing idleness	"*Always take time to rest…I feel like a lot of people who are chronically ill can forget that, they can feel selfish taking time for themselves…you need to do it, it is for your health.*"—Woman, age 22, aSLE (diagnosed at age 21)
		"*Take the opportunity to take those rest days. Take those mental check-in days because you will need them and you shouldn't be afraid of realizing them and taking in those moments – they'll make the rest of the times much easier.*"—Woman, age 24, aSLE (diagnosed at age 20)
	Importance of advocacy	"*We can't will ourselves into being able-bodied and be able to function as a normal person. We can't force ourselves to do that, if you try, you will hurt yourself and you will burn out and you need to find a job that will accommodate you and you need to ask for those accommodations up front and if a workplace does not approve your accommodations or starts treating you badly because of your disability put them on blast.*"—Gender nonconforming, age 30, aSLE (diagnosed at age 27)

* SLE, systemic lupus erythematosus; aSLE, adult-onset SLE; cSLE, childhood-onset SLE.

of several factors, including optimal management of SLE, constant self-management of symptoms, work autonomy, and having supportive relationships to help them with daily living. Most participants reported making conscious efforts to prioritize their health over other activities or social roles to continue working. Surprisingly but importantly, the themes did not differ by gender.

To date, several studies have quantitatively examined workforce participation and risk factors for work disability among patients with SLE.[31] However, few have focused on the lived employment experiences of people with lupus. Previous qualitative research had focused primarily on middle-aged adults with SLE.[32–34] As such, the current study provides new insight into the barriers and facilitators for employment among young adults with SLE. These young adults are at an earlier stage of their careers and have less experience and fewer resources to deal with work challenges compared with older adults. Understanding their unique barriers and facilitators will inform the development of supports aimed at helping them enter and stay in the workforce, which could, in turn, have impacts on subsequent socioeconomic attainments.

Many participants reported making various sacrifices to their social lives, daily routines, and careers to ensure they do not exceed their physical and emotional limits to continue working. In fact, exercising moderation at work and at home was deemed so critical that many participants cited this as a key piece of advice that should be provided to young adults with SLE before they enter the workforce. Although previous studies have connected SLE to social role restrictions, few have framed these as a proactive strategy aimed at promoting a balance among one's work, health, and other social responsibilities.[35–39] Considering these findings, young adults with SLE who are entering the workforce should be encouraged to exercise mental and physical balance whenever possible. However, these conversations should also strive to address the guilt that can beset individuals when they are engaged in activities that are not inherently productive (eg, rest, leisure activities).

In addition to moderation, work autonomy was also found to have an important influence on the work outcomes of participants. Work autonomy allowed participants the ability to modify their job in response to their physical and mental states (especially fatigue and pain) during work and thereby maintain employment. Unsurprisingly, remote work (during the pandemic) was cited as particularly beneficial because it afforded participants with control over where, when, and how they chose to work. Given the accumulating evidence that work autonomy is beneficial for patients with chronic illnesses, it stands to reason that teens and young adults with SLE should be counseled on the benefits of autonomy to ensure they make an informed choice when deciding on a career.[12,33,34,40] However, employers can also play a role in improving the employment outcomes of individuals with SLE and other chronic conditions by offering flexible working conditions, including the option for remote work.

When discussing career options, it may be useful for young adults with SLE to consider the relationship among education, skills, and workplace autonomy. Findings from the current study suggest that individuals with higher levels of education and more specialized skills were better able to access positions that afforded them greater latitude over their time and tasks. This finding is consistent with previous research.[41] With this in mind, young people with SLE may want to consider whether their educational plans place them in a position to obtain a career that provides them enough autonomy to manage their symptoms throughout the workday.

Our study did not find gender differences in the work-related experiences of young adults with SLE. Our study is the only one, to our knowledge, that attempted to examine the effect of gender on the lived employment experiences of young adults with SLE. We wonder whether the lack of gender effect stems from the fact that the participants are young and few had started their own families, as gender roles are often related to caregiving (eg, caring for elderly or children).[42,43] We recognize the numbers of men and nonbinary gendered individuals in our study were small, and this analysis is only exploratory. Given the well-recognized role that gender plays in employment, the impact of gender in the employment of patients with SLE deserves to be studied as a primary question in a larger study with more balanced gender groups.[44]

Given that SLE is more common in individuals from marginalized racial/ethnic groups and in women, we also specifically asked the participants whether they experienced any other perceived barriers to their employment related to their race, gender or other characteristics. This did not reveal any consistent trend.

Our study has limitations. Individuals with higher levels of education were overrepresented. Approximately 90% of participants in the current sample had reported finishing high school. This high school completion rate was similar to the 1000 Faces of Lupus cohort of Canadian adult patients with SLE (~80%) and the general Canadian population.[45–47] However, the proportion of postgraduate degree holders was about double (19% vs 9%) that of the Canadian population. Although this finding is not totally surprising because less educated patients are known to be less likely to consent to research studies,[48] the disproportionate number of educated participants does mean that some caution is needed when generalizing the findings to individuals with lower levels of education.

Given the relationship between education and work autonomy, it is also not surprising that most participants had jobs that afforded them significant autonomy and were more skilled and psychologically demanding but less physically demanding. Even in this relatively high-achieving group, working was not easy. It is conceivable that working is even more challenging for those with less education, working less skilled and more physically demanding jobs because they would probably have less autonomy in their work. There remains a need for more qualitative investigations into the employment experiences of less educated or skilled

patients with SLE because their coping strategies and the abilities to negotiate with employers will be different.

Our results may be more biased toward those with severe disease because participants were recruited from university lupus clinics. This should be borne in mind when interpreting our results. About one-third of our participants had recent hospitalization. By comparison, the 1000 Faces of Lupus Cohort patients had a mean annual hospitalization rate of 7.6%.[49] Given that we were aiming at a smaller subgroup (ie, young adults) of patients with SLE, it was more efficient to focus our recruitment efforts to a few centers with higher volumes of patients with SLE.

Our study provides insights into the lived work experiences of young patients with SLE at the critical life stage of young adulthood. Although there was a lot of enthusiasm for work, many participants struggled to balance all of their competing responsibilities, forcing them to prioritize their health by sacrificing their work, social responsibilities, or both. Efforts aimed at promoting the employment success of young adults with SLE should inform individuals of these challenges and offer potential coping strategies to help mitigate them. Although many will likely come to learn how to best balance their work, health, and social life in time, providing information on navigating SLE and work in an effective and in a timely manner before workforce entry or soon after diagnosis could aid in realistic career choices and life planning.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Lim confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Gladman D, Ginzler E, Goldsmith C, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. Arthritis Rheum 1996;39(3):363–369.
2. Lim LSH, Pullenayegum E, Lim L, et al. From childhood to adulthood: the trajectory of damage in patients with juvenile-onset systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69(11):1627–1635.
3. Ugarte-Gil MF, Mak A, Leong J, et al. Impact of glucocorticoids on the incidence of lupus-related major organ damage: a systematic literature review and meta-regression analysis of longitudinal observational studies. Lupus Sci Med 2021;8(1):e000590.
4. Chalhoub NE, Luggen ME. Depression-, pain-, and health-related quality of life in patients with systemic lupus erythematosus. Int J Rheumatol 2022;2022:6290736.
5. Liao J, Kang J, Li F, et al. A cross-sectional study on the association of anxiety and depression with the disease activity of systemic lupus erythematosus. BMC Psychiatry 2022;22(1):591.
6. Elefante E, Tani C, Stagnaro C, et al. Impact of fatigue on health-related quality of life and illness perception in a monocentric cohort of patients with systemic lupus erythematosus. RMD Open 2020;6(1):e001133.
7. Tench CM, McCurdie I, White PD, et al. The prevalence and associations of fatigue in systemic lupus erythematosus. Rheumatology (Oxford) 2000;39(11):1249–1254.
8. Hergenrather KC, Zeglin RJ, McGuire-Kuletz M, et al. Employment as a social determinant of health: a review of longitudinal studies exploring the relationship between employment status and mental health. Rehab Res Policy Educ 2015;29(3):261–290.
9. Wray LR. The social and economic importance of full employment. Levy Economics Institute of Bard College working paper 560. 2009. Accessed [August 8, 2023]. https://www.levyinstitute.org/publications/the-social-and-economic-importance-of-full-employment
10. Baker K, Pope J. Employment and work disability in systemic lupus erythematosus: a systematic review. Rheumatology (Oxford) 2009;48(3):281–284.
11. Yelin E, Tonner C, Trupin L, et al. Work loss and work entry among persons with systemic lupus erythematosus: comparisons with a national matched sample. Arthritis Rheum 2009;61(2):247–258.
12. Yelin E, Trupin L, Katz P, et al. Work dynamics among persons with systemic lupus erythematosus. Arthritis Rheum 2007;57(1):56–63.
13. Campbell R Jr, Cooper GS, Gilkeson GS. The impact of systemic lupus erythematosus on employment. J Rheumatol 2009;36(11):2470–2475.
14. Panopalis P, Clarke AE, Yelin E. The economic burden of systemic lupus erythematosus. Best Pract Res Clin Rheumatol 2012;26(5):695–704.
15. Panopalis P, Julian L, Yazdany J, et al. Impact of memory impairment on employment status in persons with systemic lupus erythematosus. Arthritis Rheum 2007;57(8):1453–1460.
16. Appenzeller S, Cendes F, Costallat LTL. Cognitive impairment and employment status in systemic lupus erythematosus: a prospective longitudinal study. Arthritis Rheum 2009;61(5):680–687.
17. Jetha A, Badley E, Beaton D, et al. Unpacking early work experiences of young adults with rheumatic disease: an examination of absenteeism, job disruptions, and productivity loss. Arthritis Care Res (Hoboken) 2015;67(9):1246–1254.
18. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. Am Psychol 2000;55(5):469–480.
19. Lawson EF, Hersh AO, Trupin L, et al. Educational and vocational outcomes of adults with childhood- and adult-onset systemic lupus erythematosus: nine years of followup. Arthritis Care Res (Hoboken) 2014;66(5):717–724.
20. Jetha A, Badley E, Beaton D, et al. Transitioning to employment with a rheumatic disease: the role of independence, overprotection, and social support. J Rheumatol 2014;41(12):2386–2394.
21. Lacaille D, White MA, Backman CL, et al. Problems faced at work due to inflammatory arthritis: new insights gained from understanding patients' perspective. Arthritis Rheum 2007;57(7):1269–1279.
22. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725.
23. Petri M, Orbai AM, Alarcón GS, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum 2012;64(8):2677–2686.

24. Aringer M, Costenbader K, Daikh D, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71(9):1400–1412.
25. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52:1893–1907.
26. Elera-Fitzcarrald C, Vega K, Gamboa-Cárdenas RV, et al. Reliability of visual analog scale and numeric rating scale for the assessment of disease activity in systemic lupus erythematosus. J Clin Rheumatol 2020;26(7S):S170–S173.
27. Lando ME, Cutler RR, Gamber E. 1978 Survey of Disability and Work: Data Book. US Social Security Administration; 1982.
28. Bye B, Schechter E. 1978 Survey of Disability and Work: Technical Introduction. Social Security Administration Publication; 1982.
29. Karasek R, Brisson C, Kawakami N, et al. The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. J Occup Health Psychol 1998;3(4):322–355.
30. Clarke V, Braun V. Successful Qualitative Research: A Practical Guide for Beginners. SAGE Publications; 2013.
31. Blomjous BS, Gajadin GRS, Voskuyl AE, et al. Work participation in patients with systematic lupus erythematosus: a systematic review. Rheumatology (Oxford) 2022;61(7):2740–2754.
32. Calderon O, Giffords ED, Kane K. What do we need? Accommodations in the workplace: people with lupus share their work experiences. Disabil Stud Q 2023;42(3–4).
33. Cardwell FS, Dixon J, Clarke AE, et al. Canadian workplace experiences of systemic lupus erythematosus (SLE). Lupus Sci Med 2021;8(1):e000536.
34. Ubhi M, Dubey S, Gordon C, et al. Understanding the impact of systemic lupus erythematosus on work amongst South Asian people in the UK: an explorative qualitative study. Lupus 2021;30(9):1492–1501.
35. Dixon J, Cardwell FS, Clarke AE, et al. Choices are inevitable: a qualitative exploration of the lifecosts of systemic lupus erythematosus. Chronic Illn 2022;18(1):125–139.
36. Robinson D Jr, Aguilar D, Schoenwetter M, et al. Impact of systemic lupus erythematosus on health, family, and work: the patient perspective. Arthritis Care Res (Hoboken) 2010;62(2):266–273.
37. Sterling KL, Gallop K, Swinburn P, et al. Patient-reported fatigue and its impact on patients with systemic lupus erythematosus. Lupus 2014;23(2):124–132.
38. Aim MA, Queyrel V, Tieulié N, et al. Importance of temporality and context in relation to life habit restrictions among patients with systemic lupus erythematosus: a psychosocial qualitative study. Lupus 2022;31(12):1423–1433.
39. Mattsson M, Möller B, Stamm T, et al. Uncertainty and opportunities in patients with established systemic lupus erythematosus: a qualitative study. Musculoskeletal Care 2012;10(1):1–12.
40. Booth S, Price E, Walker E. Fluctuation, invisibility, fatigue–the barriers to maintaining employment with systemic lupus erythematosus: results of an online survey. Lupus 2018;27(14):2284–2291.
41. Edgerton JD, Roberts LW, von Below S. Education and quality of life. In: Land KC, Michalos AC, Sirgy MJ, eds. Handbook of Social Indicators and Quality of Life Research. Springer; 2011:265–296.
42. Katz-Wise SL, Priess HA, Hyde JS. Gender-role attitudes and behavior across the transition to parenthood. Dev Psychol 2010;46(1):18–28.
43. Kuersten-Hogan R, McHale JP. The transition to parenthood: a theoretical and empirical overview. In: Kuersten-Hogan R, McHale JP, eds. Prenatal Family Dynamics: Couple and Coparenting Relationships During and Postpregnancy. Springer; 2021:3–21.
44. Moyer M. Women and paid work. In: Women in Canada: A Gender-Based Statistical Report. Statistics Canada; 2017:1–38.
45. Canada leads the G7 for the most educated workforce, thanks to immigrants, young adults and a strong college sector, but is experiencing significant losses in apprenticeship certificate in key trades. Statistics Canada. 2022. Accessed [September 4, 2024]. https://www150.statcan.gc.ca/n1/daily-quotidien/221130/dq221130a-eng.htm
46. Baker K, Pope J, Fortin P, et al; 1000 Faces of Lupus Investigators; CaNIOS (Canadian Network for Improved Outcomes in SLE). Work disability in systemic lupus erythematosus is prevalent and associated with socio-demographic and disease related factors. Lupus 2009;18(14):1281–1288.
47. George A, Wong-Pak A, Peschken CA, et al; 1000 Canadian Faces of Lupus Investigators. Influence of education on disease activity and damage in systemic lupus erythematosus: data from the 1000 Canadian Faces of Lupus. Arthritis Care Res (Hoboken) 2017;69(1):124–132.
48. Baquet CR, Commiskey P, Daniel Mullins C, et al. Recruitment and participation in clinical trials: socio-demographic, rural/urban, and health care access predictors. Cancer Detect Prev 2006;30(1):24–33.
49. Lee J, Peschken CA, Muangchan C, et al. The frequency of and associations with hospitalization secondary to lupus flares from the 1000 Faces of Lupus Canadian cohort. Lupus 2013;22(13):1341–1348.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 720–726
DOI 10.1002/acr.25486

Gout Flares After Stopping Anti-Inflammatory Prophylaxis: A Rapid Literature Review and Meta-Analysis

Lisa K. Stamp,[1] Christopher Frampton,[2] Jeff A. Newcomb,[3] James R. O'Dell,[4] Ted R. Mikuls,[4] and Nicola Dalbeth[5]

Objective. The aim of this research was to determine how common gout flares are after ceasing anti-inflammatory prophylaxis.

Methods. A rapid literature review and meta-analysis were undertaken. PubMed was searched from inception to February 2024. Eligibility criteria included any clinical trial of people with gout with at least one arm starting or intensifying urate-lowering therapy (ULT) with coprescription of anti-inflammatory prophylaxis and that had the percentage of participants experiencing one or more gout flares reported during and after the period of prophylaxis. Random effects meta-analyses were used to generate pooled estimates of the percentage of participants experiencing one or more flares in each period.

Results. Six trials were included, together with aggregated, unpublished data from the VA STOP Gout trial (2,972 participants). Pooled random effects estimates of the percentage of participants having one or more gout flares were 14.7% (95% confidence interval [CI] 11.3–18.5%) during prophylaxis, 29.7% (95% CI 22.9–37.0%) in the three-month period after ceasing prophylaxis, and 12.2% (95% CI 6.8–19.0%) during the last study period. The mean difference in the percentage of participants having one or more gout flare while taking prophylaxis and immediately after ceasing prophylaxis was −14.8.0% (95% CI −21.2% to −8.5%; $P < 0.0001$). The mean difference from the period immediately following prophylaxis discontinuation compared to the last study period was 16.0% ($P < 0.001$). Sensitivity analyses indicated no material effects of prophylaxis duration, trial duration, ULT class, or placebo arms.

Conclusion. Gout flares are common after stopping anti-inflammatory prophylaxis but return to levels seen during prophylaxis. Patients should be cautioned about the risk of gout flares and have a plan for effective gout flare management in the three months after stopping anti-inflammatory prophylaxis.

INTRODUCTION

Gout flares are common after starting or intensifying urate-lowering therapy (ULT), leading to the recommendation that anti-inflammatory prophylaxis should be coprescribed during the initial period of ULT.[1–3] Anti-inflammatory prophylaxis is usually prescribed as daily low-dose colchicine or nonsteroidal anti-inflammatory drugs (NSAIDs). The increase in gout flares after starting ULT is thought to be due to the dissolution of monosodium urate (MSU) crystals secondary to treatment-related reductions in serum urate levels. This is supported by evidence that the likelihood of experiencing a gout flare after starting ULT is directly related to the magnitude of the serum urate level reduction from baseline.[4] Although gout flares paradoxically increase after starting ULT, with sustained reduction in serum urate levels gout flares reduce and cease, although this is delayed by many months or even years, likely due to the slow dissolution and depletion of total body urate stores.[5]

The recommended duration of anti-inflammatory prophylaxis has varied over time. In the 2012 American College of Rheumatology (ACR) guidelines, anti-inflammatory prophylaxis was

[1]Lisa K. Stamp, FRACP: University of Otago, Christchurch and Health New Zealand, Te Whatu Ora Waitaha, Christchurch, New Zealand; [2]Christopher Frampton, PhD: University of Otago, Christchurch, Christchurch, New Zealand; [3]Jeff A. Newcomb, BS: VA Nebraska-Western Iowa Health Care System, Omaha; [4]James R. O'Dell, MD, Ted R. Mikuls, MD: VA Nebraska-Western Iowa Health Care System and the University of Nebraska Medical Center, Omaha; [5]Nicola Dalbeth, FRACP: Health New Zealand, Te Whatu Ora Te Toka Tumai, and the University of Auckland, Auckland, New Zealand.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25486.

Address correspondence via email to Lisa K. Stamp, FRACP, at lisa.stamp@cdhb.health.nz.

Submitted for publication August 15, 2024; accepted in revised form December 12, 2024.

SIGNIFICANCE & INNOVATIONS

- Gout flares are common after ceasing anti-inflammatory prophylaxis, but with time they return to levels seen while taking prophylaxis.
- For some individuals, longer periods of anti-inflammatory prophylaxis may be required.
- Patients should have a plan for effective gout flare management in the three months after stopping anti-inflammatory prophylaxis.

recommended for the greater of (1) a six-month duration, (2) three months after achieving the target serum urate level for individuals without tophi, or (3) six months after achieving the target serum urate level, in which there has been resolution of tophi previously detected on physical examination.[6] In the 2016 EULAR and 2017 British Society of Rheumatology gout guidelines, anti-inflammatory prophylaxis for up to six months was recommended.[1,2] In the most recent 2020 ACR gout guidelines, anti-inflammatory prophylaxis was recommended for three to six months, with ongoing evaluation and continued prophylaxis as needed if there were ongoing gout flares.[3]

Most clinical trials examining the efficacy of anti-inflammatory prophylaxis when starting ULT have been of short duration (12–24 weeks)[7] and have not extended beyond the period of prophylaxis. However, a recent study with gout flares as the primary end point examined low-dose colchicine compared to a placebo for the first six months when starting allopurinol, with a further six months of follow-up after colchicine or the placebo was discontinued.[8] Although colchicine-treated participants experienced fewer flares during active treatment, an unexpected finding from this study was a rise in the number of participants experiencing a gout flare after stopping colchicine, which was not observed in participants receiving a placebo. Therefore, the net effect was no difference in the mean number of gout flares per month over the entire 12-month study period between those who received colchicine and those who received a placebo.[8] The aim of this study was to determine how common gout flares are after ceasing anti-inflammatory prophylaxis and whether this varies by duration of anti-inflammatory prophylaxis.

MATERIALS AND METHODS

A rapid literature review[9,10] and meta-analysis were undertaken. Eligibility criteria included any clinical trial in people with gout with at least one arm starting or intensifying ULT in which anti-inflammatory prophylaxis was coprescribed and that had the percentage of participants with at least one gout flare reported both during and after the period of anti-inflammatory prophylaxis. Ethical approval was not required.

PubMed was searched from inception to February 2024 using the individual names of ULTs ("allopurinol," "febuxostat," "probenecid," "benzbromarone," "lesinurad," or "pegloticase" AND "gout") and was limited to English language studies, but not limited by year of publication. Titles and abstracts were reviewed by one reviewer (LKS), and full texts were independently reviewed by two reviewers (LKS and ND) for data extraction. When data were not reported, an invitation was sent to the study investigators requesting aggregated unpublished data.

EndNote X9 software was used to manage records retrieved from the PubMed search. A customized data extraction form created in Microsoft Excel was used to capture data including total number of participants, male-to-female proportions, mean (SD) age, type and doses of ULT, and type and duration of anti-inflammatory prophylaxis. The percentage of participants having at least one gout flare in three distinct observation periods was collected as follows: (1) just before stopping anti-inflammatory prophylaxis, (2) within three months of stopping anti-inflammatory prophylaxis, and (3) last period of trial available. Duration of the observation periods was determined by what was reported within the individual studies and thus could not be standardized.

Statistical analysis. The point estimates of the percentages of participants with one or more gout flares were transformed using the arcsine square root transformation and summarized using a random effects model.[11] The summarized changes in the percentages were calculated as a risk difference for the studies reporting the relevant paired time periods and were also pooled using a random effects model. Sensitivity analyses based on anti-inflammatory prophylaxis duration, trial duration, ULT class, and excluding placebo arms were undertaken. Analysis by different anti-inflammatory drugs (eg, colchicine, NSAIDs) was not possible given that studies did not report data stratified by type of anti-inflammatory drug administered. Standard I^2 statistics were used to test for heterogeneity of results among studies. All statistical analysis were performed using MedCalc for Windows, version 19.4 (MedCalc Software).

RESULTS

Study selection. Literature searches revealed a total of 422 clinical trials. After duplicates were removed, titles of 330 clinical trials were reviewed, with 196 excluded. Abstracts of 134 articles were reviewed, with 66 included for full-text review. Of the 66 trials, 6 were included in the final analysis (Figure 1). Aggregated unpublished data were also provided from the VA STOP Gout trial,[12] giving a total of seven studies, which collectively included 2,972 participants starting or intensifying ULT. The main reasons for exclusion were

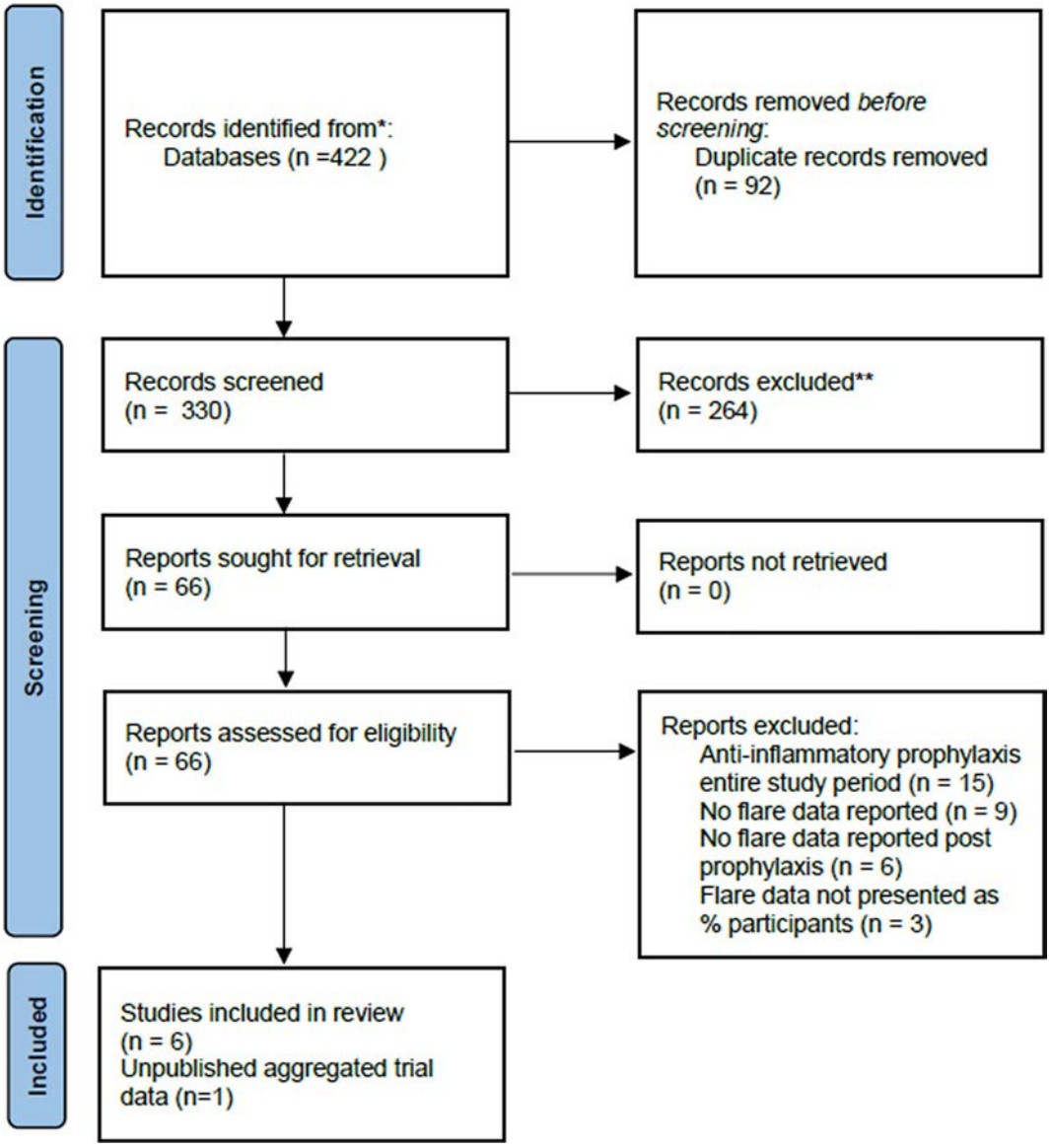

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-analyses flow chart of studies.

administration of anti-inflammatory prophylaxis for the entire study period (n = 15), no reporting of gout flare data at any time point (n = 9), no reporting of gout flares after anti-inflammatory prophylaxis was ceased (n = 6), or gout flare data reported in ways other than percentage of participants with at least one gout flare (n = 3). Details of the seven trials are outlined in Table 1. Trial duration ranged from 24 to 104 weeks and anti-inflammatory prophylaxis duration from 8 to 48 weeks. The majority of participants were male (93.3%, range 84.0–98.8%) with a mean age of 53.0 years (range of means 47.6–62.7 years). The duration of the observation three periods (1) just before stopping anti-inflammatory prophylaxis, (2) within three months of stopping anti-inflammatory prophylaxis, and (3) last period of trial available are shown in Table 2.

Percentage of participants with at least one gout flare during each of the three time periods. Pooled gout flare data from each of the three time periods are shown in Table 3 and Figure 2. During the period on anti-inflammatory prophylaxis, the pooled random effects estimate of the percentage of participants having at least one gout flare was 14.7% (95% confidence interval [CI] 11.3–18.5%; I^2 = 83.6%). During the three-month period after ceasing anti-inflammatory prophylaxis, the pooled random effects estimate of the percentage of participants having at least one gout flare was 29.7% (95% CI 22.9–37.0%; I^2 = 92.4%). During the last period of the study, the pooled random effects estimate of the percentage of participants having at least one gout flare was 12.2% (95% CI 6.8–19.0%; I^2 = 93.6%).

There was a significant mean difference in the percentage of participants having at least one gout flare while taking anti-

Table 1. Details of the seven trials included in the analysis.

First author	Trial name	Total *n*	Trial duration (weeks)	FBX		ALLO		LES (dose) + allo		Placebo	Anti-inflammatory prophylaxis		Flare definition
				Dose	*n*	Dose	*n*	Dose	*n*	*n*	Drug	Duration weeks	
Bardin (2017)[22]	CLEAR 2	610	52			Up to 800 or 900	206	200 400	204 200		NSAID or colchicine	20	Requiring treatment
Becker (2005)[23]	FACT	760	52	80 120	256 251	300	253				NSAID or colchicine	8	Requiring treatment
Becker (2005)[24]		153	4	40 80 120	37 40 38					38	Colchicine	14	Not reported
Dalbeth (2017)[25]	Early gout	314	104	40-80	157					157	Naproxen or colchicine	24	Confirmed by investigator
O'Dell (2022)[12]	STOP Gout	940	72	Up to 120	472	≤800	468				NSAID, colchicine or prednisone	24-48	Gaffo definition[19, 26]
Stamp (2023)[8]		200	52			≤900 ≤900	100 100				Colchicine Placebo	24	Self reported
Yamanaka 2018)[27]	Fortune1	255	24	10-40 40 40	96 95 50						No prophylaxis Colchicine No prophylaxis	12	Treatment for flare required

inflammatory prophylaxis and soon after ceasing prophylaxis (mean difference −14.8% [95% CI −21.2% to −8.5%]; $P < 0.0001$; $I^2 = 86.6\%$) (Figure 3). There was also a significant mean difference in the percentage of participants having at least one gout flare soon after ceasing anti-inflammatory prophylaxis compared to the last period of the study (mean difference 16.0% [95% CI −9.2% to 22.9%]; $P < 0.001$; $I^2 = 88.1\%$). However, there was no difference between the percentage of participants having a gout flare while on anti-inflammatory prophylaxis and during the last period of the study (mean difference 3.0% [95% CI −3.7% to 9.7%]; $P = 0.39$; $I^2 = 91.1\%$).

Sensitivity analyses. Sensitivity analyses are shown in Table 3. Despite considerable heterogeneity in the results, the sensitivity analyses indicated no material effects of anti-inflammatory prophylaxis duration, trial duration, ULT class, or placebo arms on the results.

DISCUSSION

This meta-analysis has demonstrated that gout flares are common after ceasing anti-inflammatory prophylaxis but ultimately return to levels seen while taking prophylaxis with time. These data have important clinical implications, indicating that the period after stopping anti-inflammatory prophylaxis is a high-risk period for gout flares, and that for some individuals, longer periods of anti-inflammatory prophylaxis may be required.

The exact mechanism of the rise in gout flares after ceasing anti-inflammatory prophylaxis is unclear. For those studies with a shorter duration of anti-inflammatory prophylaxis, namely eight weeks, it is possible this may relate to failure to achieve the target serum urate level after such a short period of ULT. However, a post hoc analysis of gout flares of three Phase 3 clinical trials of febuxostat showed the rise in gout flares after the eight weeks of anti-inflammatory prophylaxis occurred irrespective of whether participants were above or below target serum urate level of 6 mg/dL.[13] A more recent post hoc analysis of the cardiovascular safety of febuxostat or allopurinol in patients with gout also reported that the rise in gout flares after the six-month period of anti-inflammatory prophylaxis was present in all serum urate level categories (≤3.9, 4.0–5.9, 6.0–7.9, 8.0–9.9, and ≥10 mg/dL).[14] Achieving meaningful reductions in gout flare is often delayed months or even years after target serum urate goals are obtained, so these results may simply reflect the natural history of the disease. However, in our original placebo-controlled trial, we did not see a similar increase in gout flares in those participants who received a placebo, suggesting other mechanisms are likely to be important.[8]

Colchicine was the most frequently administered anti-inflammatory prophylactic agent in the analyzed trials. Colchicine prevents microtubule assembly, and thereby disrupts multiple

Table 2. Duration of the three observation periods for each included study

First author	Duration of the observation period (mo)		
	Taking prophylaxis	Within 3-mo period after stopping prophylaxis	Last period of the trial
Bardin et al,[22]	2	2	2
Becker et al,[23]	0.75	0.75	0.75
Becker et al,[24]	0.5	0.5	Not available
Dalbeth et al,[25]	6	7	8
O'Dell et al,[12]	3	3	3
Stamp et al,[8]	2	3	3.5
Yamanaka et al,[27]	3	3	3.3
Median duration	2.0	2.0	2.0
Mean duration	2.0	3.1	3.1

proinflammatory pathways involved in the pathogenesis of gout flare. Specifically, colchicine inhibits inflammasome activation, NF-κB expression, inflammatory cell chemotaxis, mast cell degranulation, generation of leukotrienes and cytokines, and phagocytosis.[15] Colchicine has been demonstrated to reduce leukocyte counts in synovial fluids from people with intercritical gout despite the presence of MSU crystals.[16] Recognizing that intraarticular MSU crystals can persist for months after the target serum urate level is achieved, particularly for people with a longer history of gout,[17] it is possible that colchicine discontinuation allows for inflammatory responses generated from resident or migrating cells targeting MSU crystals.

The optimal duration of anti-inflammatory prophylaxis has not been determined. Whether a longer period of prophylaxis would reduce these "rebound flares" is uncertain. MSU crystals can remain in joints for many months after achieving the target urate level, and the prolonged administration of colchicine may best be targeted to those who are at a higher risk of flare when the colchicine is discontinued. In this regard, we have shown that those people more likely to flare after stopping colchicine have had at least one flare in the month before stopping the study drug (odds ratio [OR] 5.39, 95% CI 2.21–13.15), and serum urate ≥0.36 mmol/L at month 6 (OR 2.85, 95% CI 1.14–7.12).[18] Given gout flares are the core element of the disease from the patient

Table 3. Pooled estimates and 95% CIs for the three periods for both the primary and sensitivity analyses*

	Taking prophylaxis, percent (95% CI)	Within 3-mo period after stopping prophylaxis, percent (95% CI)	Last period of the study, percent (95%CI)	Number of studies excluded (reference)
Primary analysis	14.7 (11.3–18.5)	29.7 (22.9–37.0)	12.2 (6.8–19.0)	0
Sensitivity analyses				
Prophylaxis ≤12 wk only	14.0 (9.7–19.1)	34.3 (28.3–40.7)	8.4 (5.9–11.4)	4 ([8,12,22,28])
Prophylaxis >12 wk only	15.5 (10.3–21.6)	25.1 (14.5–37.5)	14.1 (6.6–23.9)	2 ([23,24])
Short trial ≤4 wk	10.9 (6.5–16.2)	36.8 (29.3–44.7)		6 ([8,12,22,23,27,28])
Long trial >4 wk	15.7 (11.7–20.2)	27.7 (20.0–36.2)	12.2 (6.8–19.0)	1 ([24])
Lesinurad excluded	15.1 (11.1–19.6)	32.8 (26.0–39.9)	13.9 (7.4–22.0)	1 ([22])
Placebo arms excluded	14.7 (11.2–19.2)	29.6 (22.0–37.7)	11.8 (6.0–19.1)	2 ([24,25])
All three time points available	15.4 (11.2–20.2)	28.7 (20.43–37.7)	12.2 (6.8–19.0)	

* CI, confidence interval.

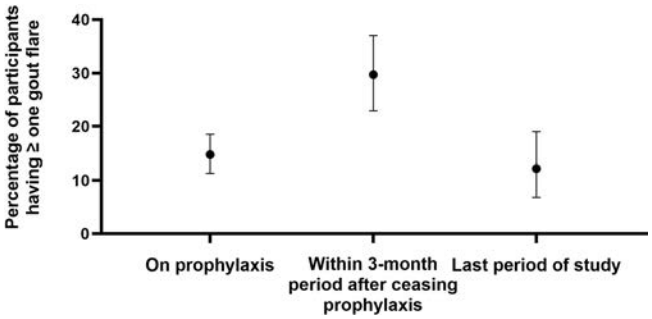

Figure 2. Pooled random effects estimate of the percentage of participants having one or more gout flares at each of the three time points.

perspective, it is likely that anything to reduce flares will be well accepted by people with gout. Further studies to determine the optimal duration of anti-inflammatory prophylaxis will be required, and it is important that future trials of new and old therapies are designed to include the period after prophylaxis ends to determine the effect on rebound flares.

Study strengths include the large sample size, the range of ULT agents administered, and consistent reporting of the gout flare outcome. Limitations include the way in which gout flares were defined and reported. Although there is a validated definition of gout flares,[19] this has not been consistently used in clinical trials, with most trials using self-reported flare requiring treatment as the definition.[20,21] The capture of gout flare data is subject to recall bias, and the reporting of gout flares is inconsistent across clinical trials, which led to many studies being excluded from this analysis and likely contributed to the large variation in flare rates across the studies.[20,21] The I^2 test indicated significant heterogeneity among study results, which did not reduce with the sensitivity analysis, indicating the heterogeneity is not explained by the study characteristics; however, the results were consistent across studies. We were unable to analyze the effects of different anti-inflammatory agents, varying duration of anti-inflammatory prophylaxis, ULT, and the postprophylaxis period. The duration of the observation period was not standardized across studies, but within each study the duration of the observation period was relatively consistent. Finally, adherence with anti-inflammatory prophylaxis was not uniformly reported, and the severity of gout flares occurring at each time period was not reported.

In conclusion, gout flares are common after stopping anti-inflammatory prophylaxis but, with time, return to levels seen during prophylaxis. People with gout should be cautioned about the risk of gout flares and have a plan for effective gout flare management in the three months after stopping anti-inflammatory prophylaxis.

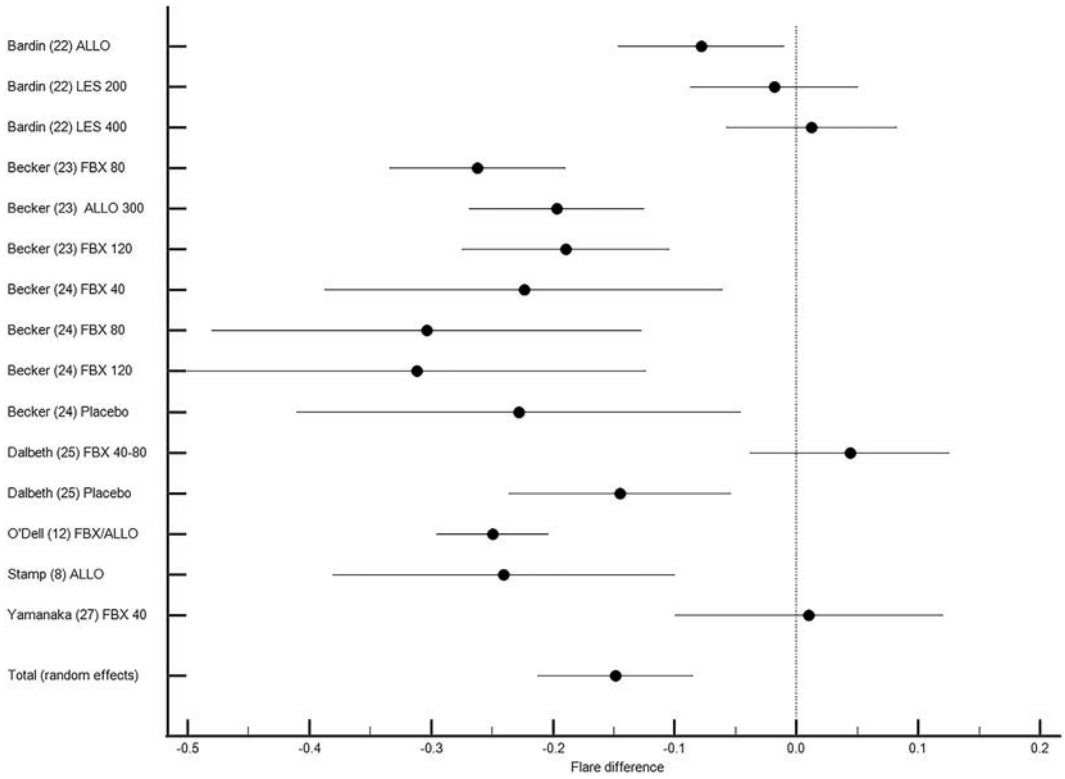

Figure 3. Forest plot of individual studies and pooled random effects showing the change in the percentage of participants having a gout flare while taking anti-inflammatory prophylaxis and soon after ceasing prophylaxis. ALLO, allopurinol; FBX, febuxostat; LES, lcaresesinurad.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Stamp confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Richette P, Doherty M, Pascual E, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis 2017;76(1):29–42.
2. Hui M, Carr A, Cameron S, et al; British Society for Rheumatology Standards, Audit and Guidelines Working Group. The British Society for Rheumatology Guideline for the Management of Gout. Rheumatology (Oxford) 2017;56(7):e1–e20.
3. FitzGerald JD, Dalbeth N, Mikuls T, et al. 2020 American College of Rheumatology Guideline for the Management of Gout. Arthritis Care Res (Hoboken) 2020;72(6):744–760.
4. Becker MA, MacDonald PA, Hunt BJ, et al. Determinants of the clinical outcomes of gout during the first year of urate-lowering therapy. Nucleosides Nucleotides Nucleic Acids 2008;27(6):585–591.
5. Stamp LK, Frampton C, Morillon MB, et al. Association between serum urate and flares in people with gout and evidence for surrogate status: a secondary analysis of two randomised controlled trials. Lancet Rheumatol 2022;4(1):e53–e60.
6. Khanna D, Khanna PP, Fitzgerald JD, et al; American College of Rheumatology. 2012 American College of Rheumatology guidelines for the management of gout. Part 2: therapy and antiinflammatory prophylaxis of acute gouty arthritis. Arthritis Care Res (Hoboken) 2012;64(10):1447–1461.
7. Latourte A, Bardin T, Richette P. Prophylaxis for acute gout flares after initiation of urate-lowering therapy. Rheumatology Oxford 2014;53(11):1920–1926.
8. Stamp L, Horne A, Mihov B, et al. Is prophylaxis required with start-low go-slow allopurinol dose escalation in gout? A randomised non-inferiority double-blind placebo-controlled trial. Ann Rheum Dis 2023;82(12):1626–1634.
9. King VJ, Stevens A, Nussbaumer-Streit B, et al. Paper 2: performing rapid reviews. Syst Rev 2022;11(1):151.
10. Klerings I, Robalino S, Booth A, et al; Cochrane Rapid Reviews Methods Group. Rapid reviews methods series: guidance on literature search. BMJ Evid Based Med 2023;28(6):412–417.
11. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177–188.
12. O'Dell JR, Brophy MT, Pillinger MH, et al. Comparative effectiveness of allopurinol and febuxostat in gout management. NEJM Evid 2022;1(3):10.1056/evidoa2100028.
13. Wortmann RL, MacDonald PA, Hunt B, et al. Effect of prophylaxis on gout flares after the initiation of urate-lowering therapy: analysis of data from three phase III trials. Clin Ther 2010;32(14):2386–2397.
14. Tedeschi SK, Hayashi K, Zhang Y, et al. Identifying the association between serum urate levels and gout flares in patients taking urate-lowering therapy: a post hoc cohort analysis of the CARES trial with consideration of dropout. Ann Rheum Dis 2024;83(10):1375–1380.
15. Dalbeth N, Lauterio TJ, Wolfe HR. Mechanism of action of colchicine in the treatment of gout. Clin Ther 2014;36(10):1465–1479.
16. Pascual E, Castellano JA. Treatment with colchicine decreases white cell counts in synovial fluid of asymptomatic knees that contain monosodium urate crystals. J Rheumatol 1992;19(4):600–603.
17. Pascual E, Sivera F. Time required for disappearance of urate crystals from synovial fluid after successful hypouricaemic treatment relates to the duration of gout. Ann Rheum Dis 2007;66(8):1056–1058.
18. Stamp LK, Horne A, Mihov B, et al. Predicting gout flares in people starting allopurinol using the start-low go-slow dose escalation strategy. Arthritis Care Res (Hoboken) 2024;76(10):1371–1378.
19. Gaffo AL, Dalbeth N, Saag KG, et al. Brief report: validation of a definition of flare in patients with established gout. Arthritis Rheumatol 2018;70(3):462–467.
20. Stamp LK, Morillon MB, Taylor WJ, et al. Variability in the reporting of serum urate and flares in gout clinical trials: need for minimal reporting requirements. J Rheumatol 2018;45(3):419–424.
21. Stewart S, Tallon A, Taylor WJ, et al. How flare prevention outcomes are reported in gout studies: a systematic review and content analysis of randomized controlled trials. Semin Arthritis Rheum 2020;50(2):303–313.
22. Bardin T, Keenan RT, Khanna PP, et al. Lesinurad in combination with allopurinol: a randomised, double-blind, placebo-controlled study in patients with gout with inadequate response to standard of care (the multinational CLEAR 2 study). Ann Rheum Dis 2017;76(5):811–820.
23. Becker MA, Schumacher HR Jr, Wortmann RL, et al. Febuxostat compared with allopurinol in patients with hyperuricemia and gout. N Engl J Med 2005;353(23):2450–2461.
24. Becker MA, Schumacher HR Jr, Wortmann RL, et al. Febuxostat, a novel nonpurine selective inhibitor of xanthine oxidase: a twenty-eight-day, multicenter, phase II, randomized, double-blind, placebo-controlled, dose-response clinical trial examining safety and efficacy in patients with gout. Arthritis Rheum 2005;52(3):916–923.
25. Dalbeth N, Saag KG, Palmer WE, et al. Effects of febuxostat in early gout: a randomized, double-blind, placebo-controlled study. Arthritis Rheumatol 2017;69(12):2386–2395.
26. Gaffo AL, Schumacher HR, Saag KG, et al. Developing a provisional definition of flare in patients with established gout. Arthritis Rheum 2012;64(5):1508–1517.
27. Yamanaka H, Tamaki S, Ide Y, et al. Stepwise dose increase of febuxostat is comparable with colchicine prophylaxis for the prevention of gout flares during the initial phase of urate-lowering therapy: results from FORTUNE-1, a prospective, multicentre randomised study. Ann Rheum Dis 2018;77(2):270–276.
28. Dalbeth N, Saag KG, Palmer WE, et al. Effects of febuxostat in early gout: a randomized, double-blind, placebo-controlled study. Arthritis Rheumatol 2017;69(12):2386–2395.

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 727–731
DOI 10.1002/acr.25494

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Relationship Between Gout Flare States and Patient-Reported Outcomes After Allopurinol Initiation

Lisa K. Stamp,[1] Chris Frampton,[2] Sarah Stewart,[3] Keith J. Petrie,[4] N. Lawrence Edwards,[5] Angelo Gaffo,[6] and Nicola Dalbeth[7]

Objective. Gout flares are the most important clinical feature of the disease. A hypothetical maximum flare occurrence in the preceding six months has been suggested to be no flares for a patient-acceptable symptom state (PASS) and only one flare for low disease activity (LDA). The aim of this analysis was to determine the relationship between gout flare states (PASS, LDA, and not in LDA or PASS [non-LDA/PASS]) and patient-reported outcomes.

Methods. Post hoc analyses of variance were undertaken using data from a 12-month randomized controlled trial involving 172 people with gout, which compared low-dose colchicine to placebo for the first 6 months while starting allopurinol with a further 6-month follow-up. Self-reported gout flares were collected monthly. Health Assessment Questionnaire (HAQ) and EuroQol 5-domain (EQ-5D-3L) were completed at 0, 3, 6 ,9, and 12 months, and the gout-specific brief illness perception questionnaire (BIPQ) was collected at months 0, 6, and 12.

Results. In the final six months of the study, 68 participants (38%) were classified as being in PASS, 34 (19%) as in LDA, and 77 (43%) as non-LDA/PASS. There was no association between gout flare states and EQ-5D-3L or HAQ. There was a statistically significant association between three of eight BIPQ items with increasing consequences, identity, and concern scores across the three states of PASS, LDA, and non-LDA/PASS.

Conclusion. The majority of people were able to achieve gout flare PASS or LDA in the second six months after commencing allopurinol. As flare burden increases, so does the impact of gout on the patient. These findings highlight the importance of flare prevention in the management of gout.

INTRODUCTION

Gout flares are the most important clinical feature for people who have gout. Gout flares affect just about every aspect of life including physical, psychological, social and family life.[1] Prevention of gout flares is therefore a key goal of management for both health care providers and people with gout. Despite this, the majority of studies of urate-lowering therapies have used serum urate (SU) as a "surrogate" measure for gout flares.[2,3] However, the burden of gout flares is multifaceted and includes the number of flares as well as the severity of each individual flare. Defining the overall flare burden for people with gout has been challenging because of variable reporting and lack of a validated flares severity definition.

Low disease activity (LDA) has been defined as "a useful target of treatment by both physician and patient, given current treatment possibilities and limitations."[4] Patient-acceptable symptom state (PASS) has been defined as the "value beyond which the patient feels well,"[5] that is, a tolerable level of symptoms for the individual. In 2021, Taylor et al[6] recruited 512 participants who answered questions about their gout flares that would classify them into one of three gout flare states: (i) remission, defined as an affirmative response to the question, "Considering

Supported by the Health Research Council of New Zealand (grant 18/151).
[1]Lisa K. Stamp, MBChB, FRACP, PhD: University of Otago, Christchurch, Christchurch, and Health New Zealand, Te Whatu Ora Waitaha, New Zealand; [2]Chris Frampton, PhD: University of Otago, Christchurch, Christchurch, New Zealand; [3]Sarah Stewart, PhD: University of Auckland and Auckland University of Technology, Auckland, New Zealand; [4]Keith J. Petrie, PhD: University of Auckland, Auckland, New Zealand; [5]N. Lawrence Edwards, MD: University of Florida, Gainesville, Florida; [6]Angelo Gaffo, MD: University of Alabama at Birmingham and Birmingham VA Medical Center, Alabama; [7]Nicola Dalbeth, MBChB, FRACP: University of Auckland, Auckland, New Zealand, and Health New Zealand, Te Whatu Ora Te Toka Tumai, New Zealand.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25494).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25494.

Address correspondence via email to Lisa K. Stamp, MBChB, FRACP, PhD, at lisa.stamp@cdhb.health.nz

Submitted for publication December 5, 2024; accepted in revised form December 27, 2024.

SIGNIFICANCE & INNOVATIONS

- Flare prevention and treatment are critical aspects of gout management.
- As gout flare burden increases, so does the impact of gout on the patient.
- Most people commencing allopurinol are able to achieve gout flare patient-acceptable symptom state or low disease activity in the second six months of treatment.

the number of attacks (flares) that you have had over the last [6 or 12] months, do you think your gout has gone away?"; (ii) LDA, defined as a negative response to the question, "Considering the number of attacks (flares) that you have had over the last [6 or 12] months, do you think you need more or stronger treatment?"; or (iii) PASS, defined as an affirmative response to the question, "Considering the number of attacks (flares) that you have had over the last [6 or 12] months, would you say that your gout control is currently satisfactory?" Participants also reported the hypothetical maximum number of flares that they could experience over 6 and 12 months and still consider themselves to be in the associated disease activity state. Based on these data, participants in LDA reported a median (interquartile range [IQR]) of 1 (0–2) flares and those in PASS 0 (0–1) flares in a six-month period. Similar results were observed over a 12-month period, with participants in PASS reporting a median (IQR) 0 (0–2) flares and in LDA 1 (0–2) flares.[6] Whether LDA and PASS are associated with patient-reported outcomes is unknown. The aim of this analysis was to determine the relationship between gout flare states (PASS, LDA, and not in LDA or PASS [non-LDA/PASS]) and patient-reported outcomes.

METHODS

Post hoc analyses of the 12-month "Is colchicine prophylaxis required with start-low go-slow allopurinol dose escalation in gout?" noninferiority randomized controlled trial were undertaken (ACTRN 12618001179224). Detailed methods and results of the full trial have been reported previously.[7] In brief, this was a one-year double-masked placebo-controlled noninferiority trial with participants randomized 1:1 to colchicine 0.5 mg daily or placebo for the first six months. All participants were required to have at least one gout flare in the preceding six months. All participants commenced allopurinol, increasing monthly to achieve target urate <0.36 mmol/L. The starting dose of allopurinol was 50 mg daily in those with estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m^2 and 100 mg daily in those with eGFR ≥60 mL/min/1.73 m^2. Allopurinol dose was increased monthly by 50 mg daily in those with eGFR <60 mL/min/1.73 m^2 and 100 mg daily in those with eGFR ≥60 mL/min/1.73 m^2 until serum urate was <0.36 mmol/L for three consecutive visits. Ethical approval was obtained from the Health and Disability Ethics Committee, New Zealand (18/STH/156), and all participants provided written informed consent.

Participants were seen every three months by study coordinators with intervening monthly telephone assessments. Gout flares, defined as self-reported gout flares requiring treatment were recorded at each assessment. Participants were categorized into three disease burden states at month 6 and month 12 as follows: (i) PASS, no gout flares in the preceding six months; (ii) LDA, one flare in the preceding six months; and (iii) non-LDA/PASS, more than one gout flare in the preceding six months. Participants were also classified into these three disease states based on the whole 12-month study period. The Health Assessment Questionnaire (HAQ), EuroQol-5D-3L (EQ-5D-3L) questionnaire, and the gout-specific brief illness perception questionnaire (BIPQ)[8] were collected at months 0, 6, and 12.

The baseline demographics and clinical features are summarized as means or medians with SDs or IQRs and frequencies and percentages for categorical measures. No missing data were imputed. The percentages of patients in the disease state groups were compared among randomized groups within each time interval using chi-square tests. The patient-reported outcome measures were compared among the disease state groups at each time using a one-way analysis of variance. To adjust for the multiple comparisons within each time interval, the *P* values presented are calculated using the Bonferroni adjustment.

RESULTS

Of the 200 participants enrolled, there were 183 remaining in the study at month 6 and 172 at month 12. The baseline demographics and clinical features of the 200 participants are outlined in Supplementary Table 1. Of the participants, 93% were male and had a mean ± SD age of 56 ± 15.7 years. The mean ± SD duration of gout before study entry was 11.2 ± 10.1 years, and the median (IQR) number of flares in the six months before study entry was 2 (2–4).

Disease activity states. Participants changed states between differing time periods depending on the number of flares they experienced. Over the entire 12-month study period, 32 participants (17.9%) were classified as being in PASS, 25 (14.0%) as in LDA, and 122 (68.2%) as non-LDA/PASS. In the first 6 months of the study, 61 participants (31.9%) were classified as being in PASS, 37 (19.4%) as in LDA, and 93 (48.7%) as non-LDA/PASS. In the final 6 months of the study, 68 participants (38%) were classified as being in PASS, 34 (19%) as in LDA, and 77 (43%) as non-LDA/PASS (Figure 1). There was no significant difference among randomized groups with respect to the proportion of participants fulfilling each of the 3 states in either the first or last 6 months or over the entire 12-month period.

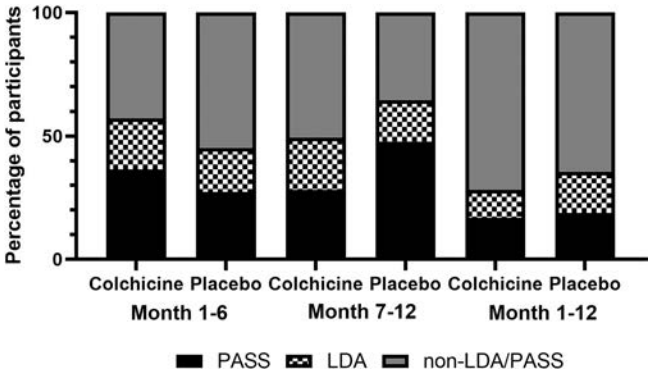

Figure 1. Percentage of participants in each disease state by randomization. LDA, low disease activity; PASS, patient-acceptable symptom state.

Association between disease activity states and patient-reported outcomes. There was no association between gout flare states and the EQ-5D-3L or HAQ (Table 1). There was a statistically significant association between three of the eight BIPQ items, namely, consequences, identity, and concern scores, with a gradient of increasing scores across the three states of PASS, LDA, and non-LDA/PASS at both months 1 to 6 and 7 to 12 (Table 1). Results were similar when PASS/LDA were combined and compared with non-LDA/PASS (Table 2).

DISCUSSION

Herein, we have shown that people with gout can achieve both gout flare LDA and PASS, but it is hard within the first year of urate-lowering therapy (ULT), and 6 months of colchicine treatment does not lead to improvements in gout flare states in the 12 months after starting ULT. Importantly, the gout flare states PASS, LDA, and non-LDA/PASS were consistently associated with three BIPQ domain scores increasing—consequences, identity, and concern—validating the impact of these gout flare states on people with gout.

Table 1. Association between PASS, LDA, and non-LDA/PASS and BIPQ items, EQ-5D-3L, and HAQ*

	PASS	LDA	Non-LDA/PASS	*P* value
Months 1–6				
Participants, n	58	35	90	
Consequences (10, severely affected)	1.4 ± 2.6	1.6 ± 2.7	3.0 ± 2.8	0.002
Timeline (10, forever)	6.2 ± 4.2	7.5 ± 3.4	7.4 ± 3.6	1.0
Personal control (10, extreme amount)	8.2 ± 2.5	8.6 ± 1.6	8.0 ± 2.2	1.0
Treatment control (10, extremely helpful)	9.0 ± 2.0	8.8 ± 2.3	8.8 ± 1.9	1.0
Identity (10, many severe symptoms)	1.7 ± 2.6	1.8 ± 1.9	3.4 ± 2.9	<0.001
Concern (10, extremely concerned)	2.1 ± 2.8	3.4 ± 2.9	4.0 ± 3.5	0.033
Understanding (10, very clearly)	7.7 ± 2.6	8.1 ± 2.1	8.2 ± 2.1	1.0
Emotional response (10, extremely affected)	1.8 ± 2.9	2.0 ± 2.5	2.7 ± 3.2	1.0
EQ-5D-3L	0.91 ± 0.17	0.95 ± 0.12	0.89 ± 0.16	0.56
HAQ	0.19 ± 0.44	0.15 ± 0.40	0.18 ± 0.41	1.0
Months 7–12				
Participants, n	66	31	75	
Consequences (10, severely affected)	1.1 ± 2.1	1.5 ± 1.9	2.7 ± 2.7	0.006
Timeline (10, forever)	7.5 ± 3.6	7.2 ± 3.6	8.3 ± 3.1	1.0
Personal control (10, extreme amount)	8.4 ± 2.5	8.7 ± 1.2	8.2 ± 2.2	1.0
Treatment control (10, extremely helpful)	8.9 ± 2.3	8.9 ± 2.0	9.2 ± 1.5	1.0
Identity (10, many severe symptoms)	0.8 ± 1.3	1.7 ± 1.9	2.8 ± 2.6	<0.001
Concern (10, extremely concerned)	1.9 ± 2.9	3.6 ± 3.4	3.8 ± 3.3	0.016
Understanding (10, very clearly)	8.4 ± 1.9	8.6 ± 1.9	8.6 ± 2.0	1.0
Emotional response (10, extremely affected)	1.4 ± 2.3	2.6 ± 2.9	1.7 ± 2.5	1.0
EQ-5D-3L	0.93 ± 0.13	0.93 ± 0.14	0.90 ± 0.16	0.98
HAQ	0.18 ± 0.38	0.12 ± 0.30	0.18 ± 0.41	1.0
Months 1–12				
Participants, n	31	23	118	
Consequences (10, severely affected)	1.1 ± 2.0	1.0 ± 1.9	2.2 ± 2.6	0.14
Timeline (10, forever)	7.6 ± 3.5	7.1 ± 3.6	8.0 ± 3.4	1.0
Personal control (10, extreme amount)	8.1 ± 2.9	8.9 ± 1.9	8.4 ± 2.1	1.0
Treatment control (10, extremely helpful)	8.3 ± 2.9	8.8 ± 2.3	9.2 ± 1.4	0.59
Identity (10, many severe symptoms)	0.5 ± 0.96	0.8 ± 1.2	2.3 ± 2.4	<0.001
Concern (10, extremely concerned)	2.1 ± 3.0	1.7 ± 2.4	3.5 ± 3.4	0.09
Understanding (10, very clearly)	8.2 ± 1.8	8.4 ± 2.2	8.6 ± 1.9	1.0
Emotional response (10, extremely affected)	1.3 ± 2.4	1.4 ± 2.4	2.0 ± 2.6	1.0
EQ-5D-3L	0.94 ± 0.13	0.97 ± 0.09	0.90 ± 0.15	0.90
HAQ	0.22 ± 0.42	0.07 ± 0.26	0.18 ± 0.39	1.0

* Data presented are mean ± SD, and *P* values are Bonferroni corrected. BIPQ, brief illness perception questionnaire; EQ-5D-3L, EuroQol 5-domain; HAQ, Health Assessment Questionnaire; LDA, low disease activity; PASS, patient-acceptable symptom state.

Table 2. Association between PASS/LDA and non-LDA/PASS and BIPQ items, EQ-5D-3L, and HAQ*

	PASS/LDA	Non-LDA/PASS	*P* value
Months 1–6			
Participants, n	93	90	
Consequences (10, severely affected)	1.5 ± 2.6	3.0 ± 2.8	0.002
Timeline (10, forever)	6.7 ± 4.0	7.4 ± 3.5	1.0
Personal control (10, extreme amount)	8.3 ± 2.2	8.0 ± 2.2	1.0
Treatment control (10, extremely helpful)	8.9 ± 2.1	8.8 ± 1.9	1.0
Identity (10, many severe symptoms)	1.7 ± 2.4	3.4 ± 2.9	<0.001
Concern (10, extremely concerned)	2.6 ± 3.0	4.0 ± 3.5	0.03
Understanding (10, very clearly)	7.9 ± 2.4	8.2 ± 2.1	1.0
Emotional response (10, extremely affected)	1.9 ± 2.7	2.7 ± 3.2	0.47
EQ-5D-3L	0.92 ± 0.15	0.89 ± 0.16	0.50
HAQ	0.17 ± 0.42	0.18 ± 0.41	1.0
Months 7–12			
Participants, n	97	75	
Consequences (10, severely affected)	1.3 ± 2.0	2.7 ± 2.7	0.001
Timeline (10, forever)	7.4 ± 3.6	8.3 ± 3.1	0.70
Personal control (10, extreme amount)	8.5 ± 2.2	8.2 ± 2.2	1.0
Treatment control (10, extremely helpful)	8.9 ± 2.2	9.2 ± 1.5	1.0
Identity (10, many severe symptoms)	1.1 ± 1.5	2.8 ± 2.6	<0.001
Concern (10, extremely concerned)	2.4 ± 3.2	3.8 ± 3.3	0.08
Understanding (10, very clearly)	8.5 ± 1.9	8.6 ± 2.0	1.0
Emotional response (10, extremely affected)	1.8 ± 2.5	1.7 ± 2.5	1.0
EQ-5D-3L	0.93 ± 0.13	0.90 ± 0.16	0.41
HAQ	0.16 ± 0.36	0.18 ± 0.41	1.0
Months 1–12			
Participants, n	54	118	
Consequences (10, severely affected)	1.1 ± 2.0	2.2 ± 2.6	0.04
Timeline (10, forever)	7.4 ± 3.5	8.0 ± 3.4	1.0
Personal control (10, extreme amount)	8.4 ± 2.5	8.4 ± 2.1	1.0
Treatment control (10, extremely helpful)	8.5 ± 2.6	9.2 ± 1.4	1.0
Identity (10, many severe symptoms)	0.7 ± 1.1	2.3 ± 2.4	<0.001
Concern (10, extremely concerned)	1.9 ± 2.8	3.5 ± 3.4	0.026
Understanding (10, very clearly)	8.3 ± 2.0	8.6 ± 1.9	1.0
Emotional response (10, extremely affected)	1.3 ± 2.4	2.0 ± 2.6	1.0
EQ-5D-3L	0.95 ± 0.12	0.90 ± 0.15	0.11
HAQ	0.16 ± 0.37	0.18 ± 0.39	1.0

* *P* values are Bonferroni corrected. BIPQ, brief illness perception questionnaire; EQ-5D-3L, EuroQol 5-domain; HAQ, Health Assessment Questionnaire; LDA, low disease activity; PASS, patient-acceptable symptom state.

Of note, the HAQ-II and EQ-5D-3L were not associated with gout flare states. Although activity limitation is recognized as an important outcome in gout studies, it has been noted that the HAQ-II has significant floor effects, which limits its ability to differentiate the spectrum of disability in people with gout.[9,10] Neither the HAQ nor EQ-5D-3L are specific for gout and may reflect the impact of comorbidities that are commonly associated with gout. Previous studies have reported an association between the number of gout flares and the gout-related Health-Related Quality of Life Gout Impact Scale, but not the HAQ-Disability Index.[11] In another study, participants with inadequately controlled gout (defined as SU >0.36 mmol/L or ≥2 flares in the previous 12 months) had worse health-related quality of life as measured by EQ-5D-3L compared with those with adequately controlled gout (defined as SU ≤0.36 mmol/L and 0 flares in the previous 12 months) (EQ-5D-3L 0.790 vs 0.877; difference –0.087; *P* <0.001).[12] However, we were unable to show a similar association with the different gout flare states.

It is important to note that the study population was 93% male. Although this reflects a typical gout trial population, women often experience higher disease severity, have more negative illness perceptions, and experience higher impact on daily activities.[13,14] Although the number of women in our study were too small to enable analysis by gender, we would expect the observed effects to be even more pronounced in women.

These data contribute to our understanding of the impact and burden of gout flares in people with gout, highlighting their concern about this core clinical manifestation of the disease. In the long term, excellent serum urate control is important. However, it is also essential that health care professionals support people with gout to prevent and manage flares as a core part of gout management.

Strengths of this study include the standardized study protocols, prospective gout flare event ascertainment, and use of outcomes of relevance to patients. Limitations include some loss to follow-up, albeit minimal; the use of subjective assessments; and

the short study design, which did not allow for assessment beyond one year. It is well recognized that gout flares can paradoxically increase after starting ULT. Given this peak in flares after starting ULT, it is not surprising that LDA and PASS are hard to achieve in the first six months of starting ULT even with anti-inflammatory prophylaxis with colchicine. Over time, there is a gradual reduction in gout flares such that, by the second year of ULT, if target serum urate is achieved, flares may cease altogether or occur less frequently.[3] Thus, it is likely that longer trials beyond 12 months are required to see the full effect of ULT on achievement of the disease activity states. Finally, we did not asked participants in the study whether they considered that their flare frequency aligned with the ascribed disease state.

The majority of people (57%) were able to achieve PASS or LDA in the second six months after commencing ULT. As the flare burden increases, so does the impact of gout on the patient. These findings highlight the importance of flare prevention in the management of gout.

ACKNOWLEDGMENTS

Open access publishing facilitated by University of Otago, as part of the Wiley - University of Otago agreement via the Council of Australian University Librarians.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Stamp confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Stewart S, Guillen AG, Taylor WJ, et al. The experience of a gout flare: a meta-synthesis of qualitative studies. Semin Arthritis Rheum 2020; 50(4):805–811.
2. Stamp L, Morillon MB, Taylor WJ, et al. Serum urate as surrogate endpoint for flares in people with gout: a systematic review and meta-regression analysis. Semin Arthritis Rheum 2018;48(2):293–301.
3. Stamp LK, Frampton C, Morillon MB, et al. Association between serum urate and flares in people with gout and evidence for surrogate status: a secondary analysis of two randomised controlled trials. Lancet Rheumatol 2022;4:e53–e60.
4. Wells G, Boers M, Tugwell P; MDA Working Group. Low disease activity state in rheumatoid arthritis: concepts and derivation of minimal disease activity. Clin Exp Rheumatol 2006;24(6 Suppl 43): S52–S59.
5. Tubach F, Wells GA, Ravaud P, et al. Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues. J Rheumatol 2005;32(10):2025–2029.
6. Taylor W, Dalbeth N, Saag KG, et al. Flare rate thresholds for patient assessment of disease activity states in gout. J Rheumatol 2021; 48(2):293–298.
7. Stamp L, Horne A, Mihov B, et al. Is prophylaxis required with start-low go-slow allopurinol dose escalation in gout? A randomised non-inferiority double-blind placebo-controlled trial. Ann Rheum Dis 2023;82(12):1626–1634.
8. Broadbent E, Petrie K, Main J, et al. The brief illness perception questionnaire. J Psychosom Res 2006;60(6):631–637.
9. Chandratre P, Roddy E, Clarson L, et al. Health-related quality of life in gout: a systematic review. Rheumatology (Oxford) 2013;52(11):2031–2040.
10. Ten Klooster PM, Oude Voshaar MAH, Taal E, et al. Comparison of measures of functional disability in patients with gout. Rheumatology (Oxford) 2011;50(4):709–713.
11. Watson L, Belcher J, Nicholls E, et al. Factors associated with change in health-related quality of life in people with gout: a 3-year prospective cohort study in primary care. Rheumatology (Oxford) 2023;62(8): 2748–2756.
12. Wood R, Fermer S, Ramachandran S, et al. Patients with gout treated with conventional urate-lowering therapy: association with disease control, health-related quality of life, and work productivity. J Rheumatol 2016;43(10):1897–1903.
13. Bergsten U, Dehlin M, Klingberg E, et al. Gender differences in illness perceptions and disease management in patients with gout, results from a questionnaire study in Western Sweden. BMC Musculoskelet Disord 2023;24(1):300.
14. Singh JA. Racial and gender disparities among patients with gout. Curr Rheumatol Rep 2013;15(2):307.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 732–743
DOI 10.1002/acr.25489

Movement Pattern Biofeedback Training After Total Knee Arthroplasty: A Randomized Controlled Trial

Michael J. Bade,[1] Cory L. Christiansen,[1] Joseph A. Zeni Jr,[2] Michael R. Dayton,[3] Jeri E. Forster,[4] Victor A. Cheuy,[5] Jesse C. Christensen,[6] Craig Hogan,[3] Ryan Koonce,[3] Doug Dennis,[7] Amy Peters,[1] and Jennifer E. Stevens-Lapsley[1]

Objective. Habitual movement compensations, such as decreased surgical peak knee extension moments (pKEM), persist years after total knee arthroplasty (TKA), are linked to poorer recovery, and may influence contralateral osteoarthritis progression. The purpose of this randomized clinical trial was to determine if a movement training program (MOVE) improves movement quality and recovery after TKA compared to a standardized rehabilitation program without movement training (CONTROL).

Methods. One hundred thirty-eight individuals were randomized to either MOVE or CONTROL groups after TKA. Participants were assessed preoperatively, 10 weeks after (end of intervention), and six months after (primary endpoint) TKA. Outcomes assessed were pKEM during walking, six-minute walk test, stair climb test, 30-second sit to stand test (30STS), timed up and go test (TUG), physical activity level, strength, range of motion, and self-reported outcomes.

Results. At six months, there were no between-group differences in surgical pKEM during walking (primary outcome). The MOVE group exhibited less contralateral pKEM compared to CONTROL during self-selected gait speed (d = 0.44, P = 0.01). CONTROL performed better on TUG and 30STS at 10 weeks ($P < 0.05$), but differences attenuated at six months.

Conclusion. The MOVE intervention did not lead to improved surgical pKEM during walking after TKA compared to CONTROL. However, the MOVE group did demonstrate less contralateral pKEM during walking. The CONTROL group demonstrated faster recovery on the TUG and 30STS, but it is unknown if this is due to improved recovery in the surgical knee or increased movement compensation relying on contralateral knee function.

INTRODUCTION

Total knee arthroplasty (TKA) is the most commonly performed surgical procedure in older adults in the United States, with projections of 3.5 million performed annually by 2040.[1,2] Although a majority of patients report improved pain and self-reported function after unilateral TKA, movement compensations developed before surgery often persist years after surgery.[3–8] These compensatory movement patterns, observed during walking and other weight-bearing tasks, are characterized by disuse of the surgical limb, resulting in decreased weight-bearing and smaller knee extension moments on the surgical limb compared to the contralateral limb. Compensatory movement patterns following unilateral TKA are associated with persistent quadriceps weakness and poor physical function.[5–8] In addition, they may also be linked to the progression of knee osteoarthritis (OA) and

ClinicalTrials.gov identifier NCT03325062.

Supported by the NIH (grants R01-AG-056585, F32-AG-066274, UL1-TR-001082, and KL2-TR-001870); the Eastern Colorado VA Health Care System Geriatric Research Education and Clinical Center (GRECC), and an Academy of Orthopaedic Physical Therapy Career Development Award.

[1]Michael J. Bade, PT, PhD, Cory L. Christiansen, PT, PhD, Amy Peters, MS, Jennifer E. Stevens-Lapsley, PT, PhD: University of Colorado Anschutz Medical Campus and Eastern Colorado VA Health Care System, Geriatric Research Education and Clinical Center (GRECC), Aurora, Colorado; [2]Joseph A. Zeni Jr, PT, PhD: Rutgers University, Newark, New Jersey; [3]Michael R. Dayton, MD, Craig Hogan, MD, Ryan Koonce, MD: University of Colorado Anschutz Medical Campus, Aurora; [4]Jeri E. Forster, PhD: University of Colorado Anschutz Medical Campus and VA Rocky Mountain Mental Illness, Research, Education and Clinical Center, VA Eastern Colorado Health Care System, Aurora; [5]Victor A. Cheuy, PhD: University of California, San Francisco; [6]Jesse C. Christensen, PT, PhD: VA Salt Lake City Healthcare System, Salt Lake City, Utah; [7]Doug Dennis, MD: Colorado Joint Replacement, Advent Health Porter, Denver, Colorado; Department of Mechanical and Materials Engineering, Ritchie School of Engineering and Computer Science, University of Denver, Denver, Colorado; Department of Orthopedics, School of Medicine, University of Colorado Anschutz Medial Campus, Aurora; and Department of Biomedical Engineering, School of Engineering, University of Tennessee, Knoxville.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25489).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25489.

[Correction added on 3 February 2025, after first online publication: affiliation of Doug Dennis has been corrected.]

Address correspondence via email to Michael J. Bade, PT, PhD, at michael.bade@cuanschutz.edu.

Submitted for publication March 11, 2024; accepted in revised form December 17, 2024.

SIGNIFICANCE & INNOVATIONS

- This randomized clinical trial investigated the effects of a novel movement training program (MOVE) that used real-time biofeedback delivered via a smart phone application from wearable insole sensors. The MOVE intervention focused on movement pattern training during dynamic functional tasks in both the clinical and home settings after total knee arthroplasty (TKA).
- The MOVE intervention did not lead to improved surgical knee extension moments during walking after TKA compared to a program without movement training (CONTROL). However, the MOVE group exhibited lower peak knee extension moments on the contralateral knee during walking after intervention compared to the CONTROL group, which may influence contralateral osteoarthritis (OA) progression.
- Future research will examine additional biomechanical tasks beyond walking and the long-term effects of the MOVE intervention on movement quality and contralateral OA progression.

subsequent need for TKA in the contralateral limb, by creating compensatory, increased loading on the contralateral knee.[3,9–14] Thus, failure to adequately remediate movement compensations after unilateral TKA may lead to long-term weakness, poor physical function, and increased risk for contralateral OA progression.

Traditionally, rehabilitation protocols after TKA have focused primarily on remediation of impairments in pain, range of motion, and strength but do not specifically include movement pattern training to address compensatory movement strategies.[15] Preliminary research has demonstrated that individuals who received movement pattern training, using weight-bearing biofeedback combined with therapist cuing to reduce movement compensation, improved physical function and strength recovery to a greater degree than those who received only traditional rehabilitation without movement pattern training.[16,17] Additionally, movement pattern training groups demonstrated greater increases in surgical peak knee extension moments (pKEM) during walking and rising from a chair, indicating improved use of the surgical knee during common daily functional tasks. Notably, improvements in the movement pattern training group were even stronger at six months (4.5 months after intervention), suggesting a lasting change in movement control may have occurred.

Prior movement pattern interventions have been delivered through the use of bilateral force plates that offer biofeedback on limb symmetry during the performance of functional tasks or games.[16,17] Although these interventions allow therapists to retrain movement patterns and encourage symmetrical limb use, they are limited in their scope of use. Force plates require constrained foot positions and do not easily allow for biofeedback on more dynamic tasks, such as walking, or allow for feedback outside of the laboratory/clinical setting. In recent years, advances in technology have led to the development of insole force sensors capable of giving real-time biofeedback via a smart phone application that enable movement pattern training during dynamic functional tasks in both laboratory and free-living conditions.[18,19] However, to date, no randomized controlled trials have evaluated the efficacy of a movement pattern training intervention after TKA using more modern biofeedback insole sensor systems.

Therefore, the purpose of this randomized clinical trial was to determine if a novel movement pattern training program (MOVE) improved movement pattern quality more than a standardized rehabilitation program without movement training (CONTROL). The secondary goal was to determine if movement pattern training improved quadriceps strength and physical function and lessened contralateral OA progression. Our hypothesis was that the MOVE group would demonstrate improved biomechanical outcomes, quadriceps strength, and physical function while reducing the incidence of contralateral OA progression. This paper reports the results for the primary trial end point (six months after TKA).

PATIENTS AND METHODS

Study design. This was a two-arm, parallel, randomized controlled trial to determine if the addition of MOVE to standard rehabilitation improved movement pattern quality more than standard rehabilitation without movement training (CONTROL). This study was prospectively registered on clinicaltrials.gov (NCT03325062), and the protocol for this trial has been previously published.[20]

Participants. Participants were consecutively recruited by 13 participating orthopedic surgeons from four institutions in the Denver, Colorado, metropolitan area from January 15, 2018, to November 29, 2022. Participants were included in the study if they were between 50 and 85 years of age and were scheduled for a primary unilateral TKA for end-stage OA. Exclusion criteria were severe contralateral knee OA (>4/10 pain on a verbal pain rating or Kellgren-Lawrence grade >3 determined by radiographs), previous contralateral TKA, comorbid condition that substantially limits physical function (eg, neurologic, cardiovascular problems, or unstable orthopedic condition other than knee OA), alternate plan of care for rehabilitation after surgery (eg, home health, skilled nursing, etc), uncontrolled diabetes (hemoglobin A1c >8.0), body mass index >40, current smoker, drug abuse, surgical complications necessitating an altered course of rehabilitation, unable to safely walk 30 m without an assistive device, or contraindications to magnetic resonance imaging. Informed consent was obtained from all participants, and this study was approved by the Colorado Multiple Institutional Review Board.

Randomization and anonymizing. Following surgery, each participant was randomized to a treatment arm (MOVE or

CONTROL) with a 1:1 allocation ratio. Randomization was conducted using a computer-generated allocation table with stratification on sex and random block sizes of two and four within each stratum by an unmasked study coordinator (MJB). Anonymized study personnel consisted of the principal investigator (JESL) and outcome assessors. Participants were only aware that they were participating in one of two potential rehabilitation programs, and they were instructed to not discuss details of their intervention with anonymized personnel. Due to the nature of the intervention, anonymizing of the physical therapists delivering the intervention was not possible.

Intervention. All patients received a unilateral tricompartmental TKA. The MOVE and CONTROL interventions were initiated at a mean ± SD of 5 ± 2 days after surgery. Participants were seen in outpatient rehabilitation twice per week for weeks 1 through 4, once per week for weeks 5 through 8, once at week 10, and for a booster visit at four months, for a total of 14 visits.[21] Participants were treated at one of five outpatient clinics throughout the Denver metropolitan area by 1 of 16 licensed physical therapists who were trained to deliver either the MOVE or CONTROL intervention. Fidelity to the intervention was assessed using published criteria.[20] Due to clinic closures or potential exposure to COVID-19, four participants from MOVE were seen via telehealth for a mean ± SD of 4 ± 3 visits. However, there were no alterations to the intervention protocol.

Both groups participated in a standardized rehabilitation program consisting of the following: activity-based exercises (eg, walking, rising from a chair, stair climbing), range of motion and flexibility exercises, strengthening exercises, balance training, and home exercises (based on clinic exercises). Clinic-based and home exercises were dosed at the same frequency and intensity based upon prespecified tolerance and progression criteria. Details of the intervention and criteria have been previously reported and additional details are reported in Supplementary File 1.[20] Participants in the MOVE group received the movement pattern training intervention (detailed below) during the performance of activity-based exercises, whereas the CONTROL group performed the same exercises without feedback.

MOVE group. The MOVE intervention was a movement pattern training program that focused on promoting surgical knee use during the performance of activity-based exercises with an emphasis on symmetry of loading between surgical and contralateral limbs. Participants in the MOVE group used the Loadsol (Novel.de) insoles and associated iOS application to receive real-time visual biofeedback on weight-bearing force symmetry during activity performance both during clinic sessions and at home during the performance of their home exercise program.[18,19,22] This amount of biofeedback provided a high number of repetitions during activity-specific practice to enhance motor learning (eg, 1,960 sit to stands prescribed over the course of the intervention).[23,24] Clinicians also used verbal, visual, auditory, and tactile cues during clinic sessions to facilitate symmetrical movement based upon the optimal strategy (internal or external cues) for each patient.[25] During each clinical session, participants were assessed on retention of motor learning from the previous session on their current activities to determine if they were ready for progression.[26] Once participants were able to complete a given task with less than 5% between-limb loading asymmetry and no movement compensations, task difficulty was progressed based upon tolerance or to a difficulty level at which the participant was unable to correctly preform the task without biofeedback. Frequency of biofeedback was faded to 50% using an intermittent biofeedback schedule to promote retention of movement patterns when participants demonstrated an improvement in movement quality in response to cuing and feedback or when tolerance limited progression (eg, unable to go deeper during squats).[27] Participants were instructed to use intermittent biofeedback throughout the day and remember their movement cues when performing these activities as a part of daily living to encourage the development of an internal representation of the movement pattern.[28]

CONTROL group. The CONTROL group intervention focused on the same exercise protocol as the MOVE program, although the physical therapists did not provide any feedback on exercise performance other than minimal cues for instruction and safety. Progression of exercise was not based upon movement quality but rather only based upon safety and tolerance of exercise. If a participant directly questioned the therapist regarding movement quality, participants were encouraged to move "naturally" or "comfortably."

Outcomes. All outcomes were assessed one to two weeks preoperatively (baseline), at the end of intervention (10 weeks), and at six months (primary endpoint) at the University of Colorado Anschutz Medical Campus. Testing details and methods for each outcome have been previously reported and additional details on testing can be found in Supplementary File 1.[20] Due to testing delays caused by the COVID-19 pandemic, two participants from MOVE and two participants from CONTROL were assessed a mean of 12 weeks and 19.5 weeks after the six-month timepoint, respectively. Additionally, four participants (one in MOVE, three in CONROL) had functional performance data on the 30-second sit to stand test (30STS) and timed up and go test (TUG) captured at 10 weeks using a telehealth interface. Video-based assessments of the TUG and 30STS have been shown to be valid and reliable in individuals with knee OA.[29]

Primary outcome. The primary outcome for this study was change in pKEM in the surgical limb from baseline to six months after TKA during walking at a fixed gait speed (FGS) of 1.0 m/second. The six-month timepoint was chosen as the primary endpoint based on preliminary data indicating response to

movement pattern training becomes more pronounced by this timepoint as opposed to immediately after intervention.[16] Biomechanical testing was conducted using an eight-camera motion capture system (Vicon Motion Systems) and two force plates (Bertec Corporation) embedded in an overground 12-m walkway. Reflective markers were placed on anatomic landmarks of the lower limbs, trunk, and upper extremities according to a modified Helen-Hayes marker set.[30] Force and marker position were sampled at 2,000 and 100 Hz and filtered using a fourth-order low-pass Butterworth filter with cutoff frequencies of 20 and 6 Hz, respectively. Before measurement, participants walked five times up or down a 10-m walkway at their customary, comfortable pace. During these warm-up trials, gait speed was measured using ceiling-mounted motion sensors as participants passed through a 5-m section in the center of the walkway. The mean of the five speeds was considered their self-selected gait speed. The subsequent recorded trials were used only if they were within ±5% of the determined self-selected speed. Participants were asked to walk at a set speed of 1.0 m/second (±5%), and five acceptable trials were collected.

Demographics. Participants' age and sex were gathered from their medical records. Race and ethnicity were ascertained by participant self-report using the National Institutes of Health race and ethnicity categories and included a response of "prefer not to answer."

Secondary outcomes. Secondary biomechanical outcome measures included pKEM during walking at the participant's self-selected gait speed in addition to pKEM in the contralateral knee during FGS and self-selected gait speed. Secondary measures of physical performance, impairments, and self-reported outcomes included a stair climb test (SCT),[31] 30STS,[32] six-minute walk test,[33] TUG,[34] average steps per day using accelerometry (Actigraph Corp), isometric quadriceps strength (measured by an electromechanical dynamometer),[35] knee range of motion,[36] Western Ontario and McMaster Universities Osteoarthritis Index,[37] and Veteran's RAND 12-item health survey.[38]

Treatment outcomes. Weekly home exercise logs, averaged over the course of the intervention for activity-based exercise interventions, and number of clinical sessions attended were used to quantify adherence to the intervention. Minutes of insole use at home was also tracked for the MOVE group as a measure of intervention dose. Timing of the initiation of higher-level activity-based exercises (lunges and step/stair training) was tracked as a measure of exercise progression rates between both groups. Finally, satisfaction with the intervention was assessed using a five-point Likert scale, and adverse events rates during the intervention were tracked and reported in Supplementary Table 1.

Sample size. Statistical power was estimated using effect sizes based on a previous pilot clinical trial.[16] The observed mean change ± SD from baseline to six months for surgical pKEM during fixed-speed walking at 1.0 m/second was 0.05 ± 0.24 Nm/kg in the experimental group vs –0.16 ± 0.30 Nm/kg in the control group. A sample size of 120 participants (60 per group) would have more than 95% power to detect this difference using a two-sided, independent samples t-test with $\alpha = 0.05$. We enrolled 138 participants to allow for loss to follow-up.

Statistical analyses. The primary analysis was an intent-to-treat comparison of differences between treatment groups in surgical limb pKEM change from baseline to six months after TKA, during fixed-speed walking. Statistical inference regarding the difference between treatment groups was based on the estimated coefficient for a treatment group indicator variable in an analysis of covariance model with the change from baseline in pKEM at six months as the response variable and additional covariates that included sex (stratification variable) and the baseline value of pKEM to improve the precision of the estimated treatment differences. The conclusion about between-group differences was determined by this single statistical test to protect against an elevated risk of false positive conclusions. The adjusted estimated difference between groups in change from baseline and its associated 95% confidence interval (CI) is reported. Secondary outcomes at six months were analyzed as described above and evaluated for consistency with the primary outcome. Cohen's d effect sizes were calculated for measures with significant results and interpreted as follows: small (d = 0.2), moderate (d = 0.5), and large (d = 0.8).[39] Measures at 10 weeks were evaluated using a repeated measures analysis of variance (RM-ANOVA) to account for the correlation between repeated observations on a participant and controlled for sex. Linear contrasts were used to estimate within and between-group differences in change over time. The six-month outcomes were also assessed using the RM-ANOVA approach to determine consistency in results when all available data were used. A sensitivity analysis was conducted for the outcomes of the 30STS and TUG at 10 weeks as four participants were evaluated via telehealth due to the COVID-19 pandemic. The 10-week analyses were rerun removing these participants. Differences in satisfaction with rehabilitation programs between groups were evaluated using a Cochran-Mantel-Haenszel test, controlling for sex. Differences in adverse event rates were evaluated using a chi-squared test. Differences in clinic session attendance were determined by an independent samples t-test. A sensitivity analysis on home exercise compliance was conducted on the primary outcome using cutoffs of 80% and 90% adherence. The association between insole wear time and change in six-month primary outcome was evaluated by adding insole wear time to the primary analysis for MOVE participants. A two-sided $P = 0.05$ without adjustment for multiple comparisons was designated a priori for statistical significance, and all analyses were run in SAS v9.4.

RESULTS

Participant flow and characteristics. A total of 3,938 individuals were assessed for eligibility in the study, and 138 individuals (mean ± SD age 64.2 ± 7.2 years, 82 women) were enrolled and randomized to either the MOVE (n = 68) or CONTROL (n = 70) groups (Figure 1). Baseline characteristics of enrolled participants were similar between groups (Table 1). At the end of intervention (10 weeks), 94% of the MOVE group and 99% of the CONTROL group completed testing. At the

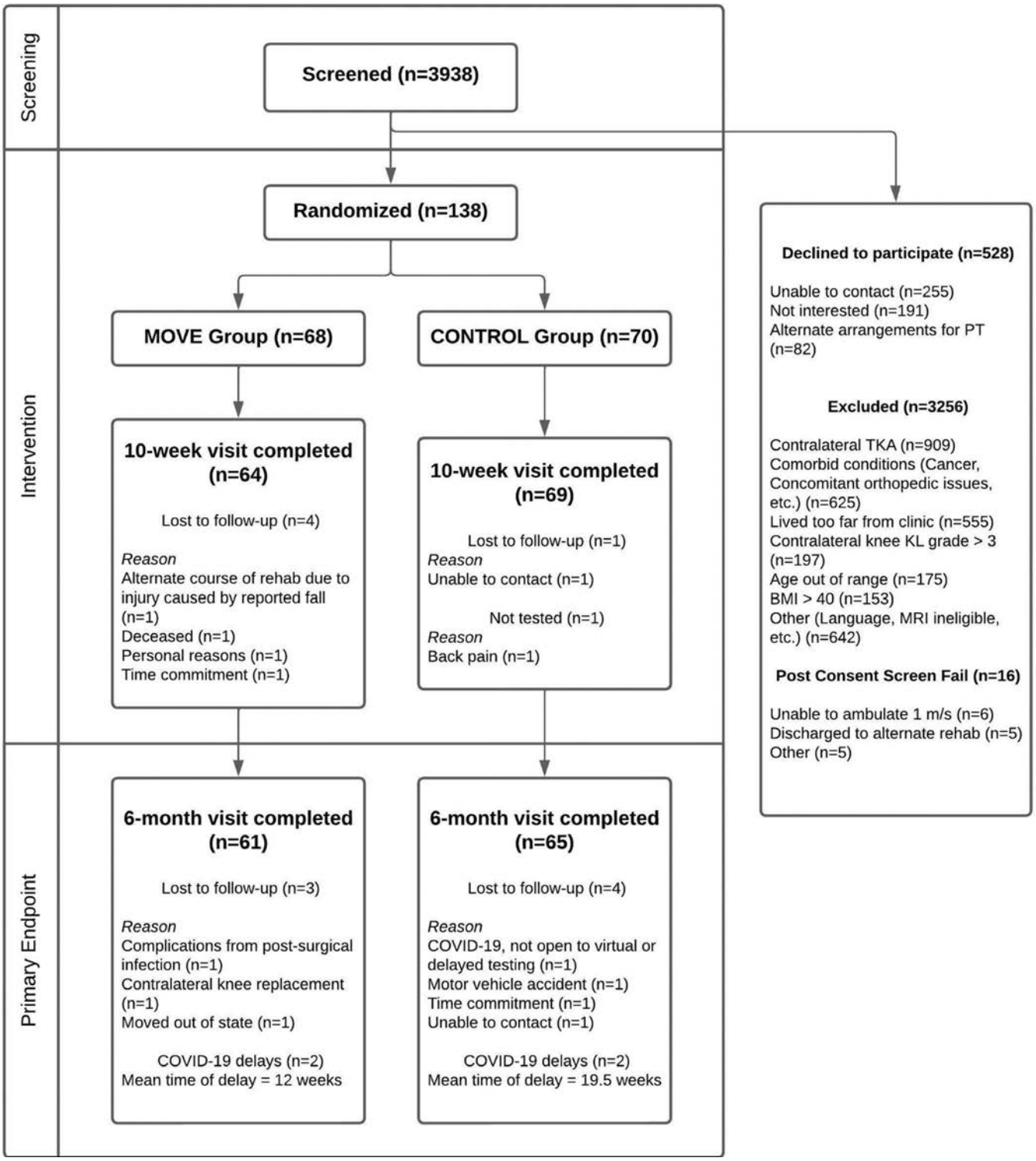

Figure 1. Consolidated Standards of Reporting Trials flow diagram. Abbreviations: BMI, body mass index; KL, Kellgren-Lawrence; MOVE; movement pattern training group; MRI, magnetic resonance imaging; PT, physical therapy; TKA, total knee arthroplasty.

Table 1. Baseline characteristics of participants*

Characteristics	Overall (n = 138)	MOVE (n = 68)	CONTROL (n = 70)	*P* value
Age, mean ± SD	64.2 ± 7.2	64.4 ± 7.5	64.1 ± 6.9	0.86
Sex, n (%)				
Female	82 (59.4)	41 (60.3)	41 (58.6)	0.84
Male	56 (40.6)	27 (39.7)	29 (41.4)	
BMI, mean ± SD	28.72 ± 4.88	28.72 ± 4.76	28.72 ± 5.04	>0.99
Race, n (%)				
Asian	2 (1.5)	1 (1.5)	1 (1.4)	0.95[a]
Black/African American	4 (2.9)	3 (4.4)	1 (1.4)	
Native Hawaiian/Other Pacific Islander	1 (0.72)	0 (0)	1 (1.4)	
White	122 (88.4)	60 (88.2)	62 (88.6)	
More than one race	7 (5.1)	2 (2.9)	5 (7.1)	
Choose not to answer	2 (1.5)	2 (2.9)	0 (0)	
Ethnicity,[b] n (%)				
Hispanic or Latino	10 (7.25)	4 (5.9)	6 (8.6)	0.83
Not Hispanic or Latino	118 (85.5)	59 (86.8)	59 (84.3)	
Choose not to answer/missing	10 (7.25)	5 (7.4)	5 (7.1)	
Functional Comorbidity Index,[b] mean ± SD	2.49 (1.5)	2.40 (1.4)	2.58 (1.6)	0.51
KL grade contralateral knee, mean ± SD	2.1 (0.84)	2.1 (0.84)	2.1 (0.86)	0.86

* BMI, body mass index; KL, Kellgren-Lawrence; MOVE, movement pattern training group.
[a] Comparison based on White/non-White.
[b] n = 137, 1 observation missing from the control group.

Table 2. Outcome measures over time by group*

Measure	MOVE			CONTROL		
	Preop, mean (SD); n	10 wk, mean (SD); n	6 mo, mean (SD); n	Preop, mean (SD); n	10 wk, mean (SD); n	6 mo, mean (SD); n
Peak knee extension moment, Nm/kgm						
Surgical FGS	0.22 (0.14); 68	0.17 (0.12); 62	0.20 (0.12); 61	0.23 (0.12); 70	0.22 (0.12); 63	0.22 (0.11); 65
Contralateral FGS	0.28 (0.10); 68	0.26 (0.10); 62	0.25 (0.09); 61	0.29 (0.13); 70	0.26 (0.11); 63	0.29 (0.12); 65
Surgical SSG	0.27 (0.17); 68	0.22 (0.14); 62	0.27 (0.16); 61	0.27 (0.13); 70	0.24 (0.13); 62	0.29 (0.13); 65
Contralateral SSG	0.36 (0.14); 68	0.34 (0.13); 62	0.35 (0.14); 61	0.35 (0.15); 70	0.33 (0.14); 62	0.40 (0.13); 65
Self-selected gait speed, m/s	1.17 (0.22); 68	1.14 (0.18); 63	1.23 (0.17); 61	1.16 (0.20); 70	1.15 (0.16); 67	1.24 (0.17); 65
6MW, m	475 (101); 68	471(83.2); 62	507 (84.2); 61	475; 70 (77.6)	479 (71.4); 63	516 (76.8); 64
SCT, s	8.45 (3.1); 58	8.44 (3.5); 53	7.19 (2.2); 51	7.97 (2.5); 56	7.73 (1.7); 59	6.47 (1.3); 52
30STS, repetitions	11.0 (2.8); 67	11.5 (3.2); 61	13.0 (3.6); 59	10.5 (3.1); 70	12.0 (3.0); 67	13.1 (3.1); 62
TUG, s	8.04 (1.7); 68	8.29 (1.6); 63	7.49 (1.3); 61	8.18 (1.7); 70	7.87 (1.6); 67	7.21 (1.3); 64
Total average step count, steps/day	5,534 (2,303); 49	5,537 (2,250); 54	6,297 (2,401); 54	5,125 (1,747); 56	4,989 (1,774); 59	5,294 (1,767); 52
Quadriceps strength, Nm/kg						
Surgical limb	1.19 (0.41); 68	0.988 (0.38); 61	1.29 (0.46); 61	1.15 (0.44); 70	1.06 (0.39); 63	1.31 (0.44); 64
Contralateral limb	1.50 (0.49); 68	1.56 (0.49); 62	1.58 (0.53); 61	1.57 (0.52); 70	1.62 (0.47); 63	1.71 (0.51); 64
Active knee ROM surgical limb, degrees						
Extension	−2.8 (3.6); 68	−2.9 (3.3); 62	−1.4 (3.3); 61	−4.0 (4.9); 70	−3.5 (4.0); 64	−2.4 (3.2); 64
Flexion	124.5 (10.2); 68	115.6 (9.9); 62	121.0 (10.1); 61	121.0 (12.2); 70	113.7 (11.5); 64	120.7 (9.6); 64
WOMAC surgical knee, points	34.5 (16.5); 68	22.6 (12.3); 62	12.3 (9.3); 60	35.2 (16.4); 69	22.9 (14.3); 67	16.0 (16.3); 65
VR-12, points						
Physical component score	36.9 (9.8); 68	42.5 (7.5); 63	48.7 (7.1); 60	37.2 (8.8); 69	41.5 (9.4); 68	48.2 (8.5); 65
Mental component score	53.6 (8.7); 68	52.8 (7.0); 63	54.2 (6.0); 60	54.3 (6.9); 69	52.1 (9.8); 68	53.1 (8.0); 65

* Secondary to the COVID-19 pandemic, four participants in the CONTROL group and two participants in the MOVE group were assessed using video for the 6MWT, TUG, and 30STS test. 30STS, 30-second sit to stand test; 6MW, six-minute walk test; FGS, fixed gait speed; MOVE, movement pattern training group; preop, preoperative; ROM, range of motion; SCT, stair climb test; SSG, self-selected gait speed; TUG, times up and go test; VR-12, Veteran's RAND 12-item health survey; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

Table 3. Estimated change and difference in change to six months*

Measure	MOVE n	MOVE Estimated change (95% CI)	CONTROL n	CONTROL Estimated change (95% CI)	Estimated difference in change (95% CI)	*P* value
Peak knee extension moment (Nm/kgm)						
Surgical FGS	61	−0.02 (−0.04 to 0.01)	65	−0.01 (−0.03 to 0.02)	0.01 (−0.03 to 0.05)	0.61
Contralateral FGS	61	−0.03 (−0.05 to −0.01)	65	0 (−0.02 to 0.02)	0.03 (0 to 0.05)	**0.045**
Surgical SSG	61	0.02 (−0.01 to 0.05)	65	0.02 (−0.01 to 0.05)	0.01 (−0.04 to 0.05)	0.84
Contralateral SSG	61	0 (−0.02 to 0.02)	65	0.04 (0.02 to 0.07)	0.04 (0.01 to 0.08)	**0.006**
Self-selected gait speed, m/s	61	0.07 (0.03 to 0.11)	64	0.08 (0.05 to 0.12)	0.01 (−0.04 to 0.06)	0.67
6MW, m	61	28.6 (14.7 to 42.5)	64	37.5 (24.0 to 50.9)	8.87 (−10.3 to 28.0)	0.36
SCT, s	47	−1.20 (−1.62 to −0.77)	44	−1.82 (−2.25 to −1.38)	−0.62 (−1.23 to −0.01)	**0.046**
30STS, repetitions	58	2.31 (1.66 to 2.95)	62	2.51 (1.89 to 3.12)	0.20 (−0.69 to 1.09)	0.66
TUG, s	61	−0.69 (−0.97 to −0.41)	64	−0.95 (−1.2 to −0.67)	−0.26 (−0.65 to 0.13)	0.19
Total average step count, steps/day	40	362 (−119 to 845)	44	27 (−429 to 482)	−336 (−1,003 to 331)	0.32
Quadriceps strength, Nm/kg						
Surgical limb	61	0.12 (0.04 to 0.21)	64	0.16 (0.07 to 0.24)	0.03 (−0.09 to 0.16)	0.57
Contralateral limb	61	0.09 (0.01 to 0.16)	64	0.15 (0.08 to 0.22)	0.07 (−0.03 to 0.17)	0.18
Active knee ROM surgical limb, degrees						
Extension	61	1.8 (1.00 to 2.53)	64	1.2 (0.45 to 1.94)	−0.6 (−1.64 to 0.50)	0.30
Flexion	61	−2.3 (−4.48 to −0.05)	64	−1.3 (−3.43 to 0.91)	1.0 (−2.09 to 4.09)	0.52
WOMAC surgical knee, points	60	−22.3 (−25.7 to −18.9)	64	−19.1 (−22.4 to −15.8)	3.2 (−1.45 to 7.92)	0.17
VR-12, points						
Physical component score	60	11.8 (9.79 to 13.7)	64	11.4 (9.52 to 13.3)	−0.3 (−3.03 to 2.37)	0.81
Mental component score	60	0.1 (−1.42 to 1.68)	64	−0.7 (−2.19 to 0.81)	−0.8 (−2.96 to 1.31)	0.45

* Models control for sex and baseline value of the outcome. All change variables were calculated as six-month value minus preoperative value, and all estimated difference in change values are CONTROL minus MOVE. 30STS, 30-second sit to stand test; 6MW, six-minute walk test; CI, confidence interval; FGS, fixed gait speed; MOVE, movement pattern training group; ROM, range of motion; SCT, stair climb test; SSG, self-selected gait speed; TUG, times up and go test; VR-12, Veteran's RAND 12-item health survey; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

six-month primary endpoint, 89.7% of the MOVE group and 92.9% of the CONTROL group completed testing.

Outcomes. Outcomes by group over time are reported in Table 2. The change from baseline in primary and secondary outcomes and the difference between groups at six months (primary endpoint) are shown in Table 3. Between-group differences on all outcomes at 10 weeks (end of intervention) are reported in Supplementary Table 2, as well as between-group differences on all outcomes at 10 weeks and six months stratified by sex are reported in Supplementary Tables 3–6.

Primary outcome. There were no between-group differences in change of surgical pKEM during FGS at the 10-week (95% CI: –0.02 to 0.06; $P = 0.38$) or six-month timepoints (95% CI: –0.026 to 0.045; $P = 0.61$) (Figure 2). Both groups returned to baseline levels of surgical pKEM by six months.

Secondary outcomes. There were no between-group differences in change of contralateral pKEM during FGS at 10 weeks; however, the MOVE group demonstrated .03 Nm/kgm less contralateral pKEM during FGS at six months compared to CONTROL ($P = 0.045$, 95% CI: –0.05 to –0.001; $d = 0.32$).

There were no between-group differences in change of surgical pKEM during self-selected gait speed at 10 weeks or six months (Figure 3). There were no between-group differences in change of contralateral pKEM during self-selected gait speed at 10 weeks; however, the MOVE group demonstrated .04 Nm/kgm less contralateral pKEM during self-selected gait speed at six months compared to CONTROL ($P < 0.01$, 95% CI: –0.076 to –0.013; $d = 0.44$).

At 10 weeks, the CONTROL group completed 1.2 repetitions more on the 30STS compared to MOVE ($P < 0.01$, 95% CI: 0.3–2.1) and was 0.5 seconds faster on the TUG than MOVE ($P = 0.04$; 95% CI: –1.0 to 0.0). Conclusions remained the same for the sensitivity analyses conducted for the 30STS and TUG. There were no differences in any other secondary outcomes.

At six months, the CONTROL group completed the SCT 0.6 seconds faster than the MOVE group ($P = 0.046$, 95% CI: –1.2 to 0.0; $d = –0.25$). There were no differences in any other secondary outcomes.

There were no differences in participant characteristics between groups in those that completed six-month testing ($P > 0.05$) (see Supplementary Table 7). Using all available data from

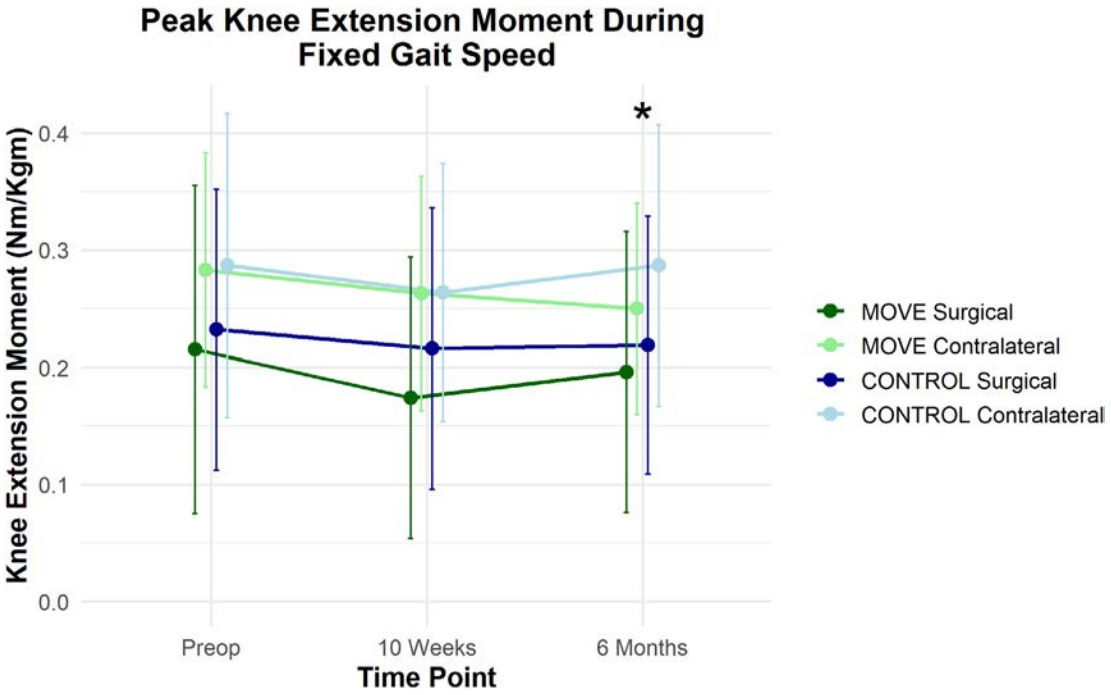

Figure 2. Peak knee extension moment in surgical and contralateral limbs during fixed gait speed of 1.0 m/s by group over time. Error bars are SD. *Significant difference between contralateral peak knee extension moments between groups ($P < 0.05$). MOVE; movement pattern retraining group.

the RM-ANOVA to examine consistency of results when using all available data, changes between groups on contralateral pKEM during FGS and SCT were no longer significant ($P = 0.11$ and $P = 0.74$, respectively). However, changes between groups on contralateral pKEM during FGS were still 0.05 Nm/kgm lower in MOVE ($P = 0.01$, 95% CI: 0.01–0.08).

Treatment outcomes. Overall fidelity to the intervention was high in both groups, with an overall fidelity rating of 98.8%. The mean number of protocol treatments was 13.8 ± 0.5, with no difference in attendance between groups ($P = 0.33$). Average home exercise program adherence was 85.9%, with no difference between groups ($P = 0.89$). HEP adherence levels of 80% and 90% were not related to changes in surgical pKEM during FGS at six months ($P = 0.26$ and $P = 0.19$, respectively). Based on exercise progression criteria, the CONTROL group began performing lunge exercises and step/stair exercises sooner than the MOVE group, at an average of five days ($P < 0.01$, 95% CI: 2–9) and nine days ($P < 0.01$, 95% CI: 4–13), respectively. Average insole wear time in the MOVE group was 11.8 ± 9.6 hours of total home use and was not related to changes in surgical pKEM during FGS at six months ($P = 0.34$). Satisfaction with the intervention was rated as "somewhat satisfied" or "very satisfied" by 87% in the MOVE group and 92% in the CONTROL group, with no difference between groups ($P = 0.32$). Adverse events occurred with nine participants in the MOVE group and nine in the CONTROL group during the 10-week intervention, with the most common event being a fall. Only one fall occurred during a clinic treatment session, and all other adverse events occurred outside of clinic treatment sessions. There were no between-group differences in the occurrence of adverse events ($P = 0.95$).

DISCUSSION

The purpose of this randomized clinical trial was to determine if MOVE improves movement pattern quality more than a standardized rehabilitation program without movement pattern training. Contrary to our hypothesis, the MOVE group demonstrated similar recovery of pKEM in the surgical knee compared to the CONTROL group during walking. Both groups recovered to baseline levels of pKEM in the surgical knee at six months. However, individuals in the MOVE group did have substantially less pKEM on the contralateral limb compared to those in the CONTROL group when walking at self-selected speeds after TKA. MOVE participants demonstrated levels of pKEM more similar to their surgical knee, which could have implications for contralateral OA progression. A secondary goal was to determine if movement pattern training improved quadriceps strength and physical function recovery. Contrary to our hypothesis, CONTROL participants demonstrated a tendency for faster functional performance recovery; however, it is unclear if this is due to improved surgical knee recovery or due to increased compensation on the contralateral knee during the performance of these tasks. Overall, both

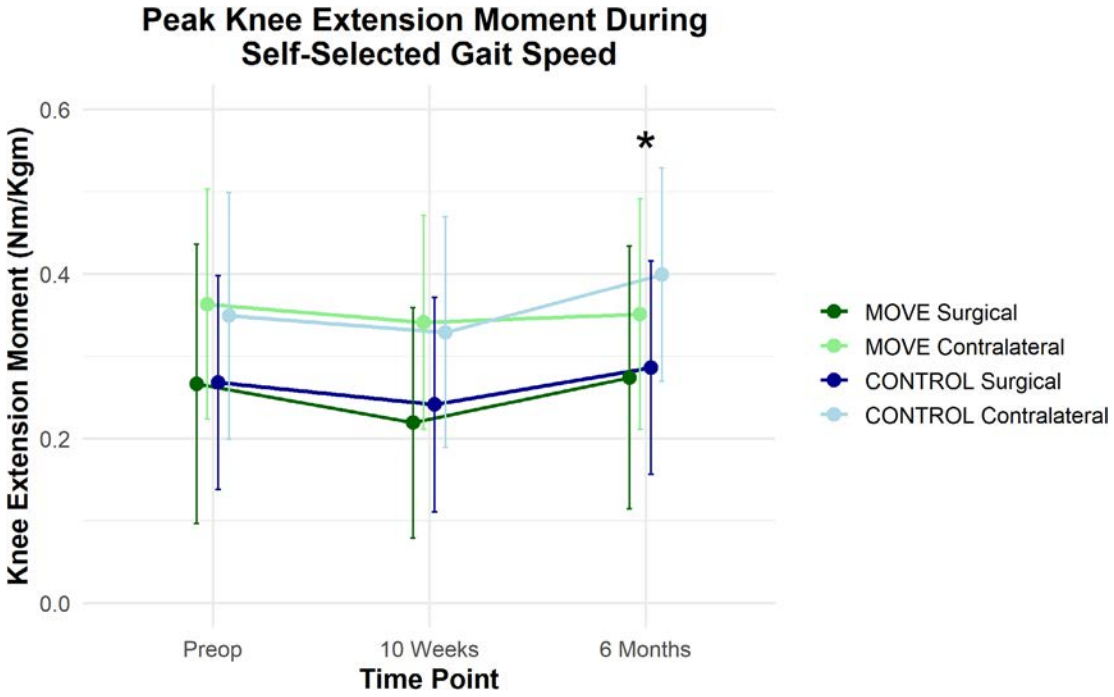

Figure 3. Peak knee extension moment in surgical and contralateral limbs during self-selected gait speed by group over time. Error bars are SD. *Significant difference between contralateral peak knee extension moments between groups ($P < 0.05$). MOVE; movement pattern retraining group. [Correction added on 3 February 2025, after first online publication: Figure 3 has been corrected.]

interventions were efficacious in improving strength and function beyond presurgical levels.

The primary outcome of surgical pKEM was chosen for this study as a clear indicator of movement quality for the following reasons: (1) speed on functional tasks can improve in the first year after TKA despite persistent movement compensations,[3,4,6,8] (2) quadriceps weakness is a chronic impairment following TKA,[35,40,41] (3) decreased surgical pKEM is a primary movement deviation following TKA,[10,16,42] and (4) the MOVE intervention specifically targeted surgical limb loading, and thus assessment of pKEM is a direct assessment of intervention efficacy. However, the MOVE intervention did not lead to improved pKEM during walking as hypothesized. One potential explanation for this finding is that the type of data provided by insoles to MOVE participants primarily consisted of weight-bearing force data as opposed to directly providing feedback on pKEM. By only providing feedback using weight-bearing force data, it is possible that participants achieved more symmetrical weight-bearing during tasks by compensating at other joints (eg, hip or ankle). Christensen et al[43] determined that the type of data provided to participants, delivered during a single session of biofeedback training three months after TKA, directly impacted responses to a movement pattern training intervention, with pKEM data providing superior biomechanical outcomes compared to weight-bearing force data. Future research should examine if providing pKEM data to individuals early after TKA leads to long-term changes in movement quality and superior outcomes compared to the current trial. However, providing individuals pKEM feedback during walking conventionally requires use of a motion capture system to combine the kinetic and kinematic aspects of the gait pattern, using force and motion measures processed with customized software to provide real-time biofeedback to participants. Thus, current technology can limit the environments with which movement pattern feedback on pKEM could be provided.

Another potential explanation for the lack of observed efficacy of the MOVE intervention in improving surgical pKEM is that the mode of feedback may have influenced motor learning during walking. For walking, typically intermittent or summary feedback was used owing to the dynamic nature of this task and safety concerns when using an iOS device for feedback (eg, looking at a device while walking). Future research should examine if providing real-time feedback (eg, using treadmill training) leads to superior outcomes, which has been suggested with findings from other studies examining strategies for retraining gait.[27,44] Furthermore, the MOVE intervention may have influenced other biomechanical aspects of movement quality during walking (eg, weight-bearing force symmetry recovery) or influenced the recovery of other tasks trained during the MOVE intervention (eg, rising from a chair). Future analyses will be conducted to determine the impact of MOVE on additional biomechanical variables and tasks.

Finally, in this trial the CONTROL intervention may have been more effective in restoring pKEM in the surgical knee than prior intervention studies. Prior research examining the recovery of pKEM after TKA has consistently found that pKEM decreases in

the six months following surgery.[10,16,42] In the current study, the CONTROL group recovered to baseline levels of pKEM in the surgical knee indicating the supervised, standardized, progressive rehabilitation protocol used in the CONTROL intervention may be more effective than traditional rehabilitation techniques. Standardization of clinical care has been shown to reduce variation in care and lead to improved outcomes.[45,46] Without a true standard of care control group for comparison, it is not possible to make conclusions regarding either program's efficacy in improving outcomes after TKA compared to standard of care interventions. Given the improved recovery of pKEM in the CONTROL group compared to historically reported outcomes, future research should examine the impact of structured rehabilitation programs similar to CONTROL compared to alternate rehabilitation pathways (eg, self-directed, mobile application-based, and group-based) and their impact on surgical limb pKEM recovery as well as long-term outcomes.

Although the MOVE intervention did not influence the recovery of pKEM in the surgical knee, MOVE led to small to moderate decreases in contralateral pKEM compared to CONTROL. The CONTROL group exhibited levels of pKEM in the contralateral knee equal to or greater than baseline depending on walking condition, whereas the MOVE group decreased contralateral pKEM to values closer to the surgical limb. Higher knee extension moments, observed in the CONTROL group, have been associated with knee OA progression, although results across studies are inconsistent.[11–13] Other biomechanical variables such as knee flexion excursion and knee adduction moment have shown stronger associations with OA progression.[3] Subsequent analyses will examine the long-term impact of MOVE intervention on additional biomechanical outcomes and OA progression at two years using magnetic resonance imaging.

A secondary hypothesis of this clinical trial was that the MOVE intervention would lead to greater recovery of quadriceps strength and functional performance compared to CONTROL. However, CONTROL participants demonstrated a tendency for faster functional performance recovery as demonstrated by 1.2 increased repetitions on the 30STS and 0.5 seconds faster TUG times at 10 weeks. At six months, differences on the 30STS and TUG attenuated, but CONTROL participants were faster on the SCT by an average of 0.6 seconds; however, secondary analysis using all available data indicated that this difference was not significant. The tendency for faster recovery might be explained by differential progression criteria used within the intervention programs. MOVE participants were required to demonstrate symmetrical movement pattern criteria to advance in exercise progression, whereas CONTROL participants needed only to safely tolerate the progression. This led to a more rapid progression of activity-based tasks in the CONTROL group. Additionally, the MOVE intervention emphasizes movement quality, which may have led to more of a focus on quality over speed during functional performance testing. Faster functional performance recovery in the CONTROL group could be due to improved recovery in the surgical knee or increased movement compensation relying on contralateral knee function.

A potential limitation to the current study is the use of a FGS 1.0 m/second for determination of our primary outcome. The rationale for choosing this speed was based on pilot data supporting used to (1) generate our sample size estimate,[16] (2) allow for population comparisons of joint moments at a consistent speed, and (3) minimize the number of participants excluded from the study and thereby increase generalizability. However, slower walking speed can minimize asymmetries and thus decrease the likelihood of detecting asymmetries.[47] Self-selected gait speed as well as higher-level biomechanical tasks such as sit to stand may better measures of asymmetries in pKEM. A second limitation is the length of testing sessions (2–3 hours) may have led to fatigue in some participants and thus biased outcome measures to be lower. However, the effects of fatigue are likely to be similar across time within groups, and thus fatigue may have less of an effect on change scores. A final limitation is that although the insoles used in the trial have been validated during walking, running, and landing,[18,19,22] they have not been validated across all tasks/environments used within this trial; plus, the bipedal calibration process used may have limited the accuracy of feedback and thus decreased intervention effectiveness.

In conclusion, the MOVE intervention did not lead to improved surgical knee use during walking or improved recovery after TKA compared to CONTROL. However, the MOVE group exhibited lower pKEM on the contralateral knee during walking after intervention, which may influence contralateral OA progression. Future research will examine if the MOVE intervention led to changes in biomechanical outcomes other than pKEM during walking and higher-demand tasks such as rising from a chair. Additionally, future research will determine the long-term effects of the MOVE intervention on movement quality and contralateral OA progression.

ACKNOWLEDGMENTS

Marc Cornier (SO), Brad Walters, Katie Seidler, Michelle Kochanek, Kurt Schulze, Katie Carbiener, Laura Baum, Lyle McKissick, Jenna Boehm, Chris Carlisle, Becca Rzepkowski, Beth Wood, Michelle Rauzi, Jeremy Graber, Aaron Jansen, Amanda Dunn, Darcy Reisman, Samantha Lewis, Karl Wojnar, Meghan Lovely, Peter Thomsen, Jessica Juan, Eli Edwards, Eliza Biondi, Michael Samelson, Margaret McCandless, Cassandra Constine, Laura Durkee, Dylan Hudson, Marie-France Izere, Jamie McMullen, Zachary Sapinsley, David Wise, Josh Scurles, Hannah Mader, Ellie Grupehhoff, Cassidy Weeks, Jacob Capin, Laura Churchill, Hope Davis-Wilson, Brandy Cuellar, Meghan Connors, Alexandra Smith, Jared Foran, Charlie Yang, Todd Miner, Lindsay Kleeman-Forsthuber, Jason Jennings, Roseann Johnson, Andrew Park, Edmund Brinkis, Braden Mayer, Ian Dickey, Jared Pate.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation

AND drafting or reviewing/editing the final draft. As corresponding author, Dr Bade confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Kurtz S. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am 2007; 89(4):780–785.
2. Fingar KR, Stocks C, Weiss AJ, et al. Most frequent operating room procedures performed in U.S. hospitals. 2003-2012. Agency for Healthcare Research and Policy website. Accessed January 13, 2025. https://www.hcup-us.ahrq.gov/reports/statbriefs/sb186-Operating-Room-Procedures-United-States-2012.jsp
3. Zeni JA Jr, Flowers P, Bade M, et al. Stiff knee gait may increase risk of second total knee arthroplasty. J Orthop Res 2019;37(2):397–402.
4. Farquhar SJ, Reisman DS, Snyder-Mackler L. Persistence of altered movement patterns during a sit-to-stand task 1 year following unilateral total knee arthroplasty. Phys Ther 2008;88(5):567–579.
5. Naili JE, Iversen MD, Esbjörnsson AC, et al. Deficits in functional performance and gait one year after total knee arthroplasty despite improved self-reported function. Knee Surg Sports Traumatol Arthrosc 2017 2017;25(11):3378–3386.
6. Yoshida Y, Zeni J, Snyder-Mackler L. Do patients achieve normal gait patterns 3 years after total knee arthroplasty? J Orthop Sports Phys Ther 2012;42(12):1039–1049.
7. Yoshida Y, Mizner RL, Ramsey DK, et al. Examining outcomes from total knee arthroplasty and the relationship between quadriceps strength and knee function over time. Clin Biomech (Bristol) 2008; 23(3):320–328.
8. Christiansen CL, Bade MJ, Judd DL, et al. Weight-bearing asymmetry during sit-stand transitions related to impairment and functional mobility after total knee arthroplasty. Arch Phys Med Rehabil 2011;92(10): 1624–1629.
9. Shakoor N, Block JA, Shott S, et al. Nonrandom evolution of end-stage osteoarthritis of the lower limbs. Arthritis Rheum 2002;46(12): 3185–3189.
10. Worsley P, Stokes M, Barrett D, et al. Joint loading asymmetries in knee replacement patients observed both pre- and six months post-operation. Clin Biomech (Bristol) 2013;28(8):892–897.
11. Chehab EF, Favre J, Erhart-Hledik JC, et al. Baseline knee adduction and flexion moments during walking are both associated with 5 year cartilage changes in patients with medial knee osteoarthritis. Osteoarthritis Cartilage 2014;22(11):1833–1839.
12. Brisson NM, Wiebenga EG, Stratford PW, et al. Baseline knee adduction moment interacts with body mass index to predict loss of medial tibial cartilage volume over 2.5 years in knee osteoarthritis. J Orthop Res 2017;35(11):2476–2483.
13. Chang AH, Moisio KC, Chmiel JS, et al. External knee adduction and flexion moments during gait and medial tibiofemoral disease progression in knee osteoarthritis. Osteoarthritis Cartilage 2015;23(7): 1099–1106.
14. Erhart-Hledik JC, Favre J, Andriacchi TP. New insight in the relationship between regional patterns of knee cartilage thickness, osteoarthritis disease severity, and gait mechanics. J Biomech 2015;48(14): 3868–3875.
15. 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: rehabilitation best practices after total joint arthroplasty. Arthritis Care Res (Hoboken) 2014;66(3):411–423.
16. Christiansen CL, Bade MJ, Davidson BS, et al. Effects of weight-bearing biofeedback training on functional movement patterns following total knee arthroplasty: a randomized controlled trial. J Orthop Sports Phys Ther 2015;45(9):647–655.
17. Zeni J, Abujaber S, Flowers P, et al. Biofeedback to promote movement symmetry after total knee arthroplasty: a feasibility study. J Orthop Sports Phys Ther 2013;43(10):715–726.
18. Renner KE, Williams DSB, Queen RM. the reliability and validity of the Loadsol® under various walking and running conditions. Sensors (Basel) 2019;19(2):265.
19. Burns GT, Deneweth Zendler J, Zernicke RF. Validation of a wireless shoe insole for ground reaction force measurement. J Sports Sci 2019;37(10):1129–1138.
20. Bade MJ, Christensen JC, Zeni JA Jr, et al. Movement pattern biofeedback training after total knee arthroplasty: randomized clinical trial protocol. Contemp Clin Trials 2020;91:105973.
21. Fitzgerald GK, Fritz JM, Childs JD, et al. Exercise, manual therapy, and use of booster sessions in physical therapy for knee osteoarthritis: a multi-center, factorial randomized clinical trial. Osteoarthritis Cartilage 2016;24(8):1340–1349.
22. Peebles AT, Maguire LA, Renner KE, et al. Validity and repeatability of single-sensor Loadsol insoles during landing. Sensors (Basel) 2018; 18(12):4082.
23. Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res 2008;51(1):S225–S239.
24. Yang JF, Musselman KE, Livingstone D, et al. Repetitive mass practice or focused precise practice for retraining walking after incomplete spinal cord injury? A pilot randomized clinical trial. Neurorehabil Neural Repair 2014;28(4):314–324.
25. Sturmberg C, Marquez J, Heneghan N, et al. Attentional focus of feedback and instructions in the treatment of musculoskeletal dysfunction: a systematic review. Man Ther 2013;18(6):458–467.
26. Lin CH, Winstein C, Fisher B, Wu A. Neural correlates of the contextual interference effect in motor learning: a transcranial magnetic stimulation investigation. J Mot Behav 2010;42(4):223–232.
27. Davis IS, Tenforde AS, Neal BS, et al. Gait retraining as an intervention for patellofemoral pain. Curr Rev Musculoskelet Med 2020;13(1): 103–114.
28. Wulf G. Self-controlled practice enhances motor learning: implications for physiotherapy. Physiotherapy 2007;93(2):96–101.
29. Aily JB, da Silva AC, de Noronha M, et al. Concurrent validity and reliability of video-based approach to assess physical function in adults with knee osteoarthritis. Phys Ther 2024;104(6):pzae039.
30. Murray AM, Gaffney BM, Davidson BS, et al. Biomechanical compensations of the trunk and lower extremities during stepping tasks after unilateral transtibial amputation. Clin Biomech (Bristol) 2017;49: 64–71.
31. Mizner RL, Petterson SC, Clements KE, et al. Measuring functional improvement after total knee arthroplasty requires both performance-based and patient-report assessments: a longitudinal analysis of outcomes. J Arthroplasty 2011;26(5):728–737.
32. Gill S, McBurney H. Reliability of performance-based measures in people awaiting joint replacement surgery of the hip or knee. Physiother Res Int 2008;13(3):141–152.
33. Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: six-minute walk test, berg balance scale, timed up & go test, and gait speeds. Phys Ther 2002;82(2):128–137.

34. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39(2):142–148.

35. Mizner RL, Petterson SC, Snyder-Mackler L. Quadriceps strength and the time course of functional recovery after total knee arthroplasty. J Orthop Sports Phys Ther 2005;35(7):424–436.

36. Norkin CC, White DJ. Measurement of Joint Motion: A Guide to Goniometry, 4th ed. F.A. Davis Company; 2009.

37. Bellamy N, Buchanan WW, Goldsmith CH, et al. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15(12):1833–1840.

38. Jones D, Kazis L, Lee A, et al. Health status assessments using the Veterans SF-12 and SF-36: methods for evaluating outcomes in the Veterans Health Administration. J Ambul Care Manage 2001; 24(3):68–86.

39. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Lawrence Eearlbaum Associates; 1988.

40. 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(9):559–567.

41. Meier WA, Marcus RL, Dibble LE, et al. The long-term contribution of muscle activation and muscle size to quadriceps weakness following total knee arthroplasty. J Geriatr Phys Ther 2009;32(2):79–82.

42. Mandeville D, Osternig LR, Chou LS. The effect of total knee replacement on dynamic support of the body during walking and stair ascent. Clin Biomech (Bristol) 2007;22(7):787–794.

43. Christensen JC, Foreman KB, LaStayo PC, et al. Comparison of 2 forms of kinetic biofeedback on the immediate correction of knee extensor moment asymmetry following total knee arthroplasty during decline walking. J Orthop Sports Phys Ther 2019;49(2):105–111.

44. White SC, Lifeso RM. Altering asymmetric limb loading after hip arthroplasty using real-time dynamic feedback when walking. Arch Phys Med Rehabil 2005;86(10):1958–1963.

45. Capin JJ, Minick KI, Stevens-Lapsley JE, et al. Variation in outcomes and number of visits following care guideline implementation: part 2 of an analysis of 12 355 patients after total knee arthroplasty. J Orthop Sports Phys Ther 2023;53(3):151–158.

46. Minick KI, Hunter SJ, Capin JJ, et al. Improved outcomes following a care guideline implementation: part 1 of an analysis of 12 355 patients after total knee arthroplasty. J Orthop Sports Phys Ther 2023;53(3): 143–150.

47. Lai H, Chen X, Huang W, et al. Whether patients with anterior cruciate ligament reconstruction walking at a fast speed show more kinematic asymmetries? Orthop Surg 2024;16(4):864–872.

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 744–752
DOI 10.1002/acr.25492

American College of Rheumatology
Empowering Rheumatology Professionals

Estimating Budget Impact and Joint Replacement Avoidance by Implementing a Standardized Education and Exercise Therapy Program for Hip and Knee Osteoarthritis in a Publicly Insured Health Care System

Darren R. Mazzei,[1] Jackie L. Whittaker,[2] Peter Faris,[3] Tracy Wasylak,[3] and Deborah A. Marshall[1]

Objective. The study objective was to estimate the budget impact of funding a standardized education and exercise therapy program, Good Life with osteoArthritis in Denmark (GLA:D) for people with hip and knee osteoarthritis (OA) waiting for total joint replacement (TJR) consultation in a universal publicly insured health care system in Canada.

Methods. We built a budget impact analysis model to estimate the annual cost (Canadian dollars) of providing the GLA:D program to people waiting for a TJR consultation and then forecasted a three-year budget cycle. The base case assumes that 40% attend GLA:D sessions, that 11% avoid surgery, uniform care delivery, that training costs are incurred separately, and that the health care system has enough trained staff to meet demand. The population of people with hip and knee OA waiting for a TJR consultation was estimated with government statistics, peer-reviewed evidence, and routinely collected data from five orthopedic centralized intake clinics (serving 80% of people seeking TJR). Patient-level costs were collected prospectively. International published evidence informed the TJR avoidance estimates. A one-way sensitivity analysis of key parameters evaluated model robustness. Four scenarios were analyzed: public funding for everyone (base case), low-income, rural, and uninsured persons.

Results. Funding GLA:D would cost $4.3 million, serve 12,500 people, and save $8.5 million by avoiding 1,300 TJRs in year one. Savings grow to $8.8 and $8.7 million in years two and three. The number of TJRs performed annually produced the most uncertainty in budget impact (–$15.3 million, –$1.8 million). The most cautious parameter estimates still produce cost savings.

Conclusion. Publicly funding standardized education and exercise therapy programs for everyone waiting for a TJR consultation would avoid surgeries, improve access to evidence-based treatments, and save more than the program costs.

INTRODUCTION

Osteoarthritis (OA) is one of the most common chronic conditions globally.[1] Health care system resources are strained by an aging population, obesity, and high OA prevalence.[2,3] In Canada, $1.26 billion is spent annually performing over 100,000 total joint replacements (TJRs),[4] and demand is expected to increase.[2] Many publicly funded health care systems struggle with long wait times for TJR. Wait times have also worsened because surgery volumes were reduced during the COVID-19 pandemic to maintain hospital bed capacity.[5] National targets in Canada recommend that the 90th percentile of wait times for TJR surgery should be within 26 weeks after the orthopedic surgeon and patient agree surgery is necessary. However, the 90th percentile is currently being seen within 89.6 weeks for consultation and undergoing TJR surgery within 91.6 weeks.[6]

Dr Mazzei's work was supported by the Arthritis Society Training Graduate PhD Salary Award, Cumming School of Medicine Graduate Student Scholarship, Arthur J. E. Child Chair Rheumatology Outcomes Research, and Canadian Institute for Health Research Health System Impact Fellowship co-funded by the Alberta Health Services Bone and Joint Health Strategic Clinical Network. Dr Whittaker's work was supported by the Michael Smith Foundation for Health Research and the Arthritis Society. Dr Marshall's work was supported by the Arthur J. E. Child Chair in Rheumatology and a Canada Research Chair in Health Systems and Services Research (2008–2018).

[1]Darren R. Mazzei, PhD, Deborah A. Marshall, PhD: University of Calgary, Calgary, Alberta, and Arthritis Research Canada, Vancouver, British Columbia, Canada; [2]Jackie L. Whittaker, PhD: University of Calgary, Calgary, Alberta, and Arthritis Research Canada and the University of British Columbia, Vancouver, British Columbia, Canada; [3]Peter Faris, PhD, Tracy Wasylak, MSc: University of Calgary and Alberta Health Services, Calgary, Alberta, Canada.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25492).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25492.

Address correspondence via email to Deborah A. Marshall, PhD, at damarsha@ucalgary.ca.

Submitted for publication June 21, 2024; accepted in revised form December 20, 2024.

SIGNIFICANCE & INNOVATIONS

- Universal publicly insured health care systems like Canada's could spend $4 million offering GLA:D to everyone waiting for a TJR consultation.
- Publicly funding a structured education and exercise program like GLA:D is an affordable solution that could help decision-makers improve access to evidence-based treatments.

Decision-makers have increased surgical capacity,[7] but long wait times persist. Alternative solutions are necessary to address the wait time crisis for TJR.

TJRs are appropriate and effective for end-stage OA after all other treatment options have been exhausted.[8] International clinical guidelines recommend that everyone with hip and knee OA receive education, exercise therapy, and weight management as first-line treatment with adjunctive pharmacological pain management as needed.[9–12] These guidelines have existed for 25 years,[13] but first-line treatments are consistently underused whereas medication and surgery are overused.[14–17] Almost 40% of people with knee OA did not attempt first-line treatments before having a TJR.[18] Standardized programs like Good Life with osteoArthritis in Denmark (GLA:D) were developed to implement high-quality hip and knee OA treatments into routine care.[19,20] GLA:D includes 2 education sessions and 12 supervised neuromuscular exercise sessions delivered twice per week.[19] Eighty-five thousand people in 10 countries have attended GLA:D,[21] with most paying out of pocket because many health care systems and reimbursement plans do not include standardized education and exercise therapy programs.[22] A randomized controlled trial (RCT) evaluating knee replacement reported that 68% of surgical candidates randomized to an education and exercise program had avoided surgery two years after the intervention.[23] Ensuring everyone undergoing TJR is end stage by optimizing nonsurgical care before surgery may help alleviate long wait times, but resource implications are an important consideration for decision-makers. We conducted a budget impact analysis (BIA) to assess the affordability of publicly funding a standardized education and exercise therapy program like GLA:D before TJR.

We used the publicly insured health care system in Alberta, Canada, as an example in our BIA because the public health care provider Alberta Health Services (AHS) has supported GLA:D implementation since 2017. In Canada, the federal government provides co-funding for each province to deliver 100% publicly insured coverage for medically necessary doctor and hospital-based services. The Ministry of Health (MOH) in each province provides additional co-funding and decides how to deliver health care to the population. The province of Alberta spends $24.5 billion annually[24] delivering health care to a population of 4.4 million. Community-based services like GLA:D are funded by a complex mixture of public and private insurance or out-of-pocket payment. Patient-level costs were also recently collected in Alberta for a cost-effectiveness analysis comparing persons receiving GLA:D versus usual care (defined as any community-based service people used to manage their OA symptoms before a TJR). Collecting patient-level costs for a standardized OA program and usual care presents an opportunity to estimate the budget impact and assess the affordability of these programs from the health care systems perspective.

METHODS

We followed the International Society for Pharmacoeconomics and Outcomes Research BIA guidelines to transparently report the parameters and methods used when estimating the budget impact (Canadian dollars) of adopting a new intervention in a health care system.[25]

Model design. Following standard practice, we programmed a cost calculator in Microsoft Excel to estimate the public health care system's (Alberta, Canada) annual budget spent delivering care to people waiting for a TJR consultation

$$\boldsymbol{Annual\ Budget} = \big((\boldsymbol{A} - \boldsymbol{B} + \boldsymbol{C}) * \boldsymbol{D}\big) + (\boldsymbol{E} * \boldsymbol{F}) - (\boldsymbol{C} * \boldsymbol{G}) + (\boldsymbol{B} * \boldsymbol{G})$$

Where:
A = Number of people waiting for TJR consultation annually
B = Number of TJRs annually
C = Number of TJRs avoided annually
D = Cost of community management annually
E = Number of people waiting for TJR consultation who participate in GLA:D® annually
F = Cost of GLA:D (per person)
G =Cost of TJR (per person)

Figure 1. Budget impact analysis formula. GLA:D, Good Life with osteoArthritis in Denmark; TJR, total joint replacement.

(Figure 1). Our model takes the MOH perspective because it includes all publicly funded health care costs and is considered the reference case in Canada.[26] Costs were extrapolated over a three-year time horizon to be consistent with MOH budget forecasts and were calculated in 2022 Canadian dollars to reflect when our cost estimates were collected. We assumed the following: (1) a proportion of people waiting for a TJR consultation will participate in GLA:D regardless of prior treatment, (2) the cost of training clinicians in GLA:D delivery will be funded by employers' professional development budgets, (3) GLA:D delivered in person or virtually is uniform across all locations, (4) each GLA:D class has six participants, (5) the health care system has reached a steady state by training enough staff to meet demand for the program, and (6) half of surgical volumes are total knee replacements and half are total hip replacements because this reflects the case mix in Alberta over five fiscal years (2017/2018 to 2021/2022).[4] Infections, revisions, delaying TJR, bilateral TJR, and TJR for a different joint were excluded from our BIA model because we assumed that GLA:D participation would not change the costs related to these clinical characteristics.

Data sources. Model inputs were estimated from peer-reviewed research, gray literature, local administrative data, and expert clinical opinion, as described below (Table 1).

Population estimates. *Population waiting for TJR consultation.* The population waiting for a TJR consultation was estimated from routinely collected data at five orthopedic centralized intake clinics that provide access to approximately 80% of TJRs throughout Alberta. We assumed that the population of people waiting for a TJR consultation would increase at the same rate as those with OA in Alberta. The population of people with OA was estimated by multiplying population growth, mortality rates, and OA prevalence in Alberta.[27–29]

Forecasted demand for TJR. The Alberta Bone & Joint Health Institute, an independent charitable organization focused on turning knowledge into better care for people with bone and joint conditions, forecasts that demand for TJRs will grow in Alberta from 13,867 to 15,028 surgeries annually over the study period.[30]

GLA:D participation rates. We extracted participation rates from peer-reviewed research evaluating exercise therapy in people with hip and knee OA and then asked experts their opinion. Eighty percent of patients eligible for TJR consented to participate in RCTs of exercise therapy,[31] but clinical experts thought participation may be lower in the real world. We conservatively estimated that participation rates would be half of what were observed in peer-reviewed research (40%) when people were invited to participate in GLA:D if it was publicly funded while they wait for a TJR consultation.

Population avoiding TJR. RCTs have demonstrated that 44%[32] to 68%[23] of people with hip and knee OA avoided TJR after being randomized to exercise therapy. However, there might be selection bias in these samples. Only 9% (127 of 1,475) of those screened were eligible to participate in the study by Skou et al,[31] and 79% (100 of 127) of eligible patients were willing to be randomized. We used real-world data from the GLA:D Canada database to estimate that 11% of participants would avoid a TJR.[33] GLA:D participants were asked, "Are you so troubled by your knee/hip problems that you want surgery?" with "yes" or "no" as possible answers. Participants who responded "yes" before the GLA:D program and then "no" at 12 months were used to estimate the percentage of the population who

Table 1. Parameters used in the budget impact analysis model*

Parameter (Alberta specific)	Value	Source
Total population, n	4.44M	Government of Alberta[33]
Annual population growth rate, %	1.5	Government of Alberta[33]
All-cause mortality rate, %	0.6	Government of Alberta[34]
OA prevalence, %	8.0	A Rowe, MSc, Alberta Health Services, personal communication (email), January 19, 2023 to T. Wasylak
OA incidence (annual), %	0.9	A Rowe, MSc, Alberta Health Services, personal communication (email), January 19, 2023 to T. Wasylak
OA population waiting for TJR consultation, n	31,227	Alberta Bone & Joint Health Institute[35]
Forecasted number of TJRs annually (2021/2022 to 2024/2025)	13,867–15,028	Alberta Bone & Joint Health Institute[36]
Per-person cost of GLA:D at private clinics, $CAD	400	GLA:D clinics
Per-person cost of GLA:D at public clinics, $CAD	304	Expert opinion
Annual cost per person to manage OA with UC, $CAD	653	Mazzei et al[30]
Average cost per TJR, $CAD	10,116	AHS[37]
Implementation, $CAD	211,920	AHS[38]
Percentage avoiding TJR, %	11	GLA:D Canada[39]
GLA:D participation rate from population waiting for a TJR, %	40	Expert opinion

* Estimates are in 2022 Canadian dollars (CAD). AHS, Alberta Health Services; GLA:D, Good Life with osteoArthritis in Denmark; M, million; OA, osteoarthritis; TJR, total joint replacement; UC, usual care.

would avoid TJR for the three-year budget cycle. This estimate is comparable to the 12% of people who reported undergoing a TJR within 12 months of participating in GLA:D although we feel unwillingness is a better predictor of avoidance than people who actually proceeded to surgery.[34]

Cost estimates. *Community management.* The cost of managing OA in the community was estimated from administrative data in a cohort of participants receiving usual care in Alberta, Canada. The average cost was applied to each person in the population of people waiting for a TJR.

GLA:D. The price to attend GLA:D ranges from $375 to $450 at private clinics in Alberta. We assumed the average price was $400 because only 1 of 68 clinics charged $450 when the study was conducted. Public facility costs were estimated by taking an average physiotherapist salary ($43.48 hourly plus 20% for benefits) multiplied by 2.5 hours per class (30-minute preparation, 60-minute class, 30-minute take-down, and 30-minute charting) for 14 classes, producing an estimated cost of $1,826 per class. Assuming six participants per class produced a per-person cost of $304. Public facility cost estimates do not include facility costs such as electricity and maintenance because these costs are incurred in a separate part of the budget whether GLA:D is delivered or not. We assumed clinics already had the necessary equipment because GLA:D was designed to use minimal equipment, and resistance bands would often be purchased by the patient for a nominal fee.

TJR surgery. In 2022, AHS estimated that the average TJR costs $10,116 (A Rowe, MSc, Alberta Health Services, personal communication (email), January 19, 2023 to T. Wasylak). Surgical costs include physician compensation, materials, staff time, and bed days in hospital. This estimate does not include rehabilitation because these costs are predominantly incurred out of pocket in Canada.[35]

Implementation. Implementation costs were estimated by the AHS Bone and Joint Health (BJH) Strategic Clinical Network (SCN), which began piloting GLA:D in 2017. SCNs are the innovation arm of Alberta's publicly funded health care system. SCNs bring together clinical experts, operational leaders, patients, and researchers to produce transformative solutions to improve health care delivery. The BJH SCN supported GLA:D implementation by taking on administrative duties as well as offering annual clinician training classes, hosting regular community of practice meetings for clinicians to learn from one another, and fidelity checks during the pilot phase. GLA:D was implemented at 45 privately funded community rehabilitation clinics, 5 of 40 primary care networks, and 18 of 106 AHS facilities.[36] The BJH SCN hired one additional staff member to support GLA:D implementation, and other team members contributed a portion of their time. Implementation costs include staff time, research grants, travel, training sessions, and event-related costs.

Sensitivity analysis. Parameter uncertainty was evaluated using one-way sensitivity analysis. The estimates are shown in Supplementary Table 1. Each parameter was varied with a high and low estimate to evaluate how variability surrounding each parameter would change the budget impact results. One standard deviation was used for parameters with distributions. The highest and lowest reported price to attend GLA:D at a private facility in Alberta was used to show how price will change the budget impact. GLA:D Denmark and GLA:D Australia ask the same question about wanting surgery before participating in GLA:D and at 12 months, so we used real-world data from these databases as high and low estimates for the percentage of people who would avoid TJR.[21,37] Standard deviations or confidence intervals did not exist in the literature to estimate parameter uncertainty for all other parameters. Therefore, parameters were varied using expert opinion with input from a senior biostatistician, a health care executive, and two clinician-scientists in the research field. Parameter uncertainty ranged from 5% to 50% based on the research team's confidence with each parameter. Results were visualized in a tornado diagram in which parameters were ordered from most to least impact on the primary results.[38]

Scenario analysis. Decision-makers may choose to publicly fund GLA:D for various subpopulations based on costs, expected benefits, clinical characteristics, or equity considerations. Operational leaders within the BJH SCN helped us select four scenarios that were relevant to decision-makers to assess how publicly funding GLA:D for different subpopulations would impact affordability: (1) low-income people to reduce economic inequities, (2) people in rural communities where there are publicly funded hospitals to reduce geographic inequities, (3) high-risk subpopulations for whom TJR surgical risks outweigh the potential benefits (eg, contraindications to general anesthetic), and (4) people who do not have private health insurance that covers allied health professional (eg, rehabilitation) visits because people with private health insurance could use these resources to access a program like GLA:D (Supplementary Table 2).

RESULTS

Base case analysis. We estimate that the MOH will spend $155.4 million in the first year delivering OA care to people waiting for a TJR consultation, and publicly funding GLA:D would reduce the annual budget to $146.7 million (Figure 2). In the first year, it would cost $4.3 million to publicly fund GLA:D, and 1,374 people would avoid surgery, producing net savings of approximately $8.5 million by reducing demand for TJR. This return of investment equals approximately $2 saved for every $1 investment. Over three years, the population waiting for a TJR consultation is expected to grow from 31,227 to 32,817 people. The number of people participating in GLA:D and avoiding TJR would also grow.

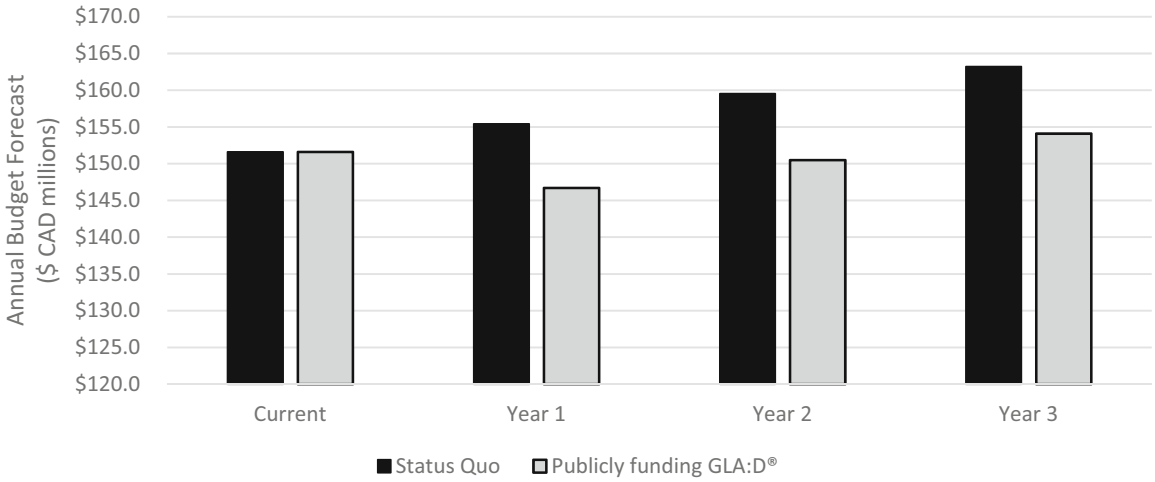

Figure 2. Annual budget forecast of publicly funding GLA:D compared to status quo. Budget impact is the difference in annual budget forecast between status quo and publicly funding GLA:D in each year. CAD, Canadian dollars; GLA:D, Good Life with osteoArthritis in Denmark.

The total budget impact would be –$8.5 million, –$8.8 million, and –$8.7 million in year 1, 2, and 3 respectively (Table 2).

Sensitivity analysis. Parameter uncertainty is shown in Figure 3. All estimates produced cost savings as shown by negative budget impacts. The annual number of TJRs produces the most uncertainty, causing the budget impact in year 1 to range from –$15.3 million if there are 5% fewer TJRs than what was forecasted to –$1.8 million if there are 5% more TJRs than what was forecasted in the base case. The budget impact ranges from –$13.8 to –$4.7 million if the percentage of people avoiding surgery changes from 15.5% to 7.8%. The budget impact will be –$12.9 million if the participation rate is 60% or –$4.7 million with participation rates of 20%. No estimates pass the breakeven point (budget impact of $0), where cost savings are less than the budget to deliver GLA:D.

Scenario analysis. All scenarios would save more than the budget needed to publicly fund GLA:D for the identified subpopulations. Publicly funding GLA:D for low-income, high-surgical-risk, rural, and uninsured subpopulations would cost $0.4 million, $0.6 million, $0.9 million, and $1.3 million while saving $0.5 million, $1.0 million, $1.5 million, and $2.4 million, respectively (Table 3). Publicly funding GLA:D for everyone saves more than when GLA:D is publicly funded for smaller subpopulations.

DISCUSSION

Investing $4.3 million will allow 12,491 people awaiting a hip and knee TJR consultation to participate in GLA:D free of charge and save the MOH approximately $8.5 million in the first year by avoiding 1,374 TJRs. A total of 4,161 TJRs will be avoided over

Table 2. Budget impact for publicly funding GLA:D*

	Current	Year 1	Year 2	Year 3
OA population waiting for TJR consultation, n	31,227	31,227	31,521	32,817
Status quo				
Total annual budget for managing OA population waiting for TJR consultation with status quo, $CAD	151.6	155.4	159.5	163.0
Forecasted number of TJRs annually	13,867	14,267	14,657	15,028
Cost to manage OA without surgery, $CAD	11.3	11.1	11.0	11.0
Publicly funding GLA:D				
Population who attend GLA:D in publicly funded scenario, n	739	12,491	12,608	12,727
Avoided TJRs, n	0	1,374	1,387	1,400
Cost of publicly funding GLA:D, $CAD	0	4.3	4.3	4.4
Cost of avoided TJRs, $CAD	0	−13.9	−14.0	−14.2
Implementation costs	—	$0.2	$0.2	$0.2
Total annual budget for managing OA population waiting for TJR consultation with publicly funding GLA:D	$151.6	$146.7	$150.5	$154.1
Budget impact, $CAD	—	−8.5	−8.8	−8.7
Budget impact (percentage of status quo annual budget), %	—	−5.5	−5.5	−5.3

* Dollar figures are in millions (2022 Canadian dollars [CAD]) and rounded to the nearest decimal so rows may not add. Negative costs indicate cost savings. GLA:D, Good Life with osteoArthritis in Denmark; OA, osteoarthritis; TJR, total joint replacement.

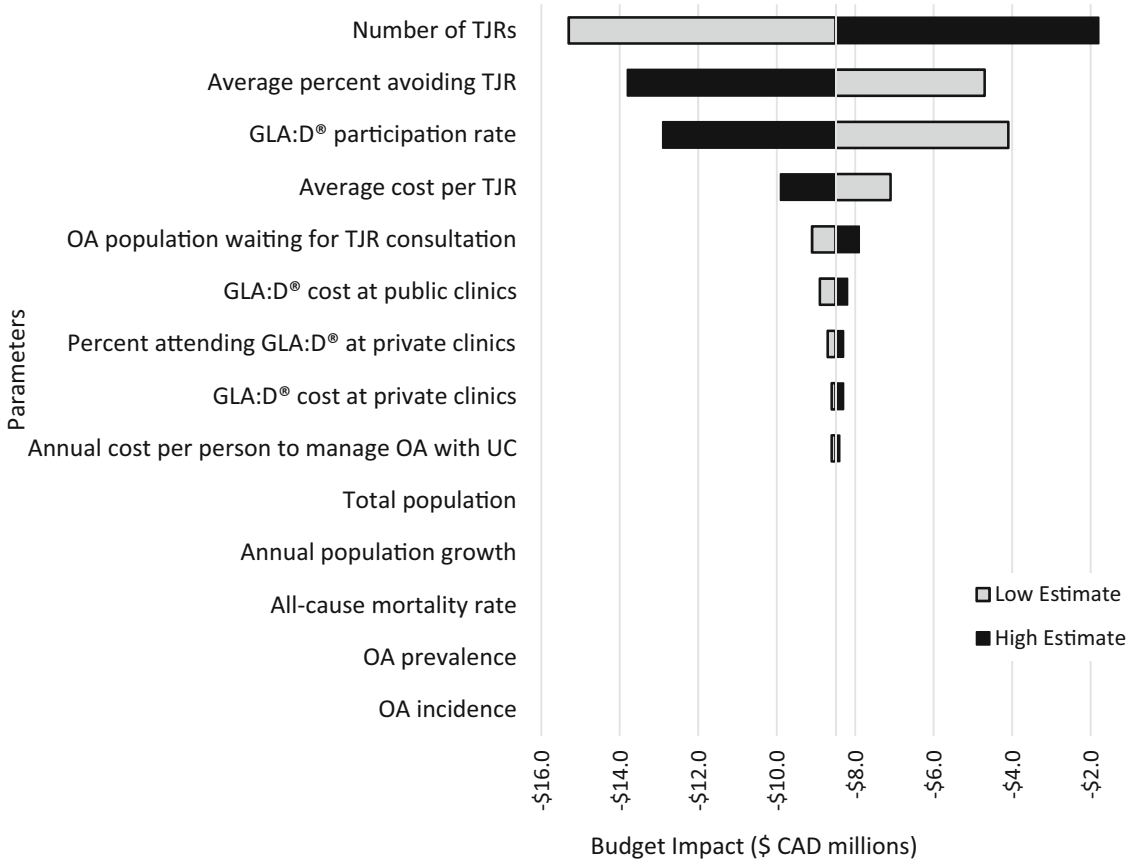

Figure 3. Change in budget impact estimates based on parameter uncertainty. Figure is centered around the Year 1 estimate (–$8.5 million) from the base case analysis. CAD, Canadian dollars; GLA:D, Good Life with osteoArthritis in Denmark; OA, osteoarthritis; TJR, total joint replacement; UC, usual care.

three years. By year three, cost savings will grow to $8.7 million annually as the population awaiting a TJR consultation grows.

We estimated that 40% of people would participate if GLA:D was offered free of charge and 11% of participants would avoid surgery. Parameter uncertainty changes the budget impact, but even pessimistic estimates for participation rates and the percentage of people who avoid surgery will still break even (shown by negative budget impacts). Based on our findings, publicly funding the GLA:D program would pay for itself if as few as 3% of people who participated in GLA:D or 1% of everyone waiting for a TJR consultation avoid surgery. Our scenario analysis showed that funding GLA:D for subpopulations instead of everyone would cost less but also produce less savings. A health care system will save more if more people participate in GLA:D. Providing universal public funding to a structured education and exercise therapy program like GLA:D ensures everyone has equitable access to evidence-based OA treatments, regardless of socioeconomic, geographic, or clinical characteristics.

Table 3. Scenario analysis for publicly funding GLA:D in select subpopulations*

Scenario[a]	Percentage of population	Number of annual GLA:D participants	Cost to deliver GLA:D, $CAD	Budget impact, $CAD
All (base case)	40	12,491	4.3	−8.5
Low income	8	1,024	0.4	−0.5
High surgical risk	14	1,749	0.6	−1.0
Rural	20	2,498	0.9	−1.5
Uninsured	30	3,747	1.3	−2.4

* Dollar figures are in millions (2022 Canadian dollars [$CAD]) and rounded to the nearest decimal. GLA:D, Good Life with osteoArthritis in Denmark; TJR, total joint replacement.
[a] Subpopulations waiting for TJR consultation. Low income is defined as the income situation below which families or persons would likely devote a higher proportion of their after-tax income than average to the necessities of food, shelter, and clothing. High surgical risk refers to subpopulations for whom TJR surgical risks outweigh the potential benefits. Rural refers to all areas outside of population centers. Uninsured is defined as people who do not have private health insurance that covers allied health professional services.

Our results align with other budget impact models evaluating standardized education and exercise therapy programs. Ackerman et al found that the Australian health care system could save $300 to $690 million if standardized education and exercise therapy programs were implemented nationally.[39] Their results showed more savings than ours because they assumed surgical avoidance and intervention costs from RCT data.[23] Populating our model with real-world data from people accessing OA care in the community is likely more generalizable to the policy options that decision-makers face. Smith et al showed that an exercise and diet intervention for OA would have a similar cost to other health promotion programs from the perspective of commercial insurance or Medicare Advantage plans in the United States.[40] Our results add to the evidence base by evaluating a standardized education and exercise therapy program from the perspective of a publicly insured health care system delivering universal access.

Other health care systems can learn from the implementation experience in Alberta, although some contexts are unique.[41] The implementation costs in our model were quite small because only one full-time equivalent staff member was hired to support implementation. The SCNs act as supportive infrastructure within AHS by providing teams and resources to support innovation. Academics and nonprofit organizations like Bone and Joint Canada also played an important role in setting up and maintaining routine data collection. Health care systems may incur additional implementation costs if innovation teams are not already embedded within their organization and partners do not offer in-kind support for common goals. Administrative costs were also not included because the GLA:D program and AHS do not have centralized referral pathways or patient navigation services for people with hip and knee OA. Our BIA model assumed that a ramp-up of training had already occurred and that the health care system has reached a steady state with enough capacity in GLA:D trained staff to deliver the program to 12,000 people annually in the first year of the program. This volume is feasible in a health care system like AHS, which has supported implementation of GLA:D for several years, but health care systems adopting a new program may have reduced volumes before reaching a steady state. We assumed that training costs were funded by employers' professional development budgets, which is common practice for allied health professionals across Canada, but other health care organizations may have to consider these costs. We estimate that it would cost $150,800 to train 260 staff ($580 per staff[42]) to meet the demand for the GLA:D program in our BIA model (assuming every clinician delivers GLA:D twice per quarter to an average of six participants, serving a total of 48 participants annually). Increasing the capacity of trained allied health professionals is a primary barrier during the initial stages of implementation. Training multiple providers at each clinic is important to deliver the program sustainably. Publicly funding a program like GLA:D may incentivize clinicians to take the training course faster than what occurred in Alberta. Program uptake was facilitated by a community of practice, prepackaged materials, and the ability to perform the exercise program without specialized equipment.[41] However, implementing GLA:D took longer than expected, with most clinicians delivering their first class three to four months after training.[41] Marketing the program is critical to increase patient uptake. Clinicians believed referral pathways would also remove barriers to the program. The GLA:D program was originally delivered in-person but was adapted to a virtual delivery model during the COVID-19 pandemic. Virtual delivery is an important option in countries with a large land mass like Canada, specifically for people in rural and remote communities. Our model shows that the difference in cost based on delivering GLA:D at private versus public clinics is marginal. When deciding the delivery location, health care systems should consider what is feasible to rapidly scale the program based on the local context of allied health professionals.

Our budget impact results are complementary to the previously published economic evaluations showing that standardized OA programs are cost-effective in many health care systems. Cost-effectiveness helps decision-makers understand whether a new intervention generates more value (ie, health benefit) for money than an alternative intervention. However, it is possible for a new intervention to be cost-effective but not affordable if the price is high and a large percentage of the population uses the new intervention. We estimated that a standardized education and exercise therapy program like GLA:D is cost-effective and affordable because it may help people avoid TJRs, which cost 25 times more than the GLA:D program per person. Combining cost-effectiveness and affordability provides a comprehensive economic picture of implementing GLA:D into a publicly insured health care system.

Publicly insured health care systems use waitlists to control demand for finite resources like TJRs. This means cost savings in the real world would be observed as reduced wait times instead of budget reductions because another person would have the TJR that was avoided. Using queuing theory,[43] we estimate that the 90th percentile wait time for TJR would be reduced by 12.3 weeks if 11% of GLA:D participants avoided surgery. This means that a publicly insured health care system in Canada could reduce the 90th percentile wait time for TJR surgery from 91.6 weeks to 79.3 weeks. Health care systems could spend $4 million offering GLA:D to everyone waiting for a TJR consultation or $14 million increasing surgical volumes to achieve the same wait time reductions. However, increased surgical volumes also assume that there is operating room capacity and trained staff (eg, orthopedic surgeons, anesthesiologists, and nurses) to meet the increased surgical demand. Publicly funding a structured education and exercise program like GLA:D is an affordable solution that could help decision-makers reduce long wait times.

Our BIA model uses real-world costs and implementation experiences within a publicly insured health care system to

showcase the financial considerations of implementing standardized education and exercise therapy programs into a large publicly insured health care system. However, health care system benefits are likely underestimated because our model only considers the benefits for OA and ignores the additional health benefits that can be gained from exercise for 35 other chronic diseases.[44,45] Policy-makers in other jurisdictions must also recognize that our cost savings are underestimated because postoperative rehabilitation accounts for 39% of health care costs related to TJR,[35] but these costs were not considered in our model because most community rehabilitation services are not publicly funded in Alberta. Although the evidence is mixed, participating in a presurgical standardized education and exercise therapy program may also offer small postsurgical health benefits and reduced length of stay,[46] which are factors not considered in our BIA model. One-way sensitivity analysis is a less rigorous method to evaluate model uncertainty, but most parameters in our model lacked standard deviations, making probabilistic sensitivity analysis not feasible. Our model assessed funding a standardized education and exercise therapy program for people waiting for a TJR consultation because long waitlists are the most relevant problem for decision-makers; however, clinical guidelines recommend education and exercise therapy right after diagnosis. Future research will need to evaluate optimal timing of education and exercise therapy to maximize clinical benefits and health care system resources. We also assumed surgeries were avoided for the entire budget cycle whereas some people may delay but still go on to have a TJR. Skou et al showed that 26% of randomized patients proceeded to surgery at one year and 9% after two years, suggesting a diminishing percentage of people delaying surgery.[23,31] We expect that delayed surgery would be an insignificant cost compared to the total annual cost of managing everyone waiting for a hip and knee TJR consultation. Assessing whether standardized education and exercise therapy programs actually help people avoid TJR in the real world also has important implications. Lastly, implementation research can help health care systems reduce other barriers like misinformation, knowledge gaps, expectations, and referral patterns to increase participation rates.[47,48]

Our results suggest that providing GLA:D to everyone waiting for a TJR consultation would avoid surgeries and save more than the program costs in a universal publicly insured health care system like Canada's. Funding GLA:D prior to TJR consultation would be an affordable solution to reduce wait times in publicly funded health care systems.

ACKNOWLEDGMENTS

The following members of the Alberta Health Services Bone and Joint Health Strategic Clinical Network provided valuable insights about implementing GLA:D in a publicly insured health care system: Jill Robert, Kira Ellis, Ania Kania-Richmond, Mel Slomp, and Jason Martyn.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Marshall confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. GBD 2021 Osteoarthritis Collaborators. Global, regional, and national burden of osteoarthritis, 1990–2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Rheumatol 2023;5:e508–e522.
2. Sloan M, Premkumar A, Sheth NP. Projected volume of primary total joint arthroplasty in the U.S., 2014 to 2030. J Bone Joint Surg Am 2018;100:1455–1460.
3. Bombardier C, Hawker G, Mosher D, et al. The Impact of Arthritis in Canada: Today and Over the Next 30 Years. Arthritis Alliance Canada; 2011.
4. Hip and Knee Replacements in Canada: CJRR Annual Report 2021–2022. Canadian Institute for Health Information; 2022. Accessed November 15, 2022. https://www.cihi.ca/sites/default/files/document/hip-knee-replacements-in-canada-cjrr-annual-report-2021-2022-en.pdf
5. Sniderman J, Khoshbin A, Wolfstadt J. The influence of the COVID-19 pandemic on total hip and knee arthroplasty in Ontario: a population-level analysis. Can J Surg 2023;66:E485–E490.
6. Alberta Bone & Joint Health Institute. Wait times [hip & knee replacement]. 2020. Accessed March 26, 2022. https://www.albertaboneandjoint.com/for-patients/wait-times/
7. Government of Alberta. Health analytics interactive portal: Alberta Surgical Initiative Performance Dashboard. 2023. Accessed September 9, 2023. https://healthanalytics.alberta.ca/health-analytics.html
8. Hawker G, Bohm ER, Conner-Spady B, et al. Perspectives of Canadian stakeholders on criteria for appropriateness for total joint arthroplasty in patients with hip and knee osteoarthritis. Arthritis Rheumatol 2015;67:1806–1815.
9. Bannuru RR, Osani MC, Vaysbrot EE, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthritis Cartilage 2019;27:1578–1589.
10. Fernandes L, Hagen KB, Bijlsma JW, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis 2013;72:1125–1135.
11. Kolasinski SL, Neogi T, Hochberg MC, et al. 2019 American College of Rheumatology/Arthritis Foundation Guideline for the Management of Osteoarthritis of the Hand, Hip, and Knee. Arthritis Care Res (Hoboken) 2020;72:149–162.
12. Bruyère O, Honvo G, Veronese N, et al. An updated algorithm recommendation for the management of knee osteoarthritis from the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO). Semin Arthritis Rheum 2019;49:337–350.
13. Hochberg MC, Altman RD, Brandt KD, et al. Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. American College of Rheumatology. Arthritis Rheum 1995;38:1541–1546.
14. Hinman RS, Nicolson PJA, Dobson FL, et al. Use of nondrug, nonoperative interventions by community-dwelling people with hip and knee osteoarthritis. Arthritis Care Res (Hoboken) 2015;67:305–309.

15. Hagen KB, Smedslund G, Østerås N, et al. Quality of community-based osteoarthritis care: a systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2016;68:1443–1452.
16. Khoja SS, Almeida GJ, Freburger JK. Recommendation rates for physical therapy, lifestyle counseling, and pain medications for managing knee osteoarthritis in ambulatory care settings: a cross-sectional analysis of the National Ambulatory Care Survey (2007–2015). Arthritis Care Res (Hoboken) 2020;72:184–192.
17. Mazzei DR, Whittaker JL, Kania-Richmond A, et al. Do people with knee osteoarthritis use guideline-consistent treatments after an orthopedic surgeon recommends nonsurgical care? A cross-sectional survey with long-term follow-up. Osteoarthr Cartil Open 2022;4:100256.
18. King LK, Marshall DA, Faris P, et al. Use of recommended non-surgical knee osteoarthritis management in patients prior to total knee arthroplasty: a cross-sectional study. J Rheumatol 2019;47:1–8.
19. 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.
20. Roos EM, Barton CJ, Davis AM, et al. GLA:D® to have a high-value option for patients with knee and hip arthritis across four continents: Good Life with osteoArthritis from Denmark. Br J Sports Med 2018;52:1544.
21. Grønne DT, Roos E, Skou ST. GLA:D® Annual Report 2017. University of Southern Denmark; 2017. Accessed October 28, 2018. https://www.glaid.dk/pdf/%C3%85rsrapport_2017_eng.pdf
22. Allen KD. Cost-effectiveness of physical activity and exercise therapy programs for knee osteoarthritis: making the case for health plan coverage. Osteoarthritis Cartilage 2020;28:719–720.
23. Skou ST, Roos E, Laursen M, et al. Cost-effectiveness of total knee replacement in addition to non-surgical treatment: a 2-year outcome from a randomised trial in secondary care in Denmark. BMJ Open 2020;10:e033495.
24. *Budget* 2023 Highlights. Government of Alberta; 2023. Accessed September 9, 2023. https://open.alberta.ca/dataset/fe2a75b5-571b-4b44-8b04-66a6d81aaba4/resource/20f11317-cc2f-4468-95c3-dae61dd516ae/download/budget-2023-highlights-strengthening-health-care.pdf
25. Sullivan SD, Mauskopf JA, Augustovski F, et al. Budget impact analysis—principles of good practice: report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. Value Health 2014;17:5–14.
26. Lee KM, McCarron CE, Bryan S, et al. CADTH Methods and Guidelines: Guidelines for the Economic Evaluation of Health Technologies: Canada. Canadian Agency for Drugs and Technologies in Health; 2017. 4th ed. Accessed November 15, 2023. https://www.cadth.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf
27. Government of Alberta. Alberta population estimates: population statistics. 2018. Accessed September 18, 2019. https://open.alberta.ca/opendata/alberta-population-estimates-data-tables/resource/5597096f-59b0-4fa5-b5a6-859a238c8350
28. Demographic Spotlight: Mortality in Alberta. Government of Alberta; 2010. Accessed November 8, 2019. https://open.alberta.ca/dataset/c04c9b36-251f-463e-a2b5-0eaafd61d2af/resource/a7441764-318d-4c4e-bcb2-81bf00508975/download/2010-1217-mortality-in-alberta.pdf
29. Marshall DA, Vanderby S, Barnabe C, et al. Estimating the burden of osteoarthritis to plan for the future. Arthritis Care Res (Hoboken) 2015;67:1379–1386.
30. Osteoarthritis Crisis in Alberta: Access, Quality, and Long-Term Planning. Alberta Bone & Joint Health Institute; 2019. Accessed February 14, 2020. https://albertaboneandjoint.com/wp-content/uploads/ABJHI_Osteoarthritis_Crisis_in_Alberta_2019.pdf
31. Skou ST, Roos EM, Laursen MB, et al. A randomized, controlled trial of total knee replacement. N Engl J Med 2015;373:1597–1606.
32. Svege I, Nordsletten L, Fernandes L, et al. 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 2013;74:164–169.
33. Zywiel MG, McGlasson R. GLA:D® Canada 2020 Annual Report. Bone and Joint Canada; 2021. Accessed November 5, 2023. https://gladcanada.ca/
34. Zywiel MG, McGlasson R. GLA:D® Canada 2020 Supplementary Report. Bone and Joint Canada; 2021. Accessed November 5, 2023. https://gladcanada.ca/
35. Marshall DA, Wasylak T, Khong H, et al. Measuring the value of total hip and knee arthroplasty: considering costs over the continuum of care. Clin Orthop Relat Res 2012;470:1065–1072.
36. List of GLA:D® Canada Clinics - Alberta. Bone and Joint Canada; 2023. Accessed November 5, 2023. https://gladcanada.ca/wp-content/uploads/2024/09/List-of-GLAD-Clinics-AB-Sept-25-2024-1.pdf
37. Crossley KM, Barton C, Kemp J, et al. GLA:D® Australia: 2021 Annual Report. La Trobe University; 2021. Accessed September 30, 2023. https://gladaustralia.com.au/annual-reports/
38. Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Oxford University Press; 2015.
39. Ackerman IN, Skou ST, Roos EM, et al. Implementing a national first-line management program for moderate-severe knee osteoarthritis in Australia: a budget impact analysis focusing on knee replacement avoidance. Osteoarthr Cartil Open 2020;2:100070.
40. Smith KC, Losina E, Messier SP, et al. Budget impact of funding an intensive diet and exercise program for overweight and obese patients with knee osteoarthritis. ACR Open Rheumatol 2020;2:26–36.
41. Kania-Richmond A, Beaupre LA, Jones CA. GLA:D® Hip & Knee Program: Alberta Feasibility Evaluation Report. Alberta Health Services Bone and Joint Health Strategic Clinical Network; 2021. Accessed March 18, 2022. https://www.albertahealthservices.ca/assets/about/scn/ahs-scn-bjh-ab-glad-feasibility-evaluation-report.pdf
42. GLA:D® Canada. Sign up for a certification course. 2024. Accessed October 22, 2024. https://gladcanada.ca/for-trainers/sign-up-for-a-certification-course/
43. Wolff RW. Stochastic Modeling and the Theory of Queues. Prentice-Hall; 1989.
44. Pedersen BK, Saltin B. Exercise as medicine—evidence for prescribing exercise as therapy in 26 different chronic diseases. Scand J Med Sci Sports 2015;25:1–72.
45. Booth FW, Roberts CK, Laye MJ. Lack of exercise is a major cause of chronic diseases. Compr Physiol 2012;2:1143–1211.
46. Moyer R, Ikert K, Long K, et al. The value of preoperative exercise and education for patients undergoing total hip and knee arthroplasty: a systematic review and meta-analysis. JBJS Rev 2017;5:e2.
47. Wallis JA, Taylor NF. Pre-operative interventions (non-surgical and non-pharmacological) for patients with hip or knee osteoarthritis awaiting joint replacement surgery—a systematic review and meta-analysis. Osteoarthritis Cartilage 2011;19:1381–1395.
48. MacKay C, Hawker GA, Jaglal SB. Qualitative study exploring the factors influencing physical therapy management of early knee osteoarthritis in Canada. BMJ Open 2018;8:e023457.

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 753–759
DOI 10.1002/acr.25485

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Treatment Response Biomarkers for Systemic Sclerosis–Associated Interstitial Lung Disease

Elizabeth R. Volkmann,[1] Holly Wilhalme,[1] Donald P. Tashkin,[1] Grace Hyun J. Kim,[1] Jonathan Goldin,[1] Alana Haussmann,[1] Masataka Kuwana,[2] Michael D. Roth,[1] and Shervin Assassi[3]

Objective. This study investigated whether changes in circulating biomarkers predict progressive pulmonary fibrosis (PPF) in patients with systemic sclerosis–associated interstitial lung disease (SSc-ILD) receiving treatment.

Methods. Participants of the Scleroderma Lung Study II, which compared receiving mycophenolate mofetil (MMF) versus cyclophosphamide (CYC) for treating SSc-ILD, who had blood samples at baseline and 12 months were included. Levels for C-reactive protein (CRP), interleukin-6, C-X-C motif chemokine ligand (CXCL) 4, CCL18, and Krebs von den Lungen (KL)-6 were measured, and a logistic regression model evaluated relationships between changes in these biomarkers and the development of PPF by 24 months.

Results. A total of 92 of the 142 randomized participants had longitudinal biomarker measurements and the required clinical outcome data, with 19 participants (21%) meeting criteria for PPF. In the whole cohort, changes in KL-6 levels were significantly correlated with PPF. KL-6 increased in patients who developed PPF and decreased in patients who did not (mean change ± SD 365.68 ± 434.41 vs –207.45 ± 670.26; $P < 0.001$). In the arm of participants who received MMF alone, changes in CRP and CXCL4 levels were also significantly correlated with PPF. When added to an existing prediction model based on baseline factors associated with PPF in this cohort (sex, baseline reflux severity, and CXCL4 levels), the change in KL-6 remained significantly associated with PPF (odds ratio 1.4; $P = 0.0002$).

Conclusion. Changes in the circulating levels of KL-6 after treatment with MMF or CYC predicted PPF, even after adjusting for baseline factors associated with PPF. Measuring longitudinal KL-6 in patients with SSc-ILD may improve how we personalize therapy in patients with SSc-ILD.

INTRODUCTION

Although both approved and nonapproved therapies appear to favorably modify the course of the forced vital capacity %-predicted (FVC) in patients with systemic sclerosis–associated interstitial lung disease (SSc-ILD),[1] observational studies demonstrate that at least 25% of patients will develop progressive pulmonary fibrosis (PPF).[2] One barrier to improving outcomes for patients with SSc-ILD is the lack of early and reliable treatment response biomarkers, potentially leading to delays in the initiation of alternative, effective therapies for patients with SSc-ILD. The present treatment response indicators include physiologic, radiologic, and clinical assessments, which are inherently difficult to interpret in the context of a patient with a systemic autoimmune disease with multiple comorbidities. For example, multiple factors outside of ILD may affect exercise tolerance (eg, arthritis, myopathy), cough (eg, gastroesophageal reflux disease), and FVC measurements due to factors outside of parenchymal fibrosis (eg, thoracic cutaneous fibrosis restricting chest wall expansion, respiratory muscle weakness) in patients with SSc-ILD. In addition, it can often take a year or more to establish a reliable trend in FVC course due to the inherent variability of FVC measurements.

Supported by the National Heart, Lung, and Blood Institute, NIH (grant K23-HL-150237 to Dr Volkmann, grant U01-HL-60587 to Dr Tashkin, and grant U01-HL-60606 to Robert Elashoff). Dr Assassi's work was supported by the Department of Defense (grant W81XWH-22-1-0162).

[1]Elizabeth R. Volkmann, MD, MS, Holly Wilhalme, MS, Donald P. Tashkin, MD, Grace Hyun J. Kim, PhD, Jonathan Goldin, MD, PhD, Alana Haussmann, MD, Michael D. Roth, MD: University of California, Los Angeles; [2]Masataka Kuwana, MD, PhD: Nippon Medical School, Tokyo, Japan; [3]Shervin Assassi, MD, MS: University of Texas Health Science Center at Houston.

Drs Roth and Assassi contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25485).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25485.

Address correspondence via email to Elizabeth R. Volkmann, MD, MS, at evolkmann@mednet.ucla.edu.

Submitted for publication July 11, 2024; accepted in revised form December 11, 2024.

SIGNIFICANCE & INNOVATIONS

- Efforts to advance personalized medicine for systemic sclerosis–associated interstitial lung disease (SSc-ILD) are limited by the lack of objective and valid therapy-specific response measures.
- This is the first study to demonstrate that patients who experience an increase in Krebs von den Lungen (KL)-6 in response to therapy with mycophenolate or cyclophosphamide are more likely to develop progressive pulmonary fibrosis in the following year.
- This study also found that in addition to KL-6, changes in C-reactive protein and C-X-C motif chemokine ligand 4 levels were associated with future progressive pulmonary fibrosis among patients who specifically received mycophenolate.
- Longitudinal measurement of these circulating biomarkers could lead to earlier introduction of alternative therapies for patients with SSc-ILD and avert the development of irreversible fibrosis.

The absence of valid biomarkers of treatment response has hindered our ability to effectively manage this condition and make informed treatment decisions. Biomarkers that detect response to treatment, such as the response to an immunomodulatory medication, relatively soon after a therapy is initiated, may help physicians determine whether or not to continue a specific treatment and/or initiate an alternative medication before irreversible lung damage occurs in patients with SSc-ILD. The present study aimed to identify treatment response biomarkers for SSc-ILD using data from the Scleroderma Lung Study (SLS) II.[3] All patients in the SLS II received treatment with either mycophenolate mofetil (MMF) or cyclophosphamide (CYC). We hypothesized that measuring changes in a select group of circulating biomarkers of inflammation, fibrosis, and epithelial injury after 12 months of therapy would predict the likelihood of PPF during the following year.

PATIENTS AND METHODS

Study participants. Participants enrolled in the SLS II (NCT00883129), a National Institutes of Health–sponsored, randomized controlled trial comparing treatment responses to receiving MMF versus CYC, were included in these post hoc analyses.[3] The SLS II enrolled an ethnically diverse population of both male and female patients with SSc-ILD from 14 sites across the United States according to inclusion and exclusion criteria that have been previously published.[3] The institutional review board of each of the 14 study sites approved this study, and informed consent was obtained by all participants.

SLS II study design. In the SLS II, patients were randomized to receive oral CYC for 12 months followed by 12 months of placebo or MMF for 24 months. The primary endpoint, the FVC percentage predicted, was measured every three months during the study, and high-resolution computed tomography (HRCT) thoracic imaging was obtained at baseline and at 24 months. A computer-aided design scoring system was used to calculate the quantitative radiologic extent of ILD (QILD).[4] The Transitional Dyspnea Index (TDI) was used to assess changes in dyspnea.[5]

Biomarker assessment. Blood samples were collected from study participants and immediately processed onsite on the day of collection, stored at –70°C, and shipped on dry ice to the central repository at the University of Texas, Houston. A select group of measured biomarkers of inflammation and fibrosis were measured at baseline and 12 months, including serum C-reactive protein (CRP; multiplex bead array), serum interleukin (IL)-6 (Simoa Planar Array), plasma CCL18 (enzyme-linked immunosorbent assay [ELISA]), plasma Krebs von den Lungen (KL)-6 (Nanopia latex-enhanced immunoturbidimetric assay), and serum C-X-C motif chemokine ligand (CXCL) 4 (ELISA). These biomarkers were selected because prior studies have demonstrated their role as either prognostic and/or predictive biomarkers in patients with SSc-ILD.[6–9] Technicians performing the assays were blinded to the clinical diagnosis and outcome data (see Supplementary Table 1 for further details on the assessment of these biomarkers).

Primary outcome: PPF. The definition of PPF in the present study was adapted from recently published guidelines.[10] Patients meeting at least two of the following were classified as PPF: (1) worsening respiratory symptoms, (2) absolute decline in FVC ≥5% predicted and/or absolute decline in diffusing capacity for carbon monoxide (DL_{CO}) corrected for hemoglobin ≥10% from baseline, and (3) radiologic evidence of disease progression. PPF was selected as the outcome of interest, rather than FVC alone, because it also incorporates the patients' experience with the disease, as well as radiologic assessment, which approximates ILD burden independently of extrapulmonary factors that may affect FVC measurements (eg, cutaneous fibrosis of the chest, respiratory muscle weakness).

Worsening of respiratory symptoms was defined as a decrease in the TDI score ≥1.5 based on the mean minimally important difference for worsening in patients with SSc-ILD.[11] Radiologic evidence of disease progression was defined as an increase in the whole-lung QILD score ≥2%, as this was the threshold of progression associated with death in two independent cohorts.[12] Patients were classified as having PPF if they fulfilled PPF criteria at 12, 18, or 24 months.

Statistical analysis. *Biomarker assessment.* Mean differences and effect sizes were calculated to determine changes in

biomarker levels from baseline to 12 months in the whole cohort (both treatment arms combined). Log transformation was performed when appropriate. Because MMF is commonly received as background therapy for current clinical trials in SSc-ILD, mean changes in biomarkers were also calculated for the arm of patients who received MMF.

Primary outcome: PPF. Univariable logistic regression was performed to evaluate the relationship between the change in biomarker levels from baseline to 12 months and the presence/absence of PPF at 12, 18, or 24 months in the whole cohort. If the change in a specific biomarker was significantly associated with PPF ($P < 0.05$), this variable was then entered into a multivariable model. This multivariable model included baseline variables associated with PPF in this cohort based on our prior publication (eg, sex, baseline reflux severity, and baseline CXCL4 levels).[13]

To assess the internal validity and quantify the optimism of the predictive model, bootstrap analyses were performed. Specifically, 2,000 sample datasets were generated from the data with replacement, and the average model coefficients were estimated. The adjusted bootstrap percentile method was used to calculate 95% confidence intervals (CIs) for the estimates. Additionally, the model optimism was assessed using Harrell's bias correction of the concordance statistic. The Brier score was calculated to assess predictive accuracy. Due to the small overall sample size and number of events in our dataset, we conducted a sensitivity analysis using Ridge regression to address the potential of overfitting and multicollinearity. We optimized the regularization strength by evaluating the model at two choices of λ: the λ that minimized the cross-validation error, and the most conservative λ within one SE of the minimum, allowing us to balance model complexity with prediction accuracy. The choice of Ridge regression supports the development of a stable, generalizable model by adjusting coefficients uniformly, which is particularly crucial in datasets with limited events. Model performance was assessed through the area under the receiver operator curve (AUC) to evaluate the predictive accuracy and robustness under Ridge regularization compared to initial analyses. All tests were two sided and performed using SAS version 9.4 (SAS Institute).

RESULTS

Participant characteristics. Among all SLS II participants (n = 142), 112 had nonmissing data for at least two of the three criteria for PPF at 12, 18, or 24 months, and 92 of these participants had biomarker measurements at baseline and 12 months. The baseline characteristics of the overall study population have been previously published,[3] and they were similar to those of the 92 participants included in these analyses (Supplementary Table 2).

Primary outcome: PPF. Among the 92 participants with PPF and biomarker data, 19 participants (21%) met our PPF criteria between 12 and 24 months (n = 10 for the arm of participants who received MMF; n = 9 for the arm of participants who received CYC). A total of 6 participants developed PPF at 12 months, 13 participants developed PPF at 24 months, and no participants developed PPF at 18 months (see Supplementary Table 3 for the frequency of patients meeting individual PPF criteria at each time point).

Biomarker changes. In the whole cohort, changes from baseline to 12 months in CRP, IL-6, and CCL18 levels were minor, and no difference between participants with or without PPF was observed (Table 1). However, the KL-6 level decreased in those without PPF, whereas it actually increased in patients who developed PPF, and the difference between these two groups was significant (effect size 1.15; $P < 0.001$; Figure 1; Table 1; Supplementary Figure 1). CXCL4 levels also numerically decreased with treatment, but the difference in changes between those with or those without PPF did not meet the predefined threshold for significance (effect size 0.44; $P = 0.091$; Figure 1; Table 1).

When considering those in the arm who received MMF alone, differences in biomarker changes between those with PPF (n = 10) and those without PPF (n = 39) were more prominent (Table 1). In this subgroup, the CRP levels decreased in those without PPF but actually increased in those with PPF, and the difference between the two groups was significant (effect size 0.68; $P = 0.040$; Table 1). The same pattern was observed for the changes in KL-6 levels (effect size 1.28; $P = 0.004$) and CXCL4 levels (effect size 0.98; $P = 0.038$).

Multivariable prediction model. For the analysis of the entire SLS II cohort, only the change from baseline to 12 months of KL-6 met predefined criteria as a potential predictor of progression to PPF. This variable was therefore added to our existing multivariable prediction model for PPF,[13] which included three baseline variables (sex, Gastrointestinal Tract (GIT) 2.0 reflux scores, CXCL4 levels). The earlier model exhibited a sensitivity and specificity of 68% and 72%, respectively, with an optimism-corrected AUC of 0.69.[14] When the change in KL-6 from baseline to 12 months was added to this model, the AUC increased to 0.89 (95% CI 0.82–0.97), with an improvement in both the sensitivity (95%) and specificity (74%; Supplementary Figure 2). The change in KL-6 remained robustly and significantly associated with the development of PPF even after adjusting for these baseline variables with an odds ratio of 1.4 for a 0.10-unit increase in KL-6 (Table 2). The optimism-corrected AUC resulting from the bootstrap analyses was 0.86; the Brier score was 0.09 (the Brier score range is 0–1; a lower score is better). The Ridge regression analysis, applied to adjust for multicollinearity and potential overfitting, resulted in varying degrees of coefficient shrinkage across model parameters when compared to the original logistic regression coefficients. Notably, for both λ

Table 1. Mean change in biomarker levels from baseline to 12 months in the SLS II cohort*

Biomarker	No PPF, mean (±SD)[a]	PPF, mean (±SD)[a]	Effect size	*P* value	Adj. *P* value[b]
CRP, µg/mL					
Entire cohort	−1.79 (±12.21)	−1.77 (±11.79)	0.19	0.449	0.651
MMF arm	−3.92 (±11.93)	1.23 (±14.70)	0.68	0.040	0.068
IL-6, pg/mL					
Entire cohort	−0.12 (±9.90)	−0.31 (±5.60)	0.12	0.632	0.651
MMF arm	−1.44 (±9.53)	−1.54 (±5.75)	0.09	0.99	0.990
KL-6, U/mL					
Entire cohort	−207.45 (±670.26)	365.68 (±434.41)	1.15	<0.001	<0.001
MMF arm	−226.76 (±432.83)	241.09 (±421.26)	1.28	0.004	0.0217
CCL18, ng/mL					
Entire cohort	−63.86 (±99.74)	−52.78 (±87.79)	0.11	0.651	0.651
MMF arm	−78.19 (±107.13)	−68.76 (±101.25)	0.08	0.74	0.922
CXCL4, ng/mL					
Entire cohort	−1,099.83 (±1,708.56)	−286.14 (±2,223.30)	0.44	0.091	0.227
MMF arm	−894.76 (±1,568.60)	695.15 (±1,857.61)	0.98	0.038	0.068

* Adj., adjusted; CRP, C-reactive protein; CXCL, C-X-C motif chemokine ligand; IL, interleukin; KL, Krebs von den Lungen; MMF, mycophenolate mofetil; PPF, progressive pulmonary fibrosis; SLS, Scleroderma Lung Study.
[a] Values are mean change (±SD) based on raw biomarker measurements before log transformation.
[b] *P* values were obtained using Student's *t*-test in the entire cohort and Wilcoxon's rank-sum test in the arm who received MMF. *P* values were adjusted using the false discovery rate method.

settings, all variables remained in the model, indicating their continued relevance to the predictive outcome despite the regularization. The model's discriminative ability, as measured by the AUC, remained unchanged (AUC 0.89) under both settings. In an exploratory analysis, we examined the relationship between the change in KL-6 level from baseline to 12 months and the change in quantitative ILD- whole lung (QILD-WL) score from baseline to 24 months. In this linear regression analysis, the change in KL-6 at 12 months was significantly associated with the change in QILD-WL at 24 months (estimate 5.86, 95% CI 2.03–9.70; $P = 0.0032$).

DISCUSSION

The present study confirms prior reports that treating patients who have active SSc-ILD with cytotoxic/immunomodulatory medications, MMF or CYC in these studies, can lead to measurable changes in circulating biomarkers of inflammation and fibrosis.[8,14] In addition, our analysis found that the magnitude and direction of change in plasma KL-6 levels over the first 12 months of therapy is associated with the development (or not) of PPF in the following year and can add significantly to the predictive power of previously defined baseline features. These

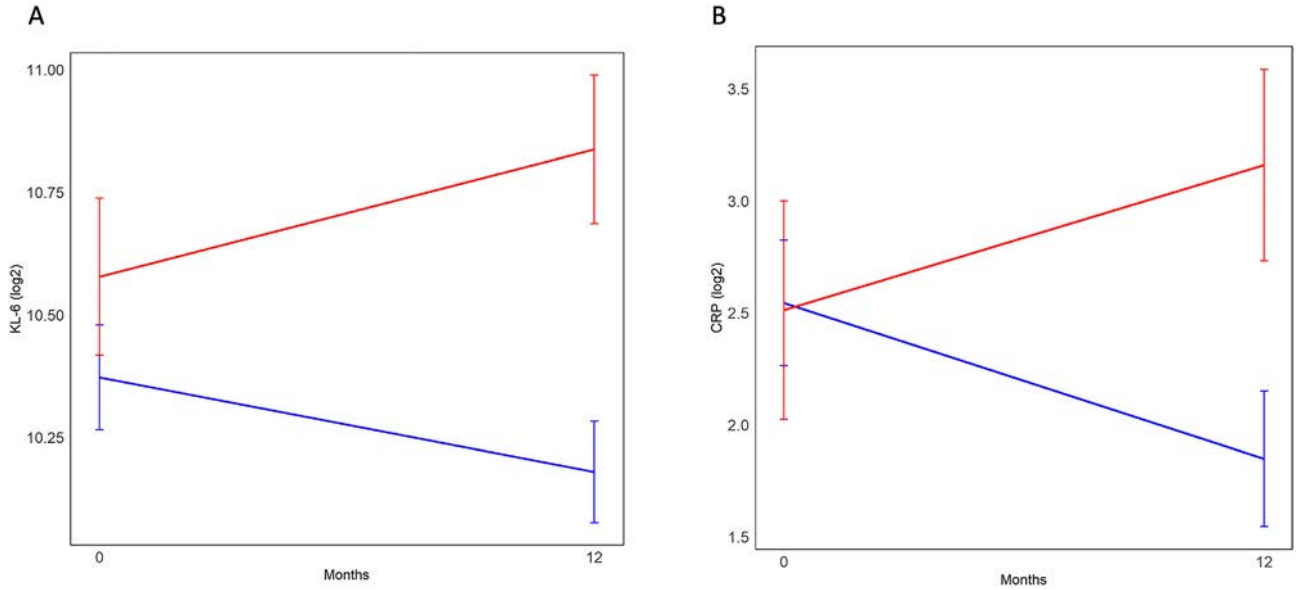

Figure 1. (A) Change in Krebs von den Lungen (KL)-6 in patients who developed progressive pulmonary fibrosis (PPF; red) and those who did not develop PPF (blue) in the combined treatment arms. (B) Change in C-reactive protein (CRP) in patients who developed PPF (red) and those who did not develop PPF (blue) in the arm who received mycophenolate mofetil alone.

Table 2. Multivariable logistic regression model for predicting PPF in the SLS II (n = 99)*

Variable	Odds ratio	95% CI	*P* value
Female	0.16	0.04–0.7	0.0148
Baseline CXCL4 level	0.53	0.3–0.93	0.0270
Baseline reflux score	4.3	1.24–14.87	0.0213
Change in KL-6	1.4	1.17–1.66	0.0002

* CI, confidence interval; CXCL, C-X-C motif chemokine ligand; KL, Krebs von den Lungen; PPF, progressive pulmonary fibrosis; SLS, Scleroderma Lung Study.

findings are in line with prior observational studies evaluating serial changes in KL-6 in patients with SSc-ILD.[15,16]

For example, in a retrospective study of patients with SSc (37% of whom received immunomodulatory treatment) who underwent KL-6 measurement every two months for six months after SSc diagnosis (n = 110), patients with ILD (n = 64) had higher baseline KL-6 levels than the entire population with SSc.[15] In the subgroup of patients, serum KL-6 levels changed over time, but in other patients, KL-6 levels remained stable.[15] Several factors affected KL-6 variability over two years, including the presence of ILD, diffuse cutaneous disease, positive anti-Scl-70 antibodies, negative anticentromere antibodies, increased ILD severity, and immunomodulatory treatment.[15] In another study of patients with ILD (n = 85, of whom 33 had connective tissue disease, ILD), sequential changes in KL-6 were significantly associated with progression of ILD (defined by death or a decline in FVC >10% or DL_{CO} >15% or more at 12 months).[17] Moreover, a study of 77 patients with SSc-ILD demonstrated that an increase in KL-6 of more than 193 U/mL from baseline was significantly associated with SSc-ILD progression (defined as ≥10% relative decline in FVC predicted or 5%–10% decline in FVC predicted along with radiologic progression on HRCT; hazard ratio 7.19, 95% CI 3.3–15.7).[18] If validated in additional prospective cohorts, the results of the present study have important implications for monitoring treatment response to immunomodulatory therapy in patients with SSc-ILD.

Although our prior studies have demonstrated that the majority of patients in the SLS II responded favorably to treatment of receiving CYC and MMF,[3,13] a significant subset of patients experienced PPF despite treatment. In a nonclinical trial population of patients with more comorbidities, the percentage of nonresponders may be larger.[19] Discerning a favorable treatment response in a timely manner is challenging in clinical practice because our current response indicators for PPF (respiratory symptoms, worsening lung function, and/or thoracic imaging) are often difficult to document or only occur late in the disease after significant lung damage has occurred. For example, the intertest variability for FVC is large enough that small changes in FVC measurements are of uncertain clinical significance unless a clear trend across multiple FVC measurements emerges. Similarly, small changes in the radiologic extent of ILD are also difficult to interpret, particularly because structural changes discerned on HRCT (eg, reticulation) cannot discriminate active from inactive fibrosis. For these reasons, it may not become clinically evident that a particular patient is not responding favorably to therapy until a year or more has passed. Because SSc-ILD progression within the first one to two years of therapy is associated with increased mortality rates,[20,21] accurately determining response to therapy within this time window is critical.

The present study demonstrated that levels of the pneumoprotein KL-6 increased in patients who developed PPF and decreased in patients who did not develop PPF in response to treatment of receiving either MMF or CYC. Even after adjusting for other risk factors of PPF, the change in KL-6 remained significantly associated with PPF under varying degrees of regularization. These findings suggest that KL-6 in particular may be an important indicator of treatment response to therapy by receiving MMF and CYC. In certain countries, such as Japan, KL-6 is already routinely used in clinical practice to monitor SSc-ILD progression at a relatively low cost (ie, the equivalent of 7 US dollars [USD]) and is covered in part by health insurance. Therefore, the addition of the change in KL-6 variable to our previously published PPF prediction model would add a negligible cost (ie, the equivalent of approximately 14 USD).

This study also found that in addition to KL-6, CRP and CXCL4 levels increased in patients who developed PPF and decreased in patients who did not develop PPF specifically in response to treatment by receiving MMF. These findings are notable because they suggest that different therapies may have distinct associated response biomarkers. Due to the relatively small sample size of the arm of patients who received MMF (n = 49), multivariable analyses could not be performed. Nevertheless, the CRP findings are consistent with a recent observational study, which demonstrated that persistent elevation in CRP was associated with an increased risk of death in 1,171 patients with SSc-ILD treated with various immunomodulatory therapies.[22] Because CRP tests are cost-effective and widely available, serial measurements of CRP may improve how we monitor treatment response to receiving MMF.

Despite the small number of patients who developed PPF in the present study, we discovered significant associations in this well-characterized cohort of patients with SSc-ILD undergoing uniform follow-up assessments and receiving standardized treatment. Even after adjusting for the false discovery rate, the change in KL-6 remained significantly associated with PPF. Although these findings require validation in other cohorts, the present results also provide insights into the key biologic pathways affected by treatment by receiving MMF and CYC. Understanding these pathways is important for future drug discovery and for the personalization of treatment of patients with SSc-ILD. For example, we have previously demonstrated that high CXCL4 at baseline is associated with a more favorable response to therapy of receiving MMF and CYC.[13]

Recognizing that the SLS II cohort was composed of patients with SSc-ILD who were treatment naive to MMF and CYC and relatively early in their disease course (mean disease duration of approximately two years), it will be important in future studies to determine whether changes in these biomarkers occur at later stages of the disease and under different therapeutic conditions. In the SENSCIS trial, KL-6 decreased in response to treatment by receiving nintedanib.[14] However, it is unknown whether the magnitude of decline in these biomarkers was associated with future ILD progression in the SENSCIS trial.

Another limitation of the present study is that the biomarkers were measured at 12 months when a response to SSc-ILD therapy may be detectable with our currently available response indicators (eg, pulmonary function tests, HRCT). Future studies are needed to determine whether changes in these biomarkers are appreciated at early time points (ie, three and six months after treatment initiation). Reliable response biomarkers detectable early in the course of SSc-ILD treatment could minimize exposure to toxic therapies that are not conferring benefit and maximize exposure to alternative therapies that do confer benefit.

In summary, treatment by receiving MMF and CYC led to measurable decreases in CRP, IL-6, CXCL4, KL-6, and CCL-18 levels in the circulation in patients with SSc-ILD. The change in KL-6 levels at 12 months was significantly associated with future PPF, even after adjusting for potential confounders. In the arm of patients who received MMF, the change in CRP was associated with PPF, suggesting that serial measurements of CRP may provide information about treatment response to this therapy. Additional studies are needed to determine how these response biomarkers perform in patients receiving alternative SSc-ILD therapies, including nintedanib, tocilizumab, and rituximab.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Volkmann confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Raghu G, Montesi S, Silver RM, et al. Treatment of systemic sclerosis-associated interstitial lung disease (SSc-ILD): evidence-based recommendations. Am J Respir Crit Care Med 2024;209:137–152.
2. Hoffmann-Vold AM, Allanore Y, Alves M, et al; EUSTAR collaborators. Progressive interstitial lung disease in patients with systemic sclerosis-associated interstitial lung disease in the EUSTAR database. Ann Rheum Dis 2021;80:219–227.
3. Tashkin DP, Roth MD, Clements PJ, et al; Scleroderma Lung Study II Investigators. 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;4:708–719.
4. Kim HJ, Brown MS, Elashoff R, et al. Quantitative texture-based assessment of one-year changes in fibrotic reticular patterns on HRCT in scleroderma lung disease treated with oral cyclophosphamide. Eur Radiol 2011;21:2455–2465.
5. Mahler DA, Ward J, Fierro-Carrion G, et al. Development of self-administered versions of modified baseline and transition dyspnea indexes in COPD. COPD 2004;1:165–172.
6. Ghuman A, Khanna D, Lin CJF, et al. Prognostic and predictive markers of systemic sclerosis-associated interstitial lung disease in a clinical trial and long-term observational cohort. Rheumatology (Oxford) 2024;63:472–481.
7. Kuwana M, Shirai Y, Takeuchi T. Elevated serum Krebs von den Lungen-6 in early disease predicts subsequent deterioration of pulmonary function in patients with systemic sclerosis and interstitial lung disease. J Rheumatol 2016;43:1825–1831.
8. Volkmann ER, Tashkin DP, Roth MD, et al. Changes in plasma CXCL4 levels are associated with improvements in lung function in patients receiving immunosuppressive therapy for systemic sclerosis-related interstitial lung disease. Arthritis Res Ther 2016;18:305.
9. Volkmann ER, Tashkin DP, Kuwana M, et al. Progression of interstitial lung disease in systemic sclerosis: the importance of pneumoproteins Krebs von den Lungen 6 and CCL18. Arthritis Rheumatol 2019;71:2059–2067.
10. Raghu G, Remy-Jardin M, Richeldi L, et al. Idiopathic pulmonary fibrosis (an update) and progressive pulmonary fibrosis in adults: an official ATS/ERS/JRS/ALAT clinical practice guideline. Am J Respir Crit Care Med 2022;205:e18–e47.
11. Khanna D, Tseng CH, Furst DE, et al; for Scleroderma Lung Study Investigators. Minimally important differences in the Mahler's Transition Dyspnoea Index in a large randomized controlled trial–results from the Scleroderma Lung Study. Rheumatology (Oxford) 2009;48:1537–1540.
12. Volkmann ER, Tashkin DP, Roth MD, et al. Early radiographic progression of scleroderma: lung disease predicts long-term mortality. Chest 2022;161:1310–1319.
13. Volkmann ER, Wilhalme H, Assassi S, et al. Combining Clinical and Biological Data to Predict Progressive Pulmonary Fibrosis in Patients With Systemic Sclerosis Despite Immunomodulatory Therapy. ACR Open Rheumatol 2023;5:547–555.
14. Assassi S, Kuwana M, Denton CP, et al; on behalf of the SENSCIS trial investigators. Effects of nintedanib on circulating biomarkers in subjects with systemic sclerosis-associated interstitial lung disease (SSc-ILD). Ann Rheum Dis 2022;81:719–720.
15. Shirai Y, Fukue R, Kaneko Y, et al. Clinical relevance of the serial measurement of Krebs von den Lungen-6 levels in patients with systemic sclerosis-associated interstitial lung disease. Diagnostics (Basel) 2021;11:2007.
16. Takei R, Yamano Y, Kataoka K, et al. Disease progression and changes in KL-6 in patients with limited cutaneous systemic sclerosis-associated interstitial lung disease. Respir Investig 2024;62:44–48.
17. Jiang Y, Luo Q, Han Q, et al. Sequential changes of serum KL-6 predict the progression of interstitial lung disease. J Thorac Dis 2018;10:4705–4714.
18. Watanabe S, Kase K, Saeki K, et al. Kinetic changes in serum KL-6 levels predict disease progression in patients with systemic sclerosis-associated interstitial lung disease. Respir Med 2022;191:106689.

19. Tashkin DP, Amin AN, Kerwin EM. Comparing randomized controlled trials and real-world studies in chronic obstructive pulmonary disease pharmacotherapy. Int J Chron Obstruct Pulmon Dis 2020;15:1225–1243.
20. Goh NS, Hoyles RK, Denton CP, et al. Short-term pulmonary function trends are predictive of mortality in interstitial lung disease associated with systemic sclerosis. Arthritis Rheumatol 2017;69:1670–1678.
21. Volkmann ER, Tashkin DP, Sim M, et al; SLS I and SLS II study groups. Short-term progression of interstitial lung disease in systemic sclerosis predicts long-term survival in two independent clinical trial cohorts. Ann Rheum Dis 2019;78:122–130.
22. Guler S, Sarbu AC, Stalder O, et al. Phenotyping by persistent inflammation in systemic sclerosis associated interstitial lung disease: a EUSTAR database analysis. Thorax 2023;78:1188–1196.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 760–766
DOI 10.1002/acr.25503

Predictors of Mortality in Antiphospholipid Antibody–Positive Patients: Prospective Results From Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking Clinical Database and Repository

Yasaman Ahmadzadeh,[1] Larry S. Magder,[2] Danieli Castro Oliveira de Andrade,[3] Diana Paredes-Ruiz,[4] Maria G. Tektonidou,[5] Vittorio Pengo,[6] Savino Sciascia,[7] Laura Andreoli,[8] Flávio Signorelli,[9] Paul R. Fortin,[10] Maria Efthymiou,[11] H. Michael Belmont,[12] Giuseppe Barilaro,[13] Ann E. Clarke,[14] Tatsuya Atsumi,[15] Chary López-Pedrera,[16] Jason S. Knight,[17] D. Ware Branch,[18] Rohan Willis,[19] Nina Kello,[20] Zhuoli Zhang,[21] Esther Rodriguez Almaraz,[22] Bahar Artim-Esen,[23] Jose Pardos-Gea,[24] Guillermo Pons-Estel,[25] Giulia Pazzola,[26] Hui Shi,[27] Alí Duarte-García,[28] Jonathan Thaler,[29] Megan R. W. Barber,[14] Leslie Skeith,[14] Massimo Radin,[7] Pier Luigi Meroni,[30] Robert Roubey,[31] Maria Laura Bertolaccini,[32] Hannah Cohen,[11] Michelle Petri,[33] and Doruk Erkan[29]

Objective. The objective was to determine the mortality rate as well as the causes and predictors of death in antiphospholipid antibody (aPL)–positive patients with and without antiphospholipid syndrome (APS) classification.

Methods. The inclusion criterion for the multicenter international Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking (APS ACTION) registry is positive aPLs according to the Revised Sapporo Classification Criteria tested within one year before enrollment. Patients are observed every 12 ± 3 months with clinical data and blood collection. For this prospective analysis, we first analyzed the causes of death for patients reported as "deceased." Secondly, we analyzed risk factors for death using the adjusted Cox proportional hazards model and calculated survival probability using the Kaplan-Meier model based on different age groups.

Results. Of 967 patients, 43 (5%) were deceased after a median follow-up of 5.3 years. Based on the univariate analysis, deceased patients, compared to living patients, were more likely to be older and have a history of arterial thrombosis, catastrophic APS, concomitant systemic autoimmune diseases (SAIDs), and baseline cardiovascular disease (CVD) risk factors. Based on the Cox proportional hazards model adjusted for age and for each of the strongest predictors of death, arterial thrombosis (hazard ratio [HR] 2.94, 95% confidence interval [CI] 1.50–5.76), concomitant SAIDs (HR 2.97, 95% 1.56–5.63), and baseline any CVD risk factor (HR 2.43, 95% CI 1.05–5.71) were significantly associated with mortality.

Conclusion. In our cohort of persistently aPL-positive patients, the mortality rate was 5% after a median follow-up of five years and was highest for patients ≥60 years old at registry entry. History of arterial thrombosis, concomitant SAIDs, and baseline any CVD risk factor independently predicted future death.

INTRODUCTION

Antiphospholipid syndrome (APS) is a systemic autoimmune disorder marked by thrombosis, pregnancy morbidity, and nonthrombotic manifestations (eg, thrombocytopenia) occurring because of antiphospholipid antibodies (aPLs).[1] The disorder is called "primary APS" when it occurs independently of other autoimmune diseases, or it can be associated with other systemic

[1]Yasaman Ahmadzadeh, MD: Roger Williams Medical Center, Providence, Rhode Island; [2]Larry S. Magder, PhD: University of Maryland School of Medicine, Baltimore; [3]Danieli Castro Oliveira de Andrade, MD: University of São Paulo, São Paulo, Brazil; [4]Diana Paredes-Ruiz, MD: Biobizkaia Health Research Institute, Barakaldo, Spain; [5]Maria G. Tektonidou, MD, PhD: National and Kapodistrian University of Athens, Athens, Greece; [6]Vittorio Pengo, MD: Padova University Hospital, Padova, Italy; [7]Savino Sciascia, MD, PhD, Massimo Radin, MD: Azienda Sanitaria Locale Città di Torino and University of Turin, Turin, Italy; [8]Laura Andreoli, MD, PhD: University of Brescia, Brescia, Italy; [9]Flávio Signorelli, MD: Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil; [10]Paul R. Fortin, MD: Centre Hospitalier Universitaire de Québec - Université Laval, Québec, Québec, Canada; [11]Maria Efthymiou, PhD, Hannah Cohen, MD: University College London, London, United Kingdom; [12]H. Michael Belmont, MD: New York University Langone Medical Center, New York; [13]Giuseppe Barilaro, MD: Hospital Clinic Barcelona, Barcelona, Catalonia, Spain; [14]Ann E. Clarke, MD, Megan R. W. Barber, MD, Leslie Skeith, MD: University of Calgary, Calgary, Alberta, Canada; [15]Tatsuya Atsumi, MD: Hokkaido University Hospital, Sapporo, Japan; [16]Chary López-Pedrera, MD: Maimonides Institute for Biomedical Research of Córdoba/Reina Sofia Hospital, University of Cordoba, Cordoba, Spain; [17]Jason S. Knight, MD, PhD: University

SIGNIFICANCE & INNOVATIONS

- Only a few studies exist addressing mortality in antiphospholipid syndrome (APS); thus, this study was designed to determine the mortality rate as well as the causes and predictors of death in antiphospholipid antibody (aPL)–positive patients with and without APS classification.
- Based on the prospective analysis of an international multicenter cohort, the mortality rate was 5% after a median follow-up of five years (estimated five-year survival probability from the registry entry as 0.96 [95% confidence interval (CI) 0.94–0.97]) and was highest for patients ≥60 years old at registry entry (estimated five-year survival probability as 0.86 [95% CI 0.77–0.92]).
- Infections, thrombosis, and malignancy were the most common causes of death; history of arterial thrombosis, concomitant systemic autoimmune diseases, and baseline any cardiovascular risk factor independently predicted future death.
- Our findings underscore the importance of considering the aforementioned factors in the management of patients with APS and hopefully will aid clinicians in identifying high-risk patients with APS.

autoimmune diseases (SAIDs), most commonly systemic lupus erythematosus (SLE).

Knowledge on the mortality of APS is limited; however, based on a small number of studies, mortality rates surpass those of the general population. Rodziewicz et al demonstrated standardized mortality ratios (SMRs) of 1.49 and 1.33 in female and male patients with APS, respectively, compared to the UK general population.[2] Although a US study[3] reported an APS SMR of 1.61, with a 10-year survival rate of 80%, in the Euro-Phospholipid project,[4] the 10-year survival rate was 90.7%, and the unadjusted SMR was 1.8.

The Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking (APS ACTION) was created in 2010 specifically to conduct large-scale multicenter clinical studies and trials in persistently aPL-positive patients. The goal of the APS ACTION Clinical Database and Repository ("registry") is to study the natural course of disease over at least 10 years in persistently aPL-positive patients with or without other SAIDs.[5] Given the limited data on APS mortality, our objective was to determine the mortality rate as well as the causes and predictors of death in aPL-positive patients with and without APS classification using the APS ACTION registry.

METHODS

The APS ACTION registry is a web-based data capturing system to store patient demographics, history, and medications (Members of the APS ACTION group are listed in Appendix A). The inclusion criterion is positive aPLs according to Revised Sapporo Classification Criteria[6] tested twice (at least 12 weeks apart) within one year before enrollment. Patients are observed every 12 ± 3 months with clinical data and blood collection. aPL-specific medical history (including microvascular or non-thrombotic aPL-related manifestations), aPL/APS-related medications, and blood samples (for aPL positivity confirmation) are collected at registry entry. At each follow-up visit, clinical data for the new aPL-related events and new SAIDs, blood samples, and medication changes are collected. For patients with center-reported early terminations, the reason for termination (and in case of death, both the primary cause and the secondary cause[s]) is collected. The registry data are managed using the REDCap electronic data capture tool, a secure web-based system designed to support research studies.[7] The registry is approved by each participating center's institutional review board.

In this prospective analysis of the registry, we grouped patients into two groups: (1) those reported as "deceased" during the follow-up and (2) those who are actively followed up ("living"). Firstly, we descriptively analyzed the causes of death (based on investigators' report) for patients reported as "deceased" during the follow-up. Secondly, we compared the clinical and laboratory characteristics of deceased versus living patients.

For the descriptive analysis, data were summarized; mean ± SD was used for continuous variables. For the comparison of mortality rates between groups, we used a Cox proportional hazards model using age as the time scale (patients with no follow-up visits, eg, those with recent registry recruitment, were excluded,

of Michigan, Ann Arbor; [18]D. Ware Branch, MD: University of Utah and Intermountain Healthcare, Salt Lake City; [19]Rohan Willis, MD: University of Texas Medical Branch, Galveston; [20]Nina Kello, MD: Northwell Health, Great Neck, New York; [21]Zhuoli Zhang, MD: Peking University First Hospital, Beijing, China; [22]Esther Rodriguez Almaraz, MD: Hospital Universitario 12 de Octubre, Madrid, Spain; [23]Bahar Artim-Esen, MD: Istanbul University School of Medicine, Istanbul, Turkey; [24]Jose Pardos-Gea, MD: Vall d'Hebron University Hospital Barcelona, Spain; [25]Guillermo Pons-Estel, MD: Centro Regional de Enfermedades Autoinmunes y Reumáticas del Grupo Oroño, Rosario, Argentina; [26]Giulia Pazzola, MD: Azienda Unita Sanitaria Locale IRCCS di Reggio Emilia, Reggio Emilia, Italy; [27]Hui Shi, MD: Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [28]Alí Duarte-García, MD: Mayo Clinic, Rochester, Minnesota; [29]Jonathan Thaler, MD, Doruk Erkan, MD, MPH: Hospital for Special Surgery, Weill Cornell Medicine, New York, New York; [30]Pier Luigi Meroni, MD: University of Milan, Milan, Italy; [31]Robert Roubey, MD: University of North Carolina at Chapel Hill; [32]Maria Laura Bertolaccini, PhD: King's College London, London, United Kingdom; [33]Michelle Petri, MD, MPH: Johns Hopkins University School of Medicine, Baltimore, Maryland.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25503).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25503.

Address correspondence via email to Doruk Erkan, MD, MPH, at erkand@hss.edu.

Submitted for publication July 10, 2024; accepted in revised form January 13, 2025.

and those with early termination were censored). To construct the multivariable model in stepwise approach, variables that were statistically significant in the univariate analysis were included. The absolute risk of death was estimated using the Kaplan-Meier approach, stratified by age group, sex, and concomitant SAIDs.

RESULTS

As of May 2023, of 1,174 patients recruited, 207 (18%) had no follow-up data. There was no major difference between the major demographic and clinical characteristics of patients with or without follow-up, except those who had follow-up were older and more likely to have SLE (Supplement Table 1).

Among the 967 participants with any follow-up data (mean ± SD age at registry entry: 45.6 ± 13.3 years; female sex: n = 723 [75%]; White race: n = 657 [68%]), 43 (5%) were reported as deceased after a median follow-up of 5.3 years (interquartile range: 2.4–7.9 years; mean ± SD: 4.9 ± 3.2 years for all participants and 2.6 ± 2.4 years for those deceased; Table 1 and Supplement Table 2). As part of the primary and/or secondary causes of death, infection was reported by investigators in 14 (33%) patients, thrombosis was reported in 10 (23%), and malignancy was reported in 8 (19%) (not mutually exclusive) (Table 2).

Based on the univariate analysis, deceased patients, compared to living patients, were more likely to be older and have a history of arterial thrombosis, catastrophic APS (CAPS), concomitant SAIDs, and baseline cardiovascular disease (CVD) risk factors (Table 1). After adjustment for age, history of arterial thrombosis (hazard ratio [HR] 2.99, 95% confidence interval [CI] 1.56–5-74), CAPS (HR 1.79, 95% CI 0.79–4.08), concomitant SAIDs (HR 2.85, 95% CI 1.52–5.36), and baseline any CVD risk factor (HR 2.85, 95% CI 1.23–6.58) were significantly more common in deceased patients compared to living patients (Table 3). Based on the Cox proportional hazards model adjusted for age and for each of the other four strongest predictors of death, history of arterial thrombosis (HR 2.94, 95% CI 1.50–5.76), concomitant SAIDs (HR 2.97, 95% CI 1.56–5.63), and baseline any CVD risk factor (HR 2.43, 95% CI 1.05–5.71) were significantly and independently associated with mortality (Table 4).

The estimated five-year survival probability from registry entry for the entire cohort was 0.96 (95% CI 0.94–0.97). The estimated five-year survival probabilities from registry entry by age group were 0.98 (95% CI 0.92–0.99) for patients <30 years old (n = 2 of 119), 0.98 (95% CI 0.95–0.99) for patients 30 to 44 years old (n = 8 of 362), 0.96 (95% CI 0.93–0.98) for patients 45 to 59 years old (n = 11 of 340), and 0.86 (95% CI 0.77–0.92) for patients ≥60 years old (n = 22 of 142) (Figure 1). The estimated five-year survival probabilities from registry entry by sex were 0.96 (95% CI 0.94–0.97) for female patients and 0.94 (95% CI 0.90–0.97) for male patients. The estimated five-year survival probabilities from registry entry by concomitant SAIDs were 0.98 (95% CI 0.96–0.99) for those without other SAIDs and 0.92 (95% CI 0.89–0.95) for those with other SAIDs.

DISCUSSION

Based on a prospective follow-up of an international cohort of 967 participants over a median follow-up of 5.3 years, our analysis revealed that 5% of individuals were reported as deceased, with infection, thrombosis, and malignancies identified as the primary causes of death. History of arterial thrombosis, concomitant SAIDs (mostly lupus), and baseline any CVD risk factor independently predicted future death.

A prospective follow-up study of 1,000 patients with APS from 13 European countries, despite 15% (at 5 years) and 42% (at 10 years) lost-to-follow-up rates, demonstrated that 53 (5%) and 93 (9%) were deceased at 5 and 10 years, respectively, thrombosis (37%) and infections (27%) being the primary contributors to death. The overall survival probability at the end of the 10-year period[4,8] was 91%. Based on another single-center prospective study of 160 patients; the mortality rate was 6.3% (primarily attributed to thrombosis, hemorrhage, and cancer), and the survival probability was 94% at 10 years.[9] Our prospective results support these studies; approximately 5% of persistently aPL-positive patients (91% with APS classification) were deceased during the five-year follow-up.

The number of studies investigating the predictors of death in patients with APS is limited. Based on retrospective follow-up of 114 patients with APS for 38 years (mortality rate 30%), low-to-moderate-level thrombocytopenia (platelet count 50–150 × 10^9/L) was linked to increased mortality.[10] In the systematic review of 16 studies with 1,740 patients, four studies (352 patients) reported that 18 deaths were directly related to recurrent thromboses (12 arterial, 5 venous, and 1 both).[11] In patients with CAPS, concomitant lupus diagnosis increases the risk of death.[12] Although no studies investigated the impact of CVD risk factors on mortality in APS, independent risks for thrombosis in aPL-positive individuals include age, male sex, hypertension, diabetes, and smoking.[13] Our study is the largest cohort study investigating the predictors of death in aPL-positive patients, which demonstrates that history of arterial thrombosis, lupus, and CVD risk factors independently increase the risk of death, even after adjustment for age. These findings build on previous studies, which were often confined to single centers, were retrospective, and/or were based on descriptive or univariate analysis. Thus, our study provides a more comprehensive understanding of the impact of APS on mortality.

Furthermore, our study highlights several important areas in patient care that could significantly influence outcomes in persistently aPL-positive patients. Infection, identified as the leading cause of death, underscores the importance of implementing robust prevention and management strategies. This includes prioritizing immunization schedules, enhancing infection surveillance,

Table 1. Selected baseline demographic and clinical characteristics, by deceased vs living patients (N = 967)*

	Deceased (n = 43)	Living (n = 924)	*P* value
Demographics			
Female, n (%)	28 (65)	695 (75)	0.14
Race, n (%)			
White (n = 690)	29 (67)	628 (68)	0.94
Latin American (n = 93)	5 (12)	92 (10)	0.72
Asian (n = 80)	4 (9)	74 (8)	0.76
Black (n = 25)	3 (7)	27 (3)	0.13
Age at registry entry, mean ± SD, y	54 ± 15	45 ± 13	**0.00**
18–30 (n = 127)	27 ± 2	26 ± 3	0.57
31–44 (n = 340)	38 ± 4	37 ± 4	0.44
45–59 (n = 343)	54 ± 5	52 ± 4	0.07
≥60 (n = 157)	69 ± 6	66 ± 5	**0.02**
Concomitant systemic autoimmune disease, n (%)	29 (67)	432 (47)	**0.007**
Systemic lupus erythematosus, n (%)	23 (54)	306 (33)	**0.005**
Disease duration, mean ± SD, y	16.0 ± 10.3	12.0 ± 10.1	**0.05**
Clinical manifestations			
No APS classification, n (%)	3 (7)	202 (22)	**0.01**
Arterial thrombosis, n (%)	28 (65)	299 (32)	**<0.001**
Venous thrombosis, n (%)	20 (47)	390 (42)	0.58
Microvascular disease,[a] n (%)	5 (12)	56 (6)	0.14
Catastrophic APS,[b] n (%)	2 (5)	9 (1)	**0.03**
Obstetric APS, n (%)	2 (7)	161 (23)	0.05
Thrombocytopenia,[c] n (%)	11 (26)	174 (19)	0.27
Duration since first aPL-related event, mean ± SD, y	10.0 ± 8.2	10.0 ± 8.6	1.0
aPL profile			
Triple aPL positivity, n (%)	11 (26)	312 (34)	0.27
Any LA positivity (except triple aPL), n (%)	10 (23)	158 (17)	0.29
Double aPL positivity (except LA), n (%)	7 (16)	107 (12)	0.35
Single aPL positivity (except LA), n (%)	15 (35)	346 (39)	0.73
Duration since first positive aPL test result, mean ± SD, mo	124 ± 214	86 ± 148	0.11
Medications (baseline), n (%)			
VKA only	20 (46.5)	335 (36)	0.17
LMWH only	1 (2.3)	19 (2)	0.90
VKA plus antiplatelet agent	8 (18.6)	141 (15)	0.55
LMWH plus antiplatelet agent	0	22 (2)	0.31
Antiplatelet agent only	10 (23.2)	281 (30)	0.32
Direct oral anticoagulants	0	30 (3)	0.23
Statin medications	18 (42)	211 (23)	**0.004**
Hydroxychloroquine	24 (55.8)	437 (47)	0.27
Glucocorticoids (ever)	12 (28)	45 (5)	**0.000**
Immunosuppression use (ever)[d]	9 (21)	221 (24)	0.86
CVD risk factors (baseline), n (%)			
Any CVD risk factor	42 (98)	714 (77)	**0.001**
Obesity	16 (37)	232 (25)	0.07
Hypertension with medication	27 (63)	291 (32)	**<0.001**
Diabetes with medication	5 (12)	50 (5)	0.09
Hyperlipidemia with medication	14 (32)	214 (23)	0.16
Hormone replacement	2 (5)	47 (5)	0.90
Nephrotic syndrome	3 (7)	8 (1)	**<0.001**
Chronic kidney disease	8 (19)	30 (3)	**<0.001**
Sedentary lifestyle[e]	28 (65)	360 (39)	**<0.001**
Family history of early CVD	11 (26)	110 (12)	**0.008**
Smoking (active/historical)	17 (39)	312 (34)	0.44

* Bold indicates statistical significance. aPL, antiphospholipid antibody; APS, antiphospholipid syndrome; CVD, cardiovascular disease; LA, lupus anticoagulant test; LMWH, low-molecular-weight heparin; VKA, Vitamin K antagonist.
[a] Diffuse pulmonary hemorrhage, aPL nephropathy, and/or livedoid vasculopathy.
[b] Based on the "definite" or "probable" catastrophic APS classification criteria.
[c] Otherwise unexplained persistent platelet count < 100×10^9/L.
[d] Excluding glucocorticoids and hydroxychloroquine and including azathioprine, mycophenolate mofetil, methotrexate, cyclosporine, rituximab, cyclophosphamide, and intravenous immunoglobulin.
[e] Less than 30 minutes of daily exercise.

Table 2. Primary or secondary causes of death, reported by investigators, in 43 patients

Causes of death	n (%)
Infection with or without sepsis[a]	14 (33)
Venous thromboembolism and/or arterial thrombosis	10 (23)
Malignancy	8 (19)
Other[b]	7 (16)
Bleeding	5 (12)
Unknown	5 (12)

[a] Three patients due to COVID-19.
[b] Congestive heart failure (n = 1), lupus flare (n = 2), multiorgan failure (n = 2), respiratory failure after a fall (n = 1), and aortic dissection (n = 1).

optimizing the use of immunosuppressive therapies, and patient awareness regarding infection risks and the importance of prompt medical attention for early symptoms. The role of CVD risk factors as a major predictor of death emphasizes the critical need for comprehensive cardiovascular risk management. Integrating targeted interventions, such as tobacco cessation programs, intensive blood pressure and cholesterol management, and lifestyle modifications (eg, diet and physical activity), into routine care is essential. Given the association between arterial thrombosis and increased risk of death, personalized approaches to thrombosis prevention, including tailored anticoagulant and/or antiplatelet therapies, should be considered to mitigate individual risks effectively.

Our study has several strengths when compared to previous studies addressing mortality in APS. The prospective nature of the data collection and analysis, with a median follow-up of 5.3 years, enhances the reliability of the findings by allowing for the longitudinal assessment of mortality and associated factors in a sizable international multicenter cohort. The inclusion criterion, requiring persistently positive aPLs according to Revised Sapporo Classification Criteria, contributes to the homogeneity of the study population and ensures the specificity of the findings to individuals with confirmed APS. The use of meticulous statistical methods, including adjusted Cox proportional hazards models and Kaplan-Meier survival analyses, strengthens the robustness of the results, allowing more accurate identification of independent predictors of death compared to the previous studies. The diverse demographic representation of the study population, including survival probabilities categorized by age, makes the findings applicable across different age groups.

Despite its strengths, this study has several limitations that warrant consideration. Selection bias toward survivors should be acknowledged given that those with severe APS clinical phenotypes (eg, CAPS, or severe SAIDs) might not have survived to be enrolled in the registry. Smaller sample sizes in specific age groups could reduce the accuracy of survival estimates and limit the applicability of findings to certain subpopulations, for example, given the small number of patients treated with different immunosuppressives, we could not explore specific immunosuppressive agents in our multivariate model. Relying on investigator-reported causes of death introduces the chance of misclassification or underreporting, impacting the accuracy of identified main causes of death, that is, there was no systematic analysis of vital statistics data given the multicenter international nature of the study. Additionally, the exclusion of individuals with no follow-up may introduce a potential source of bias because those without follow-up may have different characteristics or outcomes (eg, early death rates) than those with follow-up data; however, the clinical characteristics of our patients with and without follow-up data were mostly similar.

In summary, based on the analysis of our international multicenter registry of persistently aPL-positive patients, the mortality rate is relatively low, especially for younger age groups. Given the significant associations between mortality and CVD risk factors, arterial thrombosis, and concomitant SAIDs, our findings underscore the importance of considering these factors in the

Table 3. The association between various baseline characteristics of aPL-positive patients and mortality, adjusted for age*

Patient characteristics (N = 967)	Hazard ratio (95% CI)	*P* value
Concomitant systemic autoimmune disease	2.85 (1.52–5.36)	**0.001**
Clinical manifestations		
Arterial thrombosis	2.99 (1.56–5.74)	**0.001**
Venous thrombosis	1.24 (0.67–2.32)	0.49
Microvascular disease	1.60 (0.67–3.84)	0.29
Catastrophic APS	2.92 (1.49–5.73)	**0.018**
Thrombocytopenia	1.85 (0.65–5.27)	0.25
aPL profile		
Triple aPL positive	0.74 (0.37–1.49)	0.40
Any LA positive (except triple)	1.30 (0.70–2.40)	0.41
Medications		
Statin medications	1.73 (0.88–3.40)	0.11
Hydroxychloroquine	1.57 (0.84–2.92)	0.16
Glucocorticoids	1.94 (0.98–3.81)	0.06
Any CVD risk factor[a]	2.85 (1.23–6.58)	**0.014**

* Bold values indicate significance. aPL, antiphospholipid antibody; APS, antiphospholipid syndrome; CI, confidence interval; CVD, cardiovascular disease; LA, lupus anticoagulant test.
[a] Any CVD risk factor includes obesity, hypertension with medication, diabetes with medication, hyperlipidemia with medication, hormone replacement, nephrotic syndrome, chronic kidney disease, sedentary lifestyle, family history of early CVD, and smoking (active/historical).

Table 4. Independent associations between baseline characteristics of antiphospholipid antibody–positive patients at registry entry and mortality, adjusted for age and each other for four strongest predictors of death*

Patient characteristics (N = 967)	Hazard ratio (95% CI)	*P* value
Arterial thrombosis	2.94 (1.50–5.76)	**0.0017**
Concomitant systemic autoimmune diseases	2.97 (1.56–5.63)	**0.0009**
Catastrophic antiphospholipid syndrome	2.52 (0.57–11.3)	0.23
Any CVD risk factor	2.43 (1.05–5.71)	**0.0414**

* Results were similar when history of glucocorticoid use was also added to the model, with a hazard ratio of 1.20 (95% CI 0.58–2.46, *P* = 0.6205). Bold values indicate significance. CI, confidence interval; CVD, cardiovascular disease.

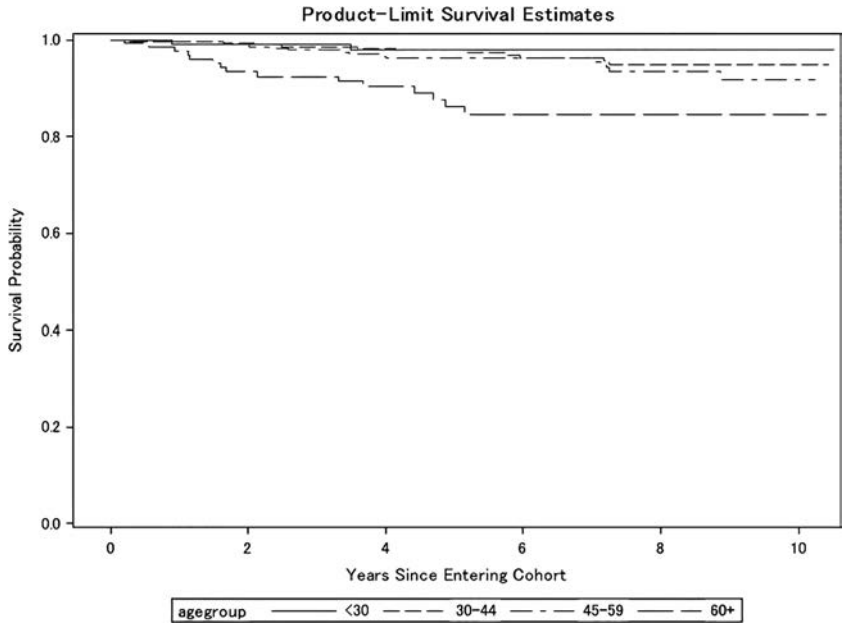

Figure 1. Estimated five-year survival probabilities of antiphospholipid antibody–positive patients (from the time of registry entry), by age groups.

management of patients with APS, contribute to a better understanding of the disease's natural course, and hopefully will aid clinicians in identifying high-risk patients.

ACKNOWLEDGMENTS

The APS ACTION registry was created using REDCap, provided by the Clinical and Translational Science Center at Weill Cornell Medical College (grant UL1-TR-000457). We want to thank JoAnn Vega, CCRC, for her administrative support as the APS ACTION global lead coordinator.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Erkan confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Barbhaiya M, Zuily S, Naden R, et al; ACR/EULAR APS Classification Criteria Collaborators. The 2023 ACR/EULAR antiphospholipid syndrome classification criteria. *Arthritis* Rheumatol 2023;75(10):1687–1702.
2. Rodziewicz MDCD, Gulliford M, Hazra M. The epidemiology of the antiphospholipid syndrome in the UK, 1990 – 2016 [abstract]. Arthritis Rheum. 2019;71(suppl 10).
3. Duarte-García A, Pham MM, Crowson CS, et al. The epidemiology of antiphospholipid syndrome: a population-based study. Arthritis Rheumatol 2019;71(9):1545–1552.
4. Cervera R, Serrano R, Pons-Estel GJ, et al; Euro-Phospholipid Project Group (European Forum on Antiphospholipid Antibodies). Morbidity and mortality in the antiphospholipid syndrome during a 10-year period: a multicentre prospective study of 1000 patients. Ann Rheum Dis 2015;74(6):1011–1018.
5. Erkan D, Sciascia S, Bertolaccini ML, et al; APS ACTION Executive Committee. Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking (APS ACTION): 10-year update. Curr Rheumatol Rep 2021;23(6):45.
6. Bobba RS, Johnson SR, Davis AM. A review of the Sapporo and revised Sapporo criteria for the classification of antiphospholipid syndrome. Where do the revised Sapporo criteria add value? J Rheumatol 2007;34(7):1522–1527.
7. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377–381.
8. Cervera R, Khamashta MA, Shoenfeld Y, et al; Euro-Phospholipid Project Group (European Forum on Antiphospholipid Antibodies). Morbidity and mortality in the antiphospholipid syndrome during a 5-year period: a multicentre prospective study of 1000 patients. Ann Rheum Dis 2009;68(9):1428–1432.
9. Serrano R, Pons-Estel GJ, Espinosa G, et al. Long-term follow-up of antiphospholipid syndrome: real-life experience from a single center. Lupus 2020;29(9):1050–1059.
10. Pardos-Gea J, Marques-Soares JR, Buján S, et al. Persistent thrombocytopenia predicts poor long-term survival in patients with antiphospholipid syndrome: a 38-year follow-up study. Rheumatology (Oxford) 2022;61(3):1053–1061.
11. Ruiz-Irastorza G, Hunt BJ, Khamashta MA. A systematic review of secondary thromboprophylaxis in patients with antiphospholipid antibodies. Arthritis Rheum 2007;57(8):1487–1495.
12. Bucciarelli S, Espinosa G, Cervera R, et al; European Forum on Antiphospholipid Antibodies. Mortality in the catastrophic antiphospholipid syndrome: causes of death and prognostic factors in a series of 250 patients. Arthritis Rheum 2006;54(8):2568–2576.
13. Zuo Y, Barbhaiya M, Erkan D. Primary thrombosis prophylaxis in persistently antiphospholipid antibody-positive individuals: where do we stand in 2018? Curr Rheumatol Rep 2018;20(11):66.

APPENDIX A: APS ACTION MEMBERS

Members of the APS ACTION group are as follows: Guillermo Pons-Estel (Santa Fe, Argentina); Bill Giannakopoulos, Steve Krilis (Sydney, Australia); Guilherme de Jesus, Roger Levy, Flávioavio Signorelli (Rio de Janeiro, Brazil); Danieli Andrade, Gustavo Balbi (Sãoao Paulo, Brazil); Ann E. Clarke, Leslie Skeith (Calgary, Canada); Paul R. Fortin (Quebec City, Canada); Lanlan Ji, Zhouli Zhang (Beijing, China); Chengde Yang, Hui Shi (Shanghai, China); Cecile Yelnik (Lilly, France); Stephane Zuily, Denis Wahl (Nancy, France); Maria G. Tektonidou (Athens, Greece); Cecilia Nalli, Laura Andreoli, Angela Tincani (Brescia, Italy); Cecilia B. Chighizola, Maria Gerosa, Pierluigi Meroni (Milan, Italy); Vittorio Pengo (Padova, Italy); Giulia Pazzola (Reggio Emilia, Italy); Savino Sciascia, Silvia Foddai, Massimo Radin (Turin, Italy); Stacy Davis (Kingston, Jamaica); Olga Amengual, Tatsuya Atsumi (Sapporo, Japan); Imad Uthman (Beirut, Lebanon); Maarten Limper, Philip de Groot (Utrecht, The Netherlands); Guillermo Ruiz-Irastorza (Barakaldo, Spain); Ignasi Rodriguez-Pinto, Ricard Cervera, Jose Pardos-Gea (Barcelona, Spain); Esther Rodriguez Almaraz (Madrid, Spain); Maria Angeles Aguirre Zamorano, Chary López-Pedreraopez-Pedrera (Cordoba, Spain); Bahar Artim-Esen (Istanbul, Turkey); Maria Laura Bertolaccini, Hannah Cohen, Maria Efthymiou, Ian Mackie, Giovanni Sanna (London, United Kingdom); Jason Knight, Yu Zuo (Ann Arbor, Michigan); Michelle Petri (Baltimore, Maryland); Rebecca K. Leaf (Boston, Massachusetts); Elena Gkrouzman (Worcester, Massachusetts); Robert Roubey (Chapel Hill, North Carolina); Thomas Ortel (Durham, North Carolina); Emilio Gonzalez, Rohan Willis (Galveston, Texas); Nina Kello (New Hyde Park, New York); Michael Belmont, Medha Barbhaiya, Doruk Erkan, Steven Levine, Michael Lockshin, Jacob Rand, Jane Salmon (New York, New York); Alíi A. Duarte- Garcíaia (Rochester, Minnesota); D. Ware Branch (Salt Lake City, Utah); and Nicola Pozzi (St Louis, Missouri).

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 767–776
DOI 10.1002/acr.25500

American College of Rheumatology
Empowering Rheumatology Professionals

Physician Views and Attitudes Regarding Tobacco-Cessation Strategies for Patients with Rheumatoid Arthritis

Maria A. Lopez-Olivo,[1] Johncy J. Kachira,[1] Sheneze Madramootoo,[1] Kaleb Michaud,[2] Rebecca Schumacher,[3] Paul Cinciripini,[1] and Maria E. Suarez-Almazor[1]

Objective. In this study, we explored physicians' level of experience with patients with rheumatoid arthritis (RA) who used tobacco, their views on the effects of tobacco use on the efficacy of RA treatments, and their experiences and attitudes with respect to tobacco-cessation programs.

Methods. We conducted qualitative, semistructured interviews of 20 physicians (10 primary care physicians [PCPs] and 10 rheumatologists).

Results. The physicians had been in clinical practice for a mean of 9.9 years. Research themes included (1) risk perception of smoking, (2) cessation aids used, (3) preferences to deliver cessation programs, and (4) barriers and facilitators for tobacco cessation. For the first theme, many PCPs did not perceive smoking as influencing RA disease activity. For the second theme, most physicians supported the use of nicotine-replacement therapy and agreed that cessation-drug therapy (eg, varenicline, bupropion) worked better than nicotine-replacement therapy or other cessation strategies, especially in patients with failed cessation attempts. For the third theme, some physicians recommended that patients join the Quitline cessation program and enroll in peer support communities; others found educational programs informing patients about the benefits of quitting and tailored with messages according to patients' specific clinical characteristics to be useful. For the fourth theme, PCPs and rheumatologists reported similar barriers to offering smoking-cessation programs (eg, lack of time, training in tobacco cessation, and financial motivation).

Conclusion. Physicians agreed with the need for tailored, multifaceted interventions to support tobacco cessation in patients with RA. However, many perceived major barriers to helping their patients quit, some of which could be overcome by training.

INTRODUCTION

Smoking is a risk factor for developing rheumatoid arthritis (RA).[1,2] The risk of developing RA is about 40% higher in smokers compared with never smokers, with the risk increasing as the number of years of smoking increases.[2–4] Tobacco use is also associated with increased RA disease activity and poorer responses to RA treatment.[5] However, compared with the general population of smokers, patients with RA are less likely to quit smoking.[4]

Although interventions like nicotine-replacement therapy (NRT), cessation-drug therapy (eg, varenicline, bupropion), health education, counseling, antismoking advertising, training physicians, and their combinations have been used to help patients stop using tobacco, the programs tested so far have been unsuccessful among most patients with RA, exposing the necessity to better understand the unique needs of this population for giving up the use of tobacco.[4,6] Some studies have found individualized support from physicians to be an important component of cessation programs; such support increases patients' adherence to cessation strategies.[7,8] For patients with RA, primary care physicians (PCPs) and rheumatologists are the first contacts for tobacco management. Therefore, the primary objective of our study was to obtain an in-depth understanding of the current practices among PCPs and rheumatologists with respect to tobacco-cessation referrals and what physicians perceive their patients need. Our overall goal

Supported by the Tobacco Settlement Funds research resource program (to Dr Suarez-Almazor) and National Cancer Institute, NIH (award P30-CA-016672).

[1]Maria A. Lopez-Olivo, MD, PhD, Johncy J. Kachira, MD, MPH, Sheneze Madramootoo, PhD, MBA, MPHIL, Paul Cinciripini, PhD, Maria E. Suarez-Almazor, MD, PhD: The University of Texas MD Anderson Cancer Center, Houston; [2]Kaleb Michaud, PhD: FORWARD, The National Databank for Rheumatic Diseases, Wichita, Kansas, and University of Nebraska Medical Center, Omaha; [3]Rebecca Schumacher, BS: FORWARD, The National Databank for Rheumatic Diseases, Wichita, Kansas.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25500).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25500.

Address correspondence via email to Maria A. Lopez-Olivo, MD, PhD, at amlopezo@mdanderson.org.

Submitted for publication September 5, 2024; accepted in revised form January 5, 2025.

SIGNIFICANCE & INNOVATIONS

- Cessation aids used by primary care physicians (PCPs) and rheumatologists include counseling and medications like nicotine-replacement therapy. However, there were discrepancies in comfort and perceived effectiveness, particularly concerning e-cigarettes.
- Both PCPs and rheumatologists expressed willingness to implement tobacco-cessation programs tailored specifically for patients with rheumatoid arthritis.
- Key barriers identified included limited time for counseling, lack of access to cessation medications, and patients' resistance to quitting.

is to develop a robust, web-based tobacco-cessation program tailored to the needs of patients with RA.

METHODS

Qualitative approach and research paradigm. This qualitative study has been reported according to the Standards for Reporting Qualitative Research.[9] We used a phenomenological approach to understand participants' experiences and perceptions of tobacco-cessation strategies for patients with RA. We used a constructivist paradigm to identify what was important to them and their patients.

Researcher characteristics and reflexivity. The research team included two nonpracticing physicians with experience in qualitative methods (MAL-O, JJK), one rheumatologist with expertise in epidemiologic designs and qualitative methods (MES-A), one expert in tobacco-cessation research (PC), two rheumatology researchers who manage a registry for patients with RA (KM, RS), and one research data coordinator with experience in qualitative research (SM). Two of the researchers had collaborated in research activities unrelated to this study's topic with four of the rheumatologists interviewed. The interviewer (JJK) was not familiar with any of the physicians interviewed.

Context and sampling strategy. Study staff worked in a hybrid format, and most study activities were done remotely. We interviewed physicians from three different practices in Houston, Texas, and its metropolitan area affiliated with academic centers, which include both hospital-based and community clinics. The names of all the PCPs and rheumatologists from each practice site were entered into a spreadsheet (2 separate lists of names) and contact information for each physician was obtained from each practice website. We excluded trainees, faculty with administrative positions, and nonphysician health care providers. Out of 582 potentially eligible participants, we generated a randomly ordered list of 300 physicians, which included individuals from 104 family practices, 63 multispecialty practices, and 133 care centers/hospitals. Physicians were contacted via email in the order they appeared on the list. Every week, 10 eligible physicians were contacted. Study staff made three weekly attempts to reach the participants, and if the participants refused to take part in the study or could not be reached, the next candidates on the randomly ordered list were contacted. On the basis of previous studies and our own experience, we believed that up to 20 interviews with physicians would be sufficient to exhaust the most relevant themes.[10]

Ethical issues pertaining to human participants. This study was approved by the Institutional Review Board of The University of Texas MD Anderson Cancer Center (protocol #2021-0760). The risk of incidents arising from the interviews was considered very low. Participants verbally consented to take part in the study and were compensated for their participation. The authors had ample experience conducting research in health care.

Data collection methods and instruments. Interviews were completed from August 2022 to February 2023. We used a semistructured interview format developed by the research team and informed by the findings of prior research on the topic.[4,8] The full interview guide is included in Supplemental Material 1. Participants were asked about their experiences with patients who were tobacco users, including their observations on the prevalence of smoking in patients with RA, their thoughts on the effects of smoking on the treatment response, and their experiences with and views on smoking-cessation programs. The questions were designed to be open ended to encourage participants to be thoughtful and to consider all aspects of their consultations with patients with RA that were important. Interviews lasted about 20 to 30 minutes on average and were conducted via Zoom or over the phone. Participants were allowed to ask questions during the interview if something was unclear. During the conduct of the study, no modifications to the data collection methods or instruments were needed.

Data collection. The interview questions and answers were recorded (audio only), and self-reported information about participants, including their demographic characteristics (ie, age, sex, self-reported race and ethnicity), practice characteristics, and years in practice, were collected. We used the NIH terminology for racial and ethnic categories.[11] The recorded sessions were transcribed verbatim, and physicians' names were replaced with codes for deidentification purposes.

Data processing and units of analysis. The transcribed interviews were entered into Dedoose, a cross-platform app, to facilitate coding and analysis.[12] Two investigators (JJK and SM) crosschecked all transcriptions against the audio files to verify

data integrity and become familiar with the raw data. The units of data were the statements made in the participants' own words. We labeled the interviews with each physician's medical specialty.

Data analysis. We organized, sorted, and interpreted the physicians' statements using a deductive approach. As a first step, the responses contained in three transcripts were independently coded by two coders (JJK and SM) and sorted into subcategories according to the guiding questions of our interviews. Responses were compared, and disagreements in terms of how a response should be categorized were managed through consensus or through discussion involving a third author (MAL-O). A preliminary report was created to identify and define the subthemes that emerged from the data analysis. The report was sent to the research team for feedback, and the team then created a focused coding handbook that refined and expanded upon the text. Thereafter, the coding handbook was used to code the remaining transcripts.[1]

Techniques to enhance trustworthiness. To enhance the credibility of this study, we interviewed PCPs (ie, family medicine, internal medicine, geriatric medicine physicians) and rheumatologists with varying demographic traits and years of experience. The participants were affiliated with organizations providing various types of health care. The heterogeneity of our sample enabled us to collect data on a diversity of experiences and perceptions of physicians and thereby enhanced the applicability of our findings to similar contexts. Additionally, we created a comprehensive description of our study procedures and approach to data analysis to further augment the dependability of our study and its potential for replication. Our data were checked and rechecked throughout the data collection and analysis process, and we developed a record of changes to the coding handbook to improve the reliability of our findings.

RESULTS

We interviewed 20 physicians: 10 PCPs and 10 rheumatologists. The mean age of the participants was 42.6 ± 7.1 years. Three participants (14%) were ethnically Hispanic, 14 of the participants (70%) were female participants, and 11 of the participants (55%) were Asian. The mean time in clinical practice was 9.9 ± 7.5 years. Sixteen participants (80%) described their practice as primarily academic. The average percent of time spent seeing patients per week was 67.5% ± 27.7%. On average, participants saw 25.4 ± 26.8 patients with RA per month (PCP range 4–20; rheumatologist range 15–100). Table 1 shows the participants' characteristics by specialty. We classified physicians' responses into 1,778 codes according to our research objectives: (1) risk perception of smoking, (2) cessation aids used, (3) preferences for the delivery of cessation programs, and (4) barriers to and facilitators for tobacco cessation.

Risk perception of smoking. Physicians shared their views on the role of smoking in the development of RA and on the RA course and prognosis in patients who smoke. Figure 1 summarizes the topics discussed under each subtheme.

Prevalence of smokers in RA population. The physicians had differing opinions about the prevalence of tobacco use in people with RA (ie, some perceived a high or low prevalence, and others were unsure). Most PCPs agreed that smoking has a negative effect on health; however, a few were unsure about there being a clear link between smoking and RA development, considering the potential multifactorial etiology of RA and the contradictory evidence in the literature, and expressed that further research is needed. Some rheumatologists believed that smoking might affect autoimmunity and the development of anticitrullinated protein antibodies and thus change the immune system. However, others were uncertain about an association between smoking

Table 1. Participants' characteristics by specialty*

Characteristic	Entire cohort (n = 20)	Primary care physicians (n = 10)[a]	Rheumatologists (n = 10)
Age, mean (SD)	42.6 (7.1)	41.2 (5.8)	44 (8.4)
Female participants, n (%)	14 (70)	9 (90)	5 (50)
Hispanic or Latino ethnicity, n (%)	3 (15)	2 (20)	1 (10)
Race, n (%)			
Asian	11 (55)	3 (30)	8 (80)
Black or African American	2 (10)	1 (10)	1 (10)
White	6 (30)	5 (50)	1 (10)
More than 1 race	1 (5)	1 (10)	—
Years of clinical practice, mean (SD)	9.9 (7.5)	8.3 (6.7)	11.5 (8.2)
Type of practice, n (%)			
Academic	16 (80)	7 (70)	9 (90)
Public	1 (5)	1 (10)	—
Private	3 (15)	2 (20)	1 (10)
Time in clinic, mean (SD)	67.5 (27.7)	73 (25.0)	62 (30.6)
Patients with RA seen per month, mean (SD)	25.4 (26.8)	8.1 (4.7)	42.6 (28.8)

* RA, rheumatoid arthritis.
[a] Including family medicine and internal medicine physicians.

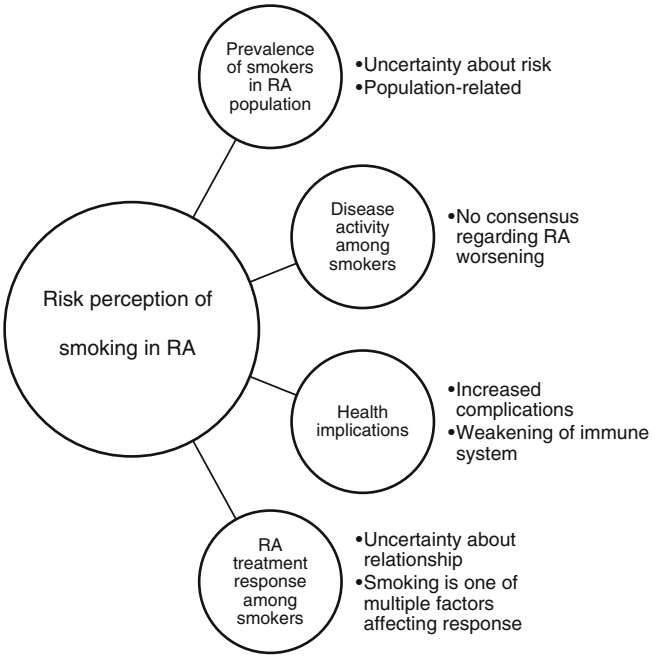

Figure 1. Mapping of topics discussed in the "risk perception of smoking" theme. RA, rheumatoid arthritis.

and autoimmunity, given that they see patients after RA has developed, but most considered that a moderate to high percentage of the people that they had diagnosed with RA smoked at the time of diagnosis (ranges expressed were between 10 and 50 percent) or had started the process of quitting.

Disease activity among smokers. Many PCPs did not perceive smoking as influencing RA disease activity. A few PCPs mentioned not being surprised that the immune dysregulation associated with smoking could potentially lead to chronic inflammation and a weakening of the immune system. Others recognized that patients with RA who smoke have a higher risk of developing heart disease and/or other comorbidities that contribute to worsening health and a decreased quality of life. Contrastingly, most rheumatologists perceived that patients with RA who smoke have a more rapid disease progression, usually with worse outcomes, and are especially likely to have higher pain levels compared with nonsmokers.

Health implications. All physicians acknowledged the potential for an increased risk of complications such as infections, cancer, and the worsening of comorbid conditions among smoking compared with nonsmoking patients with RA, stating that patients with lung disease typically do worse overall.

RA treatment response among smokers. There were also discrepant views regarding the treatment response in patients with RA who smoke. Some PCPs considered smoking as one of multiple factors affecting the treatment response, whereas others noted no correlation. Most rheumatologists described a direct relationship between a suboptimal response to treatment, a continued need to switch therapies, and difficulty in achieving remission. A few described an association between smoking and depression leading to medication noncompliance.

Use of cessation aids. Physicians discussed their roles in the management of tobacco use, including the different strategies they were using (eg, discuss the consequences of smoking, ask about readiness to quit, offer counsel, and adopt the STAR [situation, task, action, and result] method: set a quit date, tell everyone, anticipate challenges, and remove tobacco products from the environment). Figure 2 shows a list of the cessation aids physicians used in their current practice. Assessing the patients' readiness to quit without judgment, discussing the negative consequences of smoking, and offering medications to help patients quit smoking were the most used options. Most physicians supported the use of NRT, perceiving nicotine gums, patches, and e-cigarettes as effective. However, some PCPs and rheumatologists felt uncomfortable prescribing NRT because of their limited experience with it, perception that NRT has limited effectiveness, and concerns over continued physical dependence on nicotine among patients. For patients struggling with quitting, some physicians thought that getting patients to cut back was a more feasible solution. Others discussed the use of e-cigarettes for this purpose. However, there were discrepancies between the PCPs and rheumatologists regarding how e-cigarettes were used. Some PCPs thought that the effects of e-cigarettes could be similar to those of conventional cigarettes. Some rheumatologists stated that e-cigarettes could be used to transition, but most believed that there is limited evidence on the risks associated with their use and their potential benefit in aiding cessation attempts and perceived that there is a stigma associated with e-cigarette use. Most physicians had limited experience with tobacco chewing among their patients. Some PCPs and rheumatologists were concerned about their patients with RA chewing tobacco, and a few felt that, although the potential worsening of RA activity could be less with tobacco chewing compared with conventional cigarettes, there were multiple potential risks associated with this behavior such as high risk of oral cancer, tooth decay, cardiovascular problems, and decreased overall quality of life. Both PCPs and rheumatologists agreed that cessation-drug therapy (eg, varenicline, bupropion) worked better than NRT, e-cigarettes, or counseling, especially in patients with failed attempts to stop smoking.

Preferences for the delivery of cessation programs. Both rheumatologists and PCPs thought that smoking-cessation programs could be delivered by any trained personnel across all medical specialties. PCPs preferred to focus on counseling and referrals and to have their clinic staff screen for and document tobacco-use status. Most rheumatologists preferred to refer patients with RA to primary care or cessation programs for further counseling and follow-up. However, PCPs agreed that rheumatologists might be perceived by patients with RA as having more

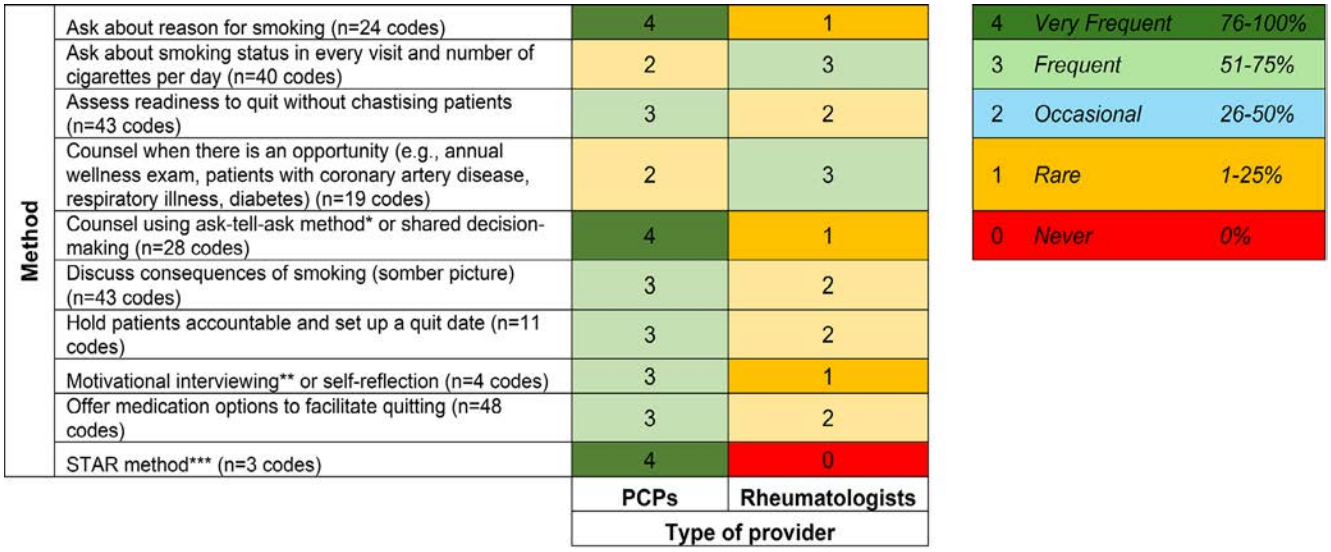

Figure 2. Cessation methods used by the physicians. *Ask what they know and what they want to know; tell them what they want to know; ask them if they understand and what else they want to know. **The physician can motivate patients to consider a quit attempt with the "5 R's": Relevance, Risks, Rewards, Roadblocks, and Repetition. Relevance: encourage the patient to indicate why quitting is personally relevant. Risks: ask the patient to identify potential negative consequences of tobacco use. Rewards: ask the patient to identify potential benefits of stopping tobacco use. Roadblocks: ask the patient to identify barriers or impediments to quitting. Repetition: the motivational intervention should be repeated every time an unmotivated patient has an interaction with a physician. Tobacco users who have failed in previous attempts to quit using tobacco should be told that most people make repeated attempts to quit before they are successful.[25] ***Set a quit date within 2 weeks. Tell family, friends, and coworkers about quitting and request understanding and support. Anticipate challenges to quitting smoking, particularly in the first few weeks. Remove tobacco products from your home, car, and office. PCP, primary care physician; STAR, situation, task, action, and result.

trustworthy information. Most PCPs preferred to use cessation methods according to what methods would be most likely to motivate patients; for example, apps could be used to encourage patients with RA to reduce disease flares by quitting. Many rheumatologists preferred programs with frequent check-ins because they considered check-ins as an important element of setting specific goals and increasing patients' accountability. A few physicians recommended using Quitline—a telephone counseling service—and enrolling in peer support communities, whereas others thought that offering educational programs would be useful in informing patients about the benefits of tobacco cessation, tailoring messages according to patients' specific clinical characteristics, and reducing the number of cigarettes gradually rather than advocating for complete and sudden abstinence. Multiple physicians were not familiar with Quitline. One PCP expressed frustration and disappointment, having previously found it difficult to navigate the system and obtain the help needed.

Physicians also expressed their preferences about the frequency in which advice to quit tobacco use should be given. In general, both types of physicians preferred to give a reminder regarding the benefits of quitting and the resources available to participants at least once per year. Some PCPs mentioned purposely avoiding asking about smoking cessation during each visit because it might aggravate patients. PCPs preferred to ask about smoking at patients' annual wellness visits, and rheumatologists preferred to ask every 6 months. Many PCPs and rheumatologists thought that placing posters, pamphlets, and/or other resources (eg, printouts from smokefree.gov or a Texas hotline) in the examination room or waiting room could replace addressing the importance of tobacco cessation during every visit.

We also asked physicians what type of tobacco-cessation information they preferred to deliver. Both PCPs and rheumatologists were in favor of tailoring the information to the patients' clinical characteristics (eg, disease state, comorbidities, age) and psychosocial needs (eg, motivation, mental health, previous failed attempts to quit). Participants were asked if a website or app could be useful to patients looking to cease smoking, and many were in favor of websites or apps with simple, user-friendly, and engaging information. They listed other features that could enhance patients' experiences such as interaction, patient testimonials, the ability to print information or summaries of recommendations to cut out tobacco use, links to YouTube videos from reputable sources, live coaching sessions/questions and answers/check-ins, and the creation of an online, multilingual, and accessible community for peer support. Some physicians suggested that having access to in-person and online resources will allow accommodation for patients' personalities; some people prefer one-on-one interactions, whereas others prefer phone/online sessions.

Barriers to and facilitators for tobacco cessation. PCPs and rheumatologists agreed on the perceived barriers to tobacco cessation (eg, a lack of time to provide counseling, a lack

of cessation medications or NRT access, patients' resistance to medical advice, and the perception of smoking as a coping mechanism). Although a few rheumatologists acknowledged that the social and emotional health of patients with RA could be barriers for successful smoking cessation, most rheumatologists believed that the barriers to quitting smoking were essentially the same among patients with RA who smoke and the general public (eg, weight gain after tobacco cessation, a lack of desire to quit, a lack of trust in their physicians, a lack of motivation to quit, discouragement after multiple failed attempts to quit, and dependency on smoking). PCPs cited additional barriers to tobacco cessation, including challenges with compliance given the potential added number of medications and fear of drug interactions and side effects, a lack of understanding of the added risks of tobacco use based on their condition, and specific screening for patients with RA. A few PCPs were concerned about the lack of targeted education materials or referral programs for patients with RA.

Both types of physicians identified several facilitators and motivators that support their delivery of tobacco-cessation programs. Facilitators include counseling, incentives, financial benefits, teaching patients to recognize and avoiding socioenvironmental triggers (such as social circles in which smoking is common or places associated with smoking, like bars), using previsit questionnaires to screen patients, and obtaining tobacco-cessation training offered as continuing medical education. Proposed motivators to help patients quit include peer support, motivational messaging, awareness of health risks and benefits associated with quitting, and life-changing events (such as upcoming social engagements, chronic disease diagnoses, or deaths). Most PCPs agreed that tailored tobacco-cessation programs and making patients aware that they should stop smoking were necessary. All physicians were willing to disseminate a newly developed program specifically tailored for patients with RA. Table 2 lists physicians' barriers, facilitators, and motivators regarding the offering of tobacco-cessation programs specifically tailored for patients with RA.

DISCUSSION

The primary objective of this study was to explore physicians' perceptions of smoking-related risks, cessation aids, program-delivery preferences, and barriers to and facilitators for tobacco cessation, specifically in the context of patients with RA. We found that physicians recognized the complex interplay between smoking and RA, including its impact on disease progression and the treatment response. However, there was variability in physicians' perceptions of smoking prevalence and the direct effects of tobacco use on RA development. Our findings suggest that, although cessation aids such as NRT and cessation-drug therapy are valued, physicians are uncomfortable about prescribing them because of NRT's perceived limitations in effectiveness and physicians' limited experience with these methods of tobacco cessation. The study also highlighted physicians' preference for personalized, multifaceted cessation programs delivered through both traditional and digital means.

PCPs and rheumatologists had differing views on smoking's role in RA. PCPs were less certain about the direct link between smoking and RA, whereas rheumatologists were more inclined to believe that smoking exacerbates RA through immune system disruptions and autoimmunity. We also observed variability in the perception of smoking's impact on disease activity and treatment response. PCPs generally did not perceive smoking as influencing RA disease activity, whereas rheumatologists linked smoking to more rapid disease progression and poorer treatment responses. This may be explained because PCPs often manage a broader range of conditions and may not see the full spectrum of RA-related complications, leading to less specific focus on smoking's impact. Rheumatologists are more likely to observe and connect smoking with disease progression and outcomes directly. This discrepancy suggests a need for improved education and communication between PCPs and rheumatologists regarding the specific impacts of smoking on RA.

Although physicians recognized the benefits of cessation programs, not all agreed that they should have a role in counseling patients regarding tobacco cessation. Both PCPs and rheumatologists believed that the PCPs should be responsible for helping patients with RA reduce their use of tobacco. Nonetheless, the PCPs believed that the rheumatologists were in a unique position to increase the likelihood of patients being willing to be educated about cessation programs, given the trust their patients with RA have in them. According to the World Health Organization, all physicians should promote tobacco-cessation programs.[13] Many barriers that the rheumatologists cited in helping their patients quit could be overcome by training. In fact, any physician could be trained to deliver three to five minute tobacco interventions such as the 5A's brief intervention model, which advises physicians to "Ask about smoking, advise cessation, assess level of readiness to quit, assist with motivation or a plan to quit (including use of tobacco-cessation medication and counseling), and arrange follow-up appointment[s] to review progress and adjust the plan."[14] Alternatively, the simpler Ask-Advice-Connect approach involves asking about smoking status, advising cessation, and directly connecting patients to a Quitline.[15] These types of interventions could also be facilitated by an evidence-based strategy specifically developed to support tobacco cessation in patients with RA. A tailored, multifaceted intervention addressing the barriers to and motivators for tobacco cessation highlighted in our study, such as NRT access, incentives, counseling, peer support, and motivational messaging, would be helpful for physicians and patients. PCPs also faced unique challenges such as managing drug interactions and lack of RA-specific resources. Addressing these barriers requires systemic changes, such as improving access to cessation aids and providing targeted education and can lead to a more unified approach to smoking cessation in RA management, improve physician

Table 2. Physicians' barriers, facilitators, and motivators regarding the offering of tobacco-cessation programs*

Barriers	Subcategories	Sample quotes
Knowledge	Lack of understanding regarding the risk of e-cigarettes and smoking	"*I think it's a lack of knowledge. I think that. Because, as a physician, also this—when this study, when this came to me, uh, association of smoking and rheumatoid arthritis, I think more awareness needs to be created in this time when it comes to smoking.*"
	Lack of understanding of added risks based on their condition	"*They don't really quite know specifically that smoking's also bad for their rheumatoid arthritis, too.*"
		"*Education … Education, addiction, denial. I don't know if those are true barriers or not, but….*"
	Lack of adequate (targeted) education materials or referral programs for people with RA	"*The first thing that comes to our mind is, of course, the lung cancer … and plus the COPD, and… asthma…and…emphysema and these things. But rheumatoid arthritis and smoking is not…a very well-known, aware… fact. For sure. Yeah.*"
Beliefs about consequences	Fear of drug interactions and side effects	"*They have concerns about drug-drug interactions, whether medications approved for smoking cessation are safe to take with their…disease-modifying antirheumatic…drugs. So,…those questions do come up.*"
	Weight gain after tobacco cessation	"*…some patients also have concerns about weight…and they worry that quitting smoking might result in weight gain, and when they already have…arthritis of any kind, including rheumatoid arthritis, they're nervous about the impact on their—on their joints, and level of functioning.*"
	Compliance given the added number of medications, fear of side effects	"*And then Chantix has so much negative publicity regarding causing cancer even though, of course, you know, smoking increases…13-fold…as opposed to Chantix, because Pfizer recalled Chantix and also, you have that black box warning of…suicide… Many patients are hesitant to start it… Yes, you'll get them on the starter pack, but beyond that, it's not like they're running back for the…continuation, you know? It's…truly is a challenge.*"
Behavioral regulations	Dependency on smoking	"*Unfortunately, the success rate is not high …the reason that is… I think that there aren't necessarily a lot of programs that I know of, right? Um, I think social work is really the…one we have. And also, there is so much… negative connotation about the treatment for smoking cessation… I'll give you an example. If you put a patient on nicotine patch, the misconception is they're gonna smoke while they have a nicotine patch on, to get more nicotine… To almost have, like, a little high or however people describe it.*"
Intentions	Lack of desire to quit	"*Um, I sometimes refer them to [..] resources, but the majority of my patients are not necessarily interested in…programs.*"
Emotions	Perception of smoking as a coping mechanism	"*…even mental health issues among that population, but I certainly think that having more medical comorbidities plus or minus additional mental health concerns may…make smoking cessation more challenging.*"
Professional roles	Counseling not considered part of the role	"*But I certainly think it's appropriate to…leave the responsibility of discussing medication treatment options to the primary care provider. That's certainly within their scope of practice and responsibilities.*"
	Lack of skills	"*I don't wanna do it. I don't have time to do it. But that's why I refer out and let the experts take care of that.*"
Environmental contexts and resources	Lack of time	"*I think primary care is…where there is space for that motivational interviewing.*"
	Loss of follow-up	"*You know, I think the loss of follow-up is why we don't probably have that success rate.*"
	Specific screening for patients with RA	"*But I don't have a specific…set of recommendations or…treatments.*"
	Avoiding environmental influences	"*Again, their social environment and their peers may influence how they elect to use tobacco or alcohol or other drugs. … so I think that's a unique challenge for younger patients.*"
	Lack of a supply of cessation medications	"*So, …that's, uh, very tricky now, but you know when Chantix was available—and… I just checked it today, it's still showing that it's not back in the market….*"
	Financial constraints	"*Insurance will not cover nicotine-replacement therapy but would cover Chantix, or same with some individuals' actual employers have quit programs who do cover specifically…Chantix or Wellbutrin, but for some reason no…nicotine-replacement therapy. So that is often another consideration when choosing which form to use.*"

(Continued)

Table 2. (*Cont'd*)

Barriers	Subcategories	Sample quotes
	Age	"*The older and socially…socioeconomic barriers are definitely…barriers to…send them to such resources.*"
	Lack of trust in medical providers	"*I basically just give them a hard time about it, I tell them hey you know that's bad for you, you know that will give you a stroke and a heart attack and you know gives you lung cancer and they say I know I gotta die from anything anyway.*"
Facilitators	**Code**	**Sample quote**
Beliefs about capabilities	Empowerment	"*I feel that if patients can learn to be empowered in learning how to quit smoking, they might actually learn to get a better handle on their disease, so, it's very much connected.*"
Social influences	Patients' resistance to medical advice	"*Well, first you have to—education and then getting the patient's trust, okay. So I think in order to get them to accept the education, if you to earn their trust first. Okay, so develop a rapport and relationship with the patient, let them know that you care and you're listening to them. … and you wanna treat their problem, help them as much as possible. Once you get to that point, then you start to make those suggestions about personal habits that may…adversely affect their health.*"
Reinforcements	Counseling and adequate referral programs	"*Unfortunately…I don't have a readily accessible tobacco-cessation counselor… at my fingertips to refer patients to. So I'm not aware of any local resources where patients can get free in-person tobacco-cessation counseling. Although that would certainly be a wonderful addition to…the resources that I'm offering patients.*"
	Incentives and financial benefits	"*But I have even compared the price of the patch and the pack's cigarette price and patch is still…economical.*"
Skills	Not being discouraged after multiple failed attempts to quit using tobacco	"*A lot of times, I don't feel like telling them they have to, as a way to get them to. It's just about constantly asking them if they're ready to discuss quitting, but telling them they have to quit…I'm not sure if that leads to more smoking cessation, rather than checking in to see if they're ready to even talk about quitting.*"
	Previsit questionnaires	"*Um, what I have found that has worked, um, in other—in—in the previous institution where I worked was including it as part of the previsit screen in MyChart prior to a patient coming to the clinic.*"
Motivators	**Code**	**Sample quote**
Social roles	Social and peer support	"*I don't think it's that successful unless a patient is super motivated, or the people around them are really motivated to get their…parents or…their spouses to quit smoking.*"
Goals	Harm reduction	"*I think yes, they're motivated to cut back, but to completely stop smoking, you know, unless there's a chronic illness that has developed, be it…COPD, be it…cancer…*"
	Decrease number of cigarettes	"*I mean not everybody is the same. I know people who they say…they can quit anytime they want to and they really can, but there are other people who are super addicted to smoking and they can't quit. I'll ask them so did you reduce it…are you down from like a whole pack a day to like a half a pack a day.…*"
	Motivational messaging	"*It's not interviewing, but something that can help patients imagine their…future self without smoking. I think that's…imperative. And then, showing them the options of what's available, maybe a brief description of why, you know—the differences between medical treatments. And then referring them back to their primary care provider, saying, 'Hey, you know, you can address this and this and this with your PCP and ask them these—the following questions.' So… prepare them for…their visit with a PCP.*"
Triggers	Life-changing events	"*I think honestly, the best thing which is…kinda sad when an adverse events happens to the patient, whether that's an MI or a stroke. I've noticed that…what ends up…causing patients to quit or some sort of…pain that they have like a patient had really bad shoulder pain and then she ended up needing some sort of surgical intervention and the surgeon wouldn't operate unless she quit smoking…, so that really motivated them and then they quit smoking.*"

* COPD, chronic obstructive pulmonary disease; MI, myocardial infarction; PCP, primary care physician; RA, rheumatoid arthritis.

education, and ultimately enhance patient care outcomes. Two recent studies evaluated the implementation of systematic screening for tobacco use with facilitated referrals to a state-supported tobacco Quitline. These interventions, which included training clinic staff, using electronic health record prompts, and establishing a streamlined referral process, significantly increased referrals to cessation services in a rheumatology clinic environment.[16,17]

In line with our findings, a US survey indicated that 15% of tobacco users were motivated to quit smoking by their doctor's advice.[18] Brief advice from medical practitioners has been shown to boost smoking-cessation rates.[19,20] In the United States, patients with RA were more likely to attempt quitting if they were new to rheumatology care, but smoking-cessation counseling was noted in only 10% of visits, particularly in patients with well-controlled disease.[21] A 2016 US study revealed that although 80% of clinical staff assessed smoking status, no follow-up occurred because of barriers like a lack of referral processes and discomfort discussing cessation.[22] Similarly, a Canadian study found that 78% of nurses and rheumatologists felt unprepared to help patients quit smoking because of limited access to resources and expertise.[23] Common facilitators of smoking cessation included patients' use of NRT and their readiness to quit, whereas cigarette pricing and visible health effects were major influences on their decision to stop.[7,24]

This study's novelty lies in its qualitative exploration of physician perceptions across multiple dimensions of smoking cessation within an RA context. It provides insights into the nuanced understanding and diverse opinions among PCPs and rheumatologists, highlighting areas of consensus and divergence. Our comprehensive approach allowed us to capture a wide range of physician perspectives that can inform the development of tailored cessation programs for patients with RA. Filling the knowledge gap in understanding physicians' perceptions of smoking cessation in patients with RA is crucial, as this would support the development of more effective and targeted interventions for these patients. Given the significant health risks associated with smoking in patients with RA, improving cessation strategies can lead to better disease management and patient outcomes.

The study has limitations. The qualitative nature of the research may limit the generalizability of findings because of potential selection bias and the subjective nature of physicians' perceptions. Additionally, the study was conducted within a specific health care context and may not fully represent the views of physicians in different regions or health care systems. The perspectives of community physicians, who may face different challenges and have different resource availability, have not been explored in this study. Academic physicians often have access to specialized resources and may be more engaged in research and evidence-based practices, which can influence their views and approaches to smoking cessation. Consequently, their experiences and opinions might not fully represent the broader range of practices and attitudes found in community settings. Including viewpoints from community physicians is essential to develop a comprehensive understanding of smoking-cessation strategies and barriers across diverse health care environments. In addition, future research should focus on exploring patient perspectives on cessation aids and program preferences; these would complement the physician insights gained in this study and inform the development of comprehensive, patient-centered cessation strategies. Longitudinal studies examining the long-term effects of tailored cessation programs on RA outcomes would also be valuable.

In conclusion, this study highlights the variability in physician perceptions and existing barriers regarding smoking cessation in patients with RA, emphasizing the need for tailored, multifaceted approaches to addressing this critical health issue. By enhancing their understanding of patients' tobacco-cessation needs and providing targeted interventions, physicians can better support patients with RA in their cessation efforts, ultimately improving their quality of life and health outcomes.

ACKNOWLEDGMENTS

We thank Laura L. Russell and Erica Goodoff, scientific editors, Research Medical Library, for editing this article.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Lopez-Olivo confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Di Giuseppe D, Discacciati A, Orsini N, et al. Cigarette smoking and risk of rheumatoid arthritis: a dose-response meta-analysis. Arthritis Res Ther 2014;16:R61.
2. Hutchinson D, Shepstone L, Moots R, et al. Heavy cigarette smoking is strongly associated with rheumatoid arthritis (RA), particularly in patients without a family history of RA. Ann Rheum Dis 2001;60:223–227.
3. Romão VC, Fonseca JE. Etiology and risk factors for rheumatoid arthritis: a state-of-the-art review. Front Med (Lausanne) 2021;8:689698.
4. Lopez-Olivo MA, Sharma G, Singh G, et al. A systematic review with meta-analysis of the effects of smoking cessation strategies in patients with rheumatoid arthritis. PLoS One 2022;17:e0279065.
5. Hyrich KL, Watson KD, Silman AJ, et al; British Society for Rheumatology Biologics Register. Predictors of response to anti-TNF-α therapy among patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register. Rheumatology (Oxford) 2006;45:1558–1565.
6. Hersi M, Traversy G, Thombs BD, et al. Effectiveness of stop smoking interventions among adults: protocol for an overview of systematic reviews and an updated systematic review. Syst Rev-London 2019;8.
7. Aimer P, Stamp LK, Stebbings S, et al. Exploring perceptions of a rheumatoid arthritis-specific smoking cessation programme. Musculoskeletal Care 2018;16:74–81.

8. Lopez-Olivo MA, Michaud K, Schumacher R, et al. Smoking cessation patterns, usefulness of quitting methods, and tobacco cessation motivators and barriers to quit in patients with rheumatoid arthritis. Clin Rheumatol 2023;42:2053–2068.
9. O'Brien BC, Harris IB, Beckman TJ, et al. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014;89:1245–1251.
10. Luckman S. Successful qualitative research: a practical guide for beginners. Fem Psychol 2016;26:387–391.
11. Office of Management and Budget Standards. Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. Office of Information and Regulatory Affairs 1997. Accessed December 2024. https://www.govinfo.gov/content/pkg/FR-1997-10-30/pdf/97-28653.pdf
12. *Dedoose version 7.0.23*. Web application for managing, analyzing, and presenting qualitative and mixed method research data. Los Angeles, CA: SocioCultural Research Consultants, LLC. Accessed December 2024. www.dedoose.com
13. World Health Organization. Toolkit for delivering the 5A's and 5R's brief tobacco interventions in primary care. Geneva, Switzerland 2014. Accessed July 23, 2024. https://iris.who.int/bitstream/handle/10665/112835/9789241506953_eng.pdf?sequence=1
14. Agency for Healthcare Research and Quality. The "5 A's" Model for Treating Tobacco Use and Dependence. Content last reviewed October 2014, Rockville, MD, 2008. Accessed December 2024. https://www.ahrq.gov/prevention/guidelines/tobacco/clinicians/presentations/2008update-overview/slide43.html
15. Vidrine JI, Shete S, Cao Y, et al. Ask-Advise-Connect: a new approach to smoking treatment delivery in health care settings. JAMA Intern Med 2013;173:458–464.
16. Bartels CM, Johnson L, Ramly E, et al. Impact of a rheumatology clinic protocol on tobacco cessation quit line referrals. Arthritis Care Res (Hoboken) 2022;74:1421–1429.
17. Brandt J, Ramly E, Lim SS, et al. Implementing a staff-led smoking cessation intervention in a diverse safety-net rheumatology clinic: a pre-post scalability study in a low-resource setting. Arthritis Care Res (Hoboken) 2024;76:1342–1350.
18. Mai Y, Soulakova JN. Retrospective reports of former smokers: receiving doctor's advice to quit smoking and using behavioral interventions for smoking cessation in the United States. Prev Med Rep 2018;11:290–296.
19. Fiore MC, Jaén CR. A clinical blueprint to accelerate the elimination of tobacco use. JAMA 2008;299:2083–2085.
20. Stead LF, Buitrago D, Preciado N, et al. Physician advice for smoking cessation. Cochrane Database Syst Rev 2013:CD000165.
21. Vreede AP, Johnson HM, Piper M, et al. Rheumatologists modestly more likely to counsel smokers in visits without rheumatoid arthritis control: an observational study. J Clin Rheumatol 2017;23:273–277.
22. Panyard D, Ramly E, Gilmore-Bykovskyi A, et al. Developing a staff-driven electronic smoking cessation referral program in rheumatology clinics. Arthritis Rheumatol 2016;68.
23. Karacabeyli D, Shojania K, Dehghan N, et al. Rheumatology health care providers' views and practices on obesity and smoking cessation management in rheumatoid arthritis. Clin Rheumatol 2021;40:2681–2688.
24. Wattiaux A, Bettendorf B, Block L, et al. Patient perspectives on smoking cessation and interventions in rheumatology clinics. Arthritis Care Res (Hoboken) 2020;72:369–377.
25. Agency for Healthcare Research and Quality. Patients not ready to make a quit attempt now (The "5 R's"). Rockville, MD 2012. Accessed July 22, 2024. https://www.ahrq.gov/prevention/guidelines/tobacco/5rs.html

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 777–784
DOI 10.1002/acr.25493

American College of Rheumatology
Empowering Rheumatology Professionals

Frequency of Spondyloarthritis Symptoms Among US Patients With Inflammatory Bowel Disease: A Cross-Sectional Multi-Center Study

Joerg Ermann,[1] Simon J. Hong,[2] Andrew Stahly,[3] Rahul S. Dalal,[1] Fardina Malik,[2] Frank I. Scott,[3] Benjamin Click,[3] Anna Way,[3] John M. Davis III ,[4] Manar Shmais,[4] Laura E. Raffals,[4] Reem Jan,[5] Alexa Silfen,[5] David T. Rubin,[5] Ishaan Dharia,[6] Abhik Bhattacharya,[6] Bahar Moghaddam,[6] Trayton Mains,[7] Jayde E. Kurland,[7] Sheena Patel,[7] Kristine A. Kuhn,[3] and Michael H. Weisman[8]

Objective. Spondyloarthritis (SpA) is the most common extraintestinal manifestation of inflammatory bowel disease (IBD). The application of screening tools to detect SpA in patients with IBD may lead to earlier recognition of SpA and affect treatment decisions.

Methods. A combination of two previously described SpA screening questionnaires, the Detection of Arthritis in Inflammatory Bowel Disease (DETAIL) and IBD Identification of Spondyloarthritis Questionnaire (IBIS-Q), was administered to consecutive patients with IBD attending IBD specialty clinics in six US academic medical centers. Demographic data, IBD, and rheumatology history were extracted by chart review.

Results. A total of 669 patients were analyzed. The median age was 40 years (interquartile range [IQR] 30–54) with a median disease duration of 12 years (IQR 6–22) and moderate to severe IBD based on medication exposure and history of bowel surgery. A total of 81 patients (12%) carried a diagnosis of an inflammatory rheumatic disease, whereas 75 (11%) had consulted a rheumatologist during the previous year. Using published cutoffs, 180 out of 669 patients (27%) screened positive with DETAIL, 266 (40%) with IBIS-Q, and 275 (41%) with either questionnaire. Axial symptoms were more frequently reported than peripheral musculoskeletal complaints. Notably, 189 out of 275 (69%) screen-positive patients had neither a documented inflammatory rheumatic disease diagnosis nor a visit with a rheumatologist within the past year.

Conclusion. A substantial proportion of patients with IBD have symptoms suggestive of SpA, and many of these may have undiagnosed SpA. The IBIS-Q questionnaire appears to identify more potential SpA cases than DETAIL. Strategies are needed to prioritize rheumatology consultations for those patients with IBD who are most likely to benefit.

INTRODUCTION

Arthritis has been known to be associated with inflammatory bowel disease (IBD) for more than a century, but it has only been formally linked to spondylarthritis (SpA) in recent decades.[1] SpA is observed in 10% to 39% of patients with IBD, and additional patients may demonstrate clinical manifestations without meeting formal SpA classification criteria.[2] Conversely, 6% to 14% of patients with ankylosing spondylitis (AS) have clinically recognized IBD, exceeding the prevalence in the general population. Notably, 60% of patients with SpA have subclinical gut inflammation, with a minority eventually developing clinically overt IBD.[2] Despite these observations, the connection between SpA and IBD has remained elusive.[3]

Supported by the Spondylitis Association of America.

[1]Joerg Ermann, MD, Rahul S. Dalal, MD, MPH: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [2]Simon J. Hong, MD, Fardina Malik, MD: New York University Langone Health, New York; [3]Andrew Stahly, MPH, Frank I. Scott, MD, MSCE, Benjamin Click, MD, Anna Way, MD, Kristine A. Kuhn, MD, PhD: University of Colorado Anschutz Medical Campus, Aurora, Colorado; [4]John M. Davis III, MD, Manar Shmais, MD, Laura E. Raffals, MD: Mayo Clinic, Rochester, Minnesota; [5]Reem Jan, MBBS, BSc, Alexa Silfen, BA, David T. Rubin, MD: University of Chicago Medicine, Chicago, Illinois; [6]Ishaan Dharia, MD, Abhik Bhattacharya, MD, Bahar Moghaddam, MD: Mt. Sinai Icahn School of Medicine, New York City, New York; [7]Trayton Mains, DO, Jayde E. Kurland, MD, Sheena Patel, MD: Mercy Health Ohio, Lima, Ohio; [8]Michael H. Weisman, MD: Stanford University School of Medicine, Palo Alto, California.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25493).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25493.

Address correspondence via email to Joerg Ermann, MD, at jermann@bwh.harvard.edu.

Submitted for publication November 21, 2024; accepted in revised form December 23, 2024.

SIGNIFICANCE & INNOVATIONS

- Two spondyloarthritis (SpA) screening tools, the Detection of Arthritis in Inflammatory Bowel Disease (DETAIL) and the IBD Identification of Spondyloarthritis Questionnaire (IBIS-Q), were administered to consecutive patients with inflammatory bowel disease (IBD) at US academic IBD centers.
- A significant proportion of patients with IBD had symptoms strongly suggestive of SpA.
- We demonstrated a feasible approach for identifying potentially undiagnosed cases of SpA.
- Strategies for the accurate diagnosis and effective management of IBD-associated SpA, including implementation of screening practices, remain an urgent unmet need.

Patients with coexisting SpA and IBD have worse physical function and quality of life.[4] Although SpA is the most frequent extraintestinal manifestation (EIM) of IBD, a substantial diagnostic delay is common.[5–7] Obstacles in identifying SpA in IBD include the heterogeneity of SpA phenotypes, the lack of diagnostic tools, and the often-confounding presence of noninflammatory musculoskeletal diseases such as fibromyalgia and osteoarthritis.[8] Finally, the large variability in study designs and the absence of a multidisciplinary approach have been shown to limit investigations into best practices for identifying SpA in patients with IBD.[9]

This study attempts to partially fill this critical knowledge gap and contribute to the goal of improving patient outcomes by further clarifying the relationship between IBD and SpA. The research presented in this article is the result of a collaborative effort between rheumatologists and gastroenterologists. We employed two previously validated questionnaires, the Detection of Arthritis in Inflammatory Bowel Disease (DETAIL) and the IBD Identification of Spondyloarthritis Questionnaire (IBIS-Q),[10–12] to screen patients at IBD clinics in six academic medical centers across the United States. Using these validated screening tools, we sought to estimate the prevalence of SpA symptoms and highlight the potential need for rheumatology evaluation in patients with IBD receiving care at US academic centers.

MATERIALS AND METHODS

Study design. The study was conducted between January 2022 and June 2023 at the following 6 US academic medical centers: Brigham and Women's Hospital (BWH) in Boston, Mayo Clinic in Rochester, Mount Sinai Medical Center in New York City, New York University Langone Health (NYU) in New York City, the University of Chicago in Chicago, and the University of Colorado Anschutz Medical Campus in Aurora. In this cross-sectional study, patients with IBD attending a routine visit at the IBD center of the respective institution were asked to answer a set of SpA screening questions. Demographic and clinical data were then extracted through chart review from the electronic health records at each site and entered into a REDCap database.[13] Deidentified data were exported to Microsoft Excel for centralized analysis. Each site independently obtained institutional review board approval (BWH, 2023P001262; Mayo Clinic, 22-004690; Mt. Sinai Medical Center, STUDY-22-00796-CR001; University of Chicago, IRB23-0264; NYU, i22-00654; University of Colorado, 22-0788). The study is reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cross-sectional studies.[14]

Recruitment. Patients with IBD were invited to participate up to two weeks before their scheduled appointment or at check-in. Enrollment details varied slightly because of local circumstances (eg, implied consent by completion of the questionnaire vs signing an informed consent form; completion of the questionnaire using a paper form vs using a tablet computer). Nonetheless, each site ensured that unselected consecutive patients were enrolled and completed the study questionnaire prior to their encounter with the treating physician. IBD diagnoses were confirmed during chart review. Patients were included in the analysis if they had a diagnosis of IBD based on documented clinical symptoms with supporting endoscopic, pathologic, and imaging findings. Our goal was to enroll 600 patients (100 per participating center) based on the sample size used in the validation study by Benfaremo et al.[12]

Questionnaire. The study questionnaire combined the DETAIL[10] and IBIS-Q[11] and included 16 YES/NO questions about peripheral arthritis, inflammatory back pain, dactylitis, enthesitis, and functional limitations (Supplemental Table 1). The DETAIL questionnaire, consisting of 6 YES/NO questions, was previously validated in a cohort of 418 consecutive patients with IBD without a history of SpA.[12] In this cohort, three or more positive screening questions had a sensitivity of 75% and a specificity of 79% for a rheumatologist diagnosis of SpA; the positive and negative predictive values were 53% and 91%, respectively. The IBIS-Q, which contains 14 YES/NO questions, had a sensitivity of 93% and specificity of 77% for SpA, similarly ascertained by a rheumatologist, with three or more positive answers in a cohort of 181 consecutive patients with IBD.[11] Questionnaire results were transcribed into an Excel table; there were no missing data. We used previously determined thresholds (≥3 YES responses with either DETAIL or IBIS-Q) to identify screen-positive patients. We also analyzed the DETAIL and IBIS-Q score distribution (sum of YES responses) in the cohort as well as the frequency of YES responses for individual questions.

Clinical data. Demographic information, IBD type, treatment history, and history of EIMs were obtained through chart review. We specifically recorded prior and current medication

exposure and history of IBD-related surgeries or hospitalization. We identified any prior diagnoses of inflammatory rheumatic disease by reviewing problem lists and clinic visit notes and determined whether the patient had consulted a rheumatologist within one year prior to enrollment in the study. Patients with a documented clinical diagnosis of an inflammatory rheumatic disease were compared with the total cohort and to those without a prior diagnosis of an inflammatory rheumatic disease.

Statistics and graphing. Statistical analysis and graphing were done using Excel, R version 4.4.0, and GraphPad Prism version 10.3. Continuous variables are presented as means with SDs or medians with interquartile ranges, whereas categorical variables are shown as percentages. The frequency of positive DETAIL and IBIS-Q tests at individual sites was compared using the paired *t*-test. The frequency of positive DETAIL or IBIS-Q tests in subsets of patients was compared using the chi-square test. The unpaired *t*-test was used to compare DETAIL and IBIS-Q scores in patients with and without a prior inflammatory rheumatic disease diagnosis. A receiver operating characteristic (ROC) curve analysis was performed to identify individual questionnaire items and combinations that could discriminate among groups.

RESULTS

Study population. A total of 669 patients with IBD were analyzed, with approximately equal contribution of patients from each of the six participating centers. Demographic details of the cohort are presented in Table 1. The median age of the patients was 40 years (interquartile range [IQR] 30–54) with a median IBD disease duration of 12 years (IQR 6–22). A total of 381 (57%) of the patients were female, and the majority was White (84%).

Table 1. Patient demographics and clinical setting*

	Overall cohort, n = 669
Patient demographics	
Age (years)	40 (30–54)
Disease duration (years)	12 (6–22)
Female	381 (57)
Race	
White	561 (84)
Black	46 (7)
Asian/Pacific Islander	23 (3)
Ethnicity	
Hispanic	40 (6)
Study sites	
US region	
Northeast	332 (50)
Midwest	236 (35)
West	101 (15)
Institution	
Brigham and Women's Hospital	116 (17)
Mayo Clinic	119 (18)
Mt. Sinai Medical Center	96 (14)
New York University	120 (18)
University of Chicago	117 (18)
University of Colorado	101 (15)

* Numbers are n (%) except for age and disease duration where numbers represent median (IQR) in years. IQR, interquartile range.

IBD characteristics. Table 2 presents baseline IBD characteristics of the study population. A total of 398 (59%) had Crohn disease (CD), 247 (37%) had ulcerative colitis (UC), and 24 (4%) had unclassified IBD. At enrollment, 520 (78%) were receiving treatment with a biologic or targeted synthetic disease-modifying antirheumatic drug. Tumor necrosis factor inhibitors were the most frequently used biologic (239 out of 669, 36%) followed by the interleukin-12p40 inhibitor ustekinumab (129 out of 669, 19%) and the α4β7 integrin antagonist vedolizumab (88 out of 669, 13%), reflecting clinical practice in 2022 to 2023.

Table 2. Baseline IBD characteristics, medication history, and EIMs*

	Overall cohort, n = 669	
IBD phenotype and behavior		
Crohn disease	398 (59)	
Ulcerative colitis	247 (37)	
IBD unclassified	24 (4)	
Treatment history		
Pharmacotherapy	**Any exposure**	**Current use**
5-ASA	437 (65)	115 (17)
Systemic corticosteroids	538 (80)	69 (10)
Thiopurine antimetabolites (6-MP, azathioprine)	293 (44)	67 (10)
Methotrexate	92 (14)	23 (3)
TNF antagonist	470 (70)	239 (36)
Vedolizumab	201 (30)	88 (13)
Anti-IL-12p40	189 (28)	129 (19)
Anti-IL-23p19	23 (3)	21 (3)
JAK inhibitor	64 (10)	43 (7)
h/o GI surgeries	261 (39)	
h/o bowel resection	220 (33)	
h/o IBD-related hospitalizations	298 (45)	
Extraintestinal manifestations		
h/o rheumatologic EIM	149 (22)	
h/o dermatologic EIM	61 (9)	
h/o ophthalmologic EIM	12 (2)	
h/o hepatobiliary EIM	30 (5)	
Rheumatology history		
Prior inflammatory rheumatic disease diagnosis[a]	81 (12)	
Peripheral	38 (6)	
Axial SpA/AS	28 (4)	
Psoriatic arthritis	9 (1)	
Rheumatoid arthritis	7 (1)	
Dactylitis	1 (<1)	
Enthesitis	1 (<1)	
Other	9 (1)	
Rheumatology visit in the past year	75 (11)	

* Values are the n (%). 6-MP, 6-mercaptopurine; AS, ankylosing spondylitis; ASA, aminosalicylic acid; EIM, extraintestinal manifestation; GI, gastrointestinal; IBD, inflammatory bowel disease; IL, interleukin; SpA, spondyloarthritis; TNF, tumor necrosis factor.
[a] Some patients had more than one diagnosis.

Additionally, 261 patients (39%) had a history of bowel surgery. These clinical characteristics are consistent with moderate to severe disease expected in this cohort of patients with IBD attending US tertiary care centers.

EIM characteristics. By chart review, 149 patients (22%) had a history of a rheumatologic EIM, making rheumatologic EIMs the most frequently recorded EIMs, followed by dermatologic (9%), hepatobiliary (5%), and ophthalmologic (2%) EIMs. Just over 10% of patients had a documented diagnosis of an inflammatory rheumatic disease in their health record (81 of 669, 12%) or had consulted a rheumatologist during the previous year (75 of 669, 11%). The subset of patients with IBD and an established inflammatory rheumatic disease diagnosis included patients with peripheral SpA (38 of 81), axial SpA or AS (28 of 81), and psoriatic arthritis (9 of 81). Seven patients carried a diagnosis of rheumatoid arthritis.

Global questionnaire results. Using at least three affirmative responses as the cutoff for a positive screen, 180 patients (27%) were positive with DETAIL, 266 (40%) with IBIS-Q, and 275 (41%) with either questionnaire (Figure 1A, Table 3). A total of 171 patients (26%) screened positive with both questionnaires. Almost all patients who screened positive with DETAIL also tested positive with IBIS-Q (171 of 180, 95%), whereas a substantially smaller fraction of the IBIS-Q positive patients also screened positive with DETAIL (171 of 266, 65%) (Figure 1B). There was a notable variability in screen positivity rates among centers, ranging from 16% to 36% for DETAIL and from 24% to 51% for IBIS-Q. Consistently, the screen positivity rate with DETAIL was lower than with IBIS-Q (mean ± SD, 28 ± 7% vs 41 ± 10%, $P = 0.001$) (Figure 1C). No significant differences in screen positivity rates were observed between patients with CD or UC (28% vs 26% with DETAIL, $P = 0.663$; 42% vs 37% with IBIS-Q, $P = 0.261$). Patients with a documented inflammatory rheumatic disease diagnosis were more likely to have positive screens than patients without such a diagnosis (62% vs 22% for DETAIL, $P < 0.001$; 84% vs 34% for IBIS-Q, $P < 0.001$). Similarly, patients who had seen a rheumatologist within the last year were more likely to have a positive screen than patients who had not seen a rheumatologist (63% vs 22% for DETAIL, $P < 0.001$; 87% vs 34% for IBIS-Q, $P < 0.001$). Among the 75 patients who had seen a rheumatologist in the last year, 54 (72%) carried a diagnosis of an inflammatory rheumatic disease, whereas 21 (28%) did not. Screen positivity rates between these two subgroups showed

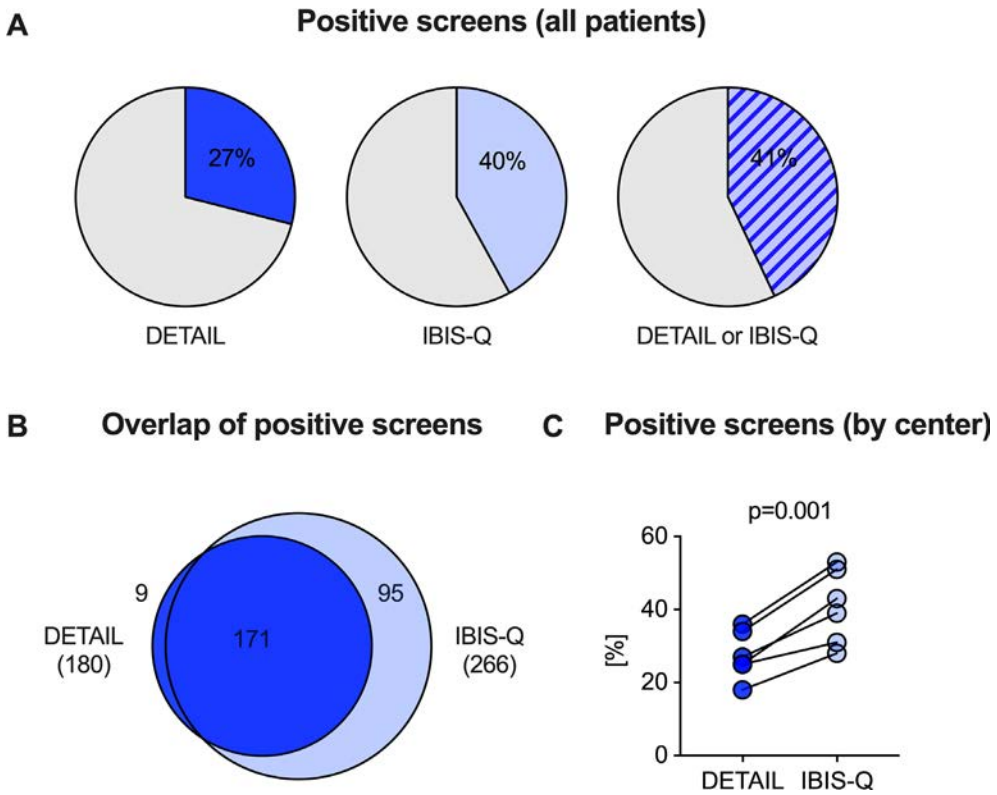

Figure 1. Screen positivity rates. (A) Proportion of consecutive patients with IBD (n = 669) who screened positive with DETAIL, IBIS-Q, or either of the two questionnaires across all centers. (B) Venn diagram demonstrating the overlap of positive screen results for DETAIL and IBIS-Q. (C) Proportion of patients who screened positive with DETAIL or IBIS-Q in the six participating IBD centers. Each pair of values represents one center. *P* value by paired *t*-test. DETAIL, Detection of Arthritis in Inflammatory Bowel Disease; IBD, inflammatory bowel disease; IBIS-Q, IBD Identification of Spondyloarthritis Questionnaire.

Table 3. Questionnaire responses*

	DETAIL	IBIS-Q	DETAIL or IBIS-Q
Overall cohort, n = 669	180 (27)	266 (40)	275 (41)
By institution			
Brigham and Women's Hospital, n = 116	18 (16)	28 (24)	30 (26)
Mayo Clinic, n = 119	43 (36)	63 (53)	66 (56)
Mt. Sinai Medical Center, n = 96	26 (27)	37 (39)	37 (39)
New York University, n = 120	30 (25)	51 (43)	54 (45)
University of Chicago, n = 117	29 (25)	36 (31)	37 (32)
University of Colorado, n = 101	34 (34)	51 (51)	51 (51)
By IBD type			
Crohn disease, n = 398	111 (28)	166 (42)	171 (43)
Ulcerative colitis, n = 247	65 (26)	92 (37)	96 (39)
By rheumatology history			
Prior inflammatory rheumatic disease diagnosis, n = 81	50 (62)	68 (84)	68 (84)
Rheumatology visit in the past year, n = 75	47 (63)	65 (87)	65 (87)

* Number (%) of patients who screened positive (≥3 YES responses) for DETAIL, IBIS-Q, or either questionnaire (DETAIL or IBIS-Q). DETAIL, Detection of Arthritis in Inflammatory Bowel Disease; IBD, inflammatory bowel disease; IBIS-Q, IBD Identification of Spondyloarthritis Questionnaire; NYU, New York University; UC, ulcerative colitis.

no significant differences (63% vs 62% for DETAIL, $P = 0.932$, 87% vs 86% for IBIS-Q, $P = 0.88$).

Score distribution. Figure 2A presents the DETAIL and IBIS-Q score distributions for the total cohort and for patients with or without a prior inflammatory rheumatic disease diagnosis. For both questionnaires, the total population and the group of patients without an inflammatory rheumatic disease diagnosis showed a Poisson pattern, whereas the distribution for the patients with a documented inflammatory rheumatic diagnosis approached a normal distribution, with modes of four positive responses for DETAIL and six positive responses for IBIS-Q. The DETAIL score (mean ± SD) for patients with a diagnosis of an inflammatory rheumatic disease (n = 81) was 3.0 ± 1.6 compared with 1.4 ± 1.5 for patients without such a diagnosis (n = 588) ($P < 0.001$). The corresponding IBIS-Q scores were 5.5 ± 2.9 for patients with and 2.3 ± 2.7 for patients without an inflammatory rheumatic disease diagnosis ($P < 0.001$).

Individual questionnaire items. The 16 questions in the study questionnaire include questions about axial symptoms (back pain and stiffness), peripheral joint symptoms and signs (joint pain and swelling), enthesitis (heel pain), and dactylitis (sausage digit). The most frequently positive screening question was "Do you have low back pain in the morning and/or after resting that improves with exercise?", a question about inflammatory back pain. This question was answered with YES by 252 out of 669 patients (38%). Three of the four questions most frequently answered with YES were questions about axial symptoms (Figure 2B). No individual question outperformed the DETAIL or IBIS-Q to discriminate between patients with and without an established inflammatory rheumatic disease diagnosis in the ROC curve analysis. Using established cutoffs, the IBIS-Q was slightly better in distinguishing patients with and without an inflammatory rheumatic disease diagnosis with an area under the curve (AUC) of 0.8 and a positive likelihood ratio of 2.49. DETAIL similarly distinguished between the two groups with an AUC of 0.76 and a positive likelihood ratio of 2.79.

Burden of undiagnosed SpA. A total of 50 out of 180 patients (28%) with a positive DETAIL and 68 out of 266 (26%) with a positive IBIS-Q had a preexisting inflammatory rheumatic disease diagnosis, whereas 47 out of 180 patients (26%) with a positive DETAIL and 65 out of 266 (24%) with a positive IBIS-Q had seen a rheumatologist within the last year. Conversely, as many as 189 out of 275 screen-positive patients (69%) had neither a documented inflammatory rheumatic disease diagnosis nor had seen a rheumatologist within the past year. Among the 588 patients without a prior inflammatory rheumatic disease diagnosis, 22% screened positive with the DETAIL and 34% with the IBIS-Q. Together, these findings support the conclusion that a substantial proportion of patients with IBD may have undiagnosed SpA.

DISCUSSION

We screened a convenience sample of 669 consecutively encountered patients with IBD from 6 IBD centers across the United States using the previously validated DETAIL and IBIS-Q questionnaires for SpA. Based on established cutoffs, 27% of patients screened positive with the DETAIL and 40% with the IBIS-Q. Among those who screened positive, a notable majority (69%) had no documented inflammatory rheumatic disease diagnosis nor seen a rheumatologist during the previous year. These findings suggest a significant burden of undiagnosed SpA in the IBD population and underscore the need for improved rheumatology care among US patients with IBD.

The frequency of screen-positive individuals in our study was comparable with the original studies that developed and validated the DETAIL questionnaire; the IBIS-Q publication did not report

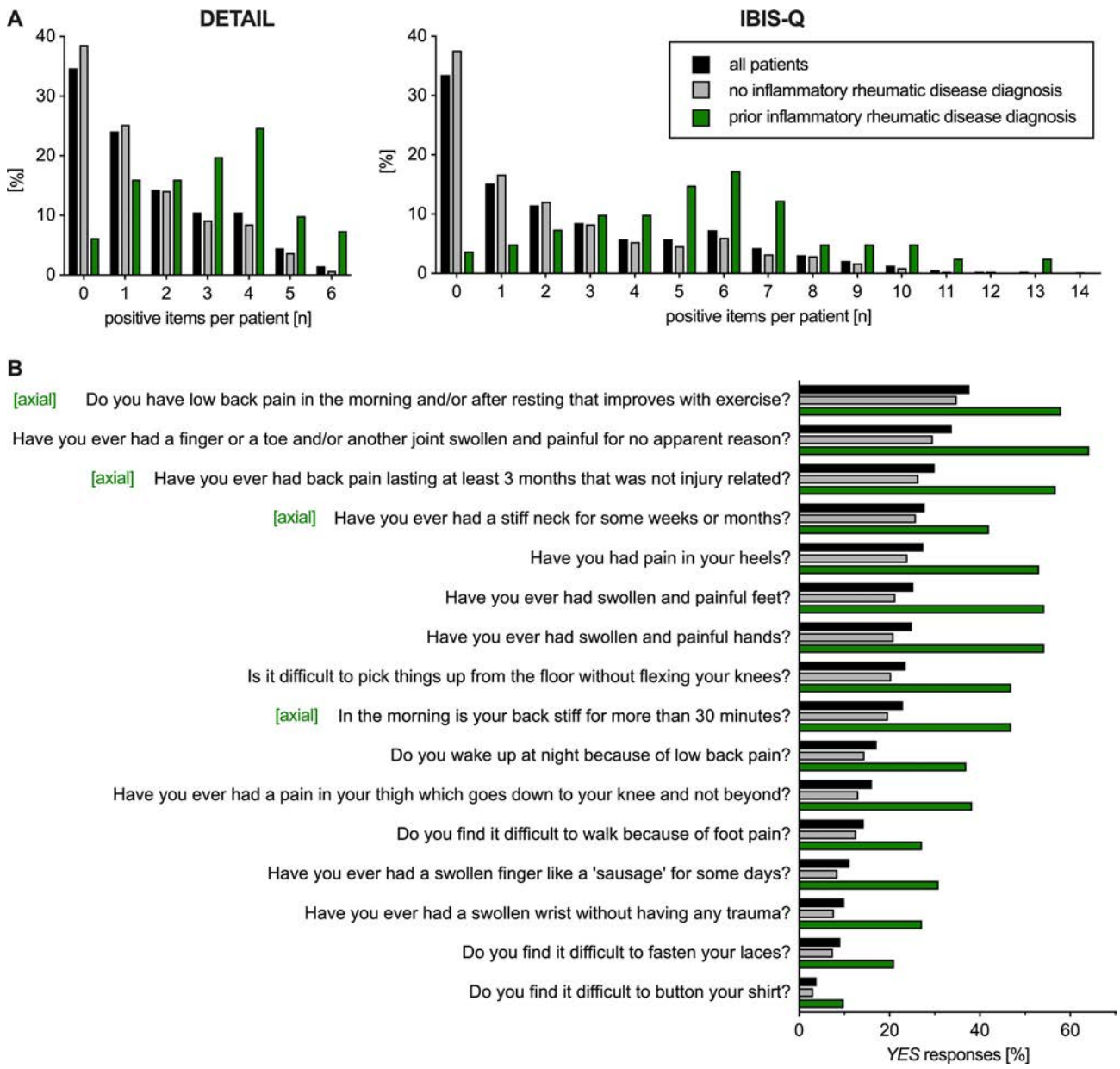

Figure 2. Score distribution and responses to individual screening questions. (A) Proportion of patients with IBD who answered YES to 0–6 DETAIL questions or 0–14 IBIS-Q questions, shown separately for all patients as well as patients without or with a prior inflammatory rheumatic disease diagnosis. (B) Frequency of YES responses for each of the 16 questions, ranked by the frequency of YES responses in all patients. The four questions about axial symptoms are highlighted. DETAIL, Detection of Arthritis in Inflammatory Bowel Disease; IBD, inflammatory bowel disease; IBIS-Q, IBD Identification of Spondyloarthritis Questionnaire. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25493/abstract.

the raw questionnaire results, limiting our ability to compare findings. Consistently across all centers, we found that the number of screen-positive patients was higher with IBIS-Q than with DETAIL, suggesting that IBIS-Q may have higher sensitivity, potentially at the cost of lower specificity.[15] We did not perform a rheumatology assessment of every screened patient, which precludes our ability to calculate the test characteristics of sensitivity, specificity, and predictive values. The significantly higher screen positivity rates in patients with a documented diagnosis of an inflammatory rheumatic disease, most of whom had SpA, supports the face validity of the screening tools. However, among the patients who had consulted a rheumatologist in the last year, screen positivity rates were similar regardless of whether they had an inflammatory rheumatic disease diagnosis, highlighting the limited specificity of the questionnaires.

Both DETAIL and IBIS-Q combine questions about current and historical symptoms. Current musculoskeletal symptoms may be influenced by IBD-directed therapies that also treat musculoskeletal inflammation, whereas responses to questions about past symptoms are subject to recall bias. The extent to which

these nuances affect the performance of the screening tools remains uncertain and requires further study. The inclusion of past symptoms is likely to increase the sensitivity of a screening tool but detracts from the focus on current problems requiring intervention. Assigning a higher weight to current symptoms may help to more accurately identify those patients with IBD who may benefit the most from a rheumatology referral. Alternatively, adding a question such as "Have you recently considered seeking medical care for any of the symptoms mentioned in this questionnaire?" may be useful.

Our data indicate that the majority of screen-positive patients did not have an existing diagnosis of an inflammatory rheumatic disease and had not seen a rheumatologist within the last year, highlighting a potential unmet need. It is possible that our data overestimate the proportion of symptomatic patients with IBD who have not previously been investigated for SpA, as we only had access to electronic medical record data for the enrolling academic medical centers. Rheumatology care provided outside these systems may have been missed. Additionally, we limited our analysis to patients who had a rheumatology visit within one year prior to enrollment, assuming that those diagnosed with SpA in an academic setting would be followed longitudinally with at least annual visits. Consequently, patients with less frequent visits or those lost to follow-up may have been overlooked. However, it is unlikely that this would occur without any indication of a rheumatic disease diagnosis in problem lists or clinic notes.

The true prevalence of SpA in patients with IBD remains uncertain, with estimates in the published literature ranging from 1% to 46% for axial SpA and from 1% to 43% for peripheral SpA.[9] Part of the challenge lies in the absence of classification criteria for IBD-SpA.[16] Efforts to develop such criteria are underway, and these tools will be essential for future clinical trials.[17] In our study, the frequency of positive screens varied from 16% to 36% for the DETAIL and from 24% to 51% for the IBIS-Q, suggesting considerable heterogeneity even among cohorts in academic centers. Contributing factors may include differences in patient demographics, local referral patterns, and other historically evolved practices. We cannot rule out study design as an additional factor; although we mandated that consecutive patients were approached before a routine visit to the IBD center, other details of enrollment were left to the individual center as a practical matter. How the study was presented to patients at each site may have influenced their individual willingness to participate (selection bias) or their response to questions (response bias). It will be important to obtain similar data from patients with IBD in different practice settings such a community practices or integrated health care systems. However, a recent analysis of population-based data from the US National Health and Nutrition Examination Survey (NHANES) 1976 to 1980 and 2009 to 2010 cycles revealed high rates of chronic axial pain, inflammatory back pain, and peripheral arthritis in persons with physician-diagnosed IBD compared with those without an IBD diagnosis, which is consistent with the substantial burden of SpA symptoms in the IBD population revealed by the current study.[18]

For both DETAIL and IBIS-Q, the proportion of patients who screened positive was substantial, raising concerns about the feasibility of providing a rheumatology assessment for all screen-positive patients in clinical practice.[19] Workforce limitations may necessitate prioritizing patients for rheumatology referral based on the likelihood of significantly improving their symptom burden and long-term outcomes. Alternatively, if it were possible to reliably predict which musculoskeletal manifestations would respond to specific IBD therapies, gastroenterologists might manage these symptoms effectively as part of a broader treatment plan. More likely, the increasing diversity and complexity of available treatments for SpA and IBD will require and benefit from the concurrent management of patients with IBD-SpA by gastroenterologists and rheumatologists in an interdisciplinary clinic, similar to the collaboration of dermatologists and rheumatologists when managing patients with psoriatic arthritis.[20, 21]

This study is the first to report SpA screening positivity rates in IBD centers across different US geographic regions, providing valuable insight into the potential burden of SpA in the US IBD population. An additional strength of the study is that we compared DETAIL and IBIS-Q in the same patient population. Limitations include the cross-sectional design and the lack of a rheumatology assessment to confirm SpA diagnoses. Despite these limitations, our study highlights an important clinical problem. Future research should focus on improving access to rheumatology care for patients with IBD with musculoskeletal symptoms and identifying those who will benefit most from such interventions.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Ermann confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Ashrafi M, Kuhn KA, Weisman MH. The arthritis connection to inflammatory bowel disease (IBD): why has it taken so long to understand it? RMD Open 2021;7:e001558.
2. Fragoulis GE, Liava C, Daoussis D, et al. Inflammatory bowel diseases and spondyloarthropathies: from pathogenesis to treatment. World J Gastroenterol 2019;25:2162–2176.
3. Gracey E, Vereecke L, McGovern D, et al. Revisiting the gut-joint axis: links between gut inflammation and spondyloarthritis. Nat Rev Rheumatol 2020;16:415–433.

4. Luchetti MM, Benfaremo D, Bendia E, et al. Clinical and patient reported outcomes of the multidisciplinary management in patients with inflammatory bowel disease-associated spondyloarthritis. Eur J Intern Med 2019;64:76–84.
5. Rogler G, Singh A, Kavanaugh A, et al. Extraintestinal manifestations of inflammatory bowel disease: current concepts, treatment, and implications for disease management. Gastroenterology 2021;161: 1118–1132.
6. Conigliaro P, D'Antonio A, Wlderk A, et al. Clinical and imaging findings in enteropathic spondyloarthritis with special emphasize in diagnostic delay: a cross-sectional study. Ther Adv Chronic Dis 2024;15: 20406223241229843.
7. Rios Rodriguez V, Duran Tl, Torgutalp M, et al. Comparing clinical profiles in spondyloarthritis with Crohn's disease or ulcerative colitis: insights from the ASAS-PerSpA study. Rheumatol Adv Pract 2024; 8(2):rkae064.
8. Conigliaro P, Chimenti MS, Ascolani M, et al. Impact of a multidisciplinary approach in enteropathic spondyloarthritis patients. Autoimmun Rev 2016;15:184–190.
9. Schwartzman M, Ermann J, Kuhn KA, et al. Spondyloarthritis in inflammatory bowel disease cohorts: systematic literature review and critical appraisal of study designs. RMD Open 2022;8:e001777.
10. Di Carlo M, Luchetti MM, Benfaremo D, et al. The DETection of Arthritis in Inflammatory boweL diseases (DETAIL) questionnaire: development and preliminary testing of a new tool to screen patients with inflammatory bowel disease for the presence of spondyloarthritis. Clin Rheumatol 2018;37:1037–1044.
11. Variola A, Zanolin ME, Cipriano G, et al. The IBIS-Q [IBd Identification of Spondyloarthritis Questionnaire]: a novel tool to detect both axial and peripheral arthritis in inflammatory bowel disease patients. J Crohns Colitis 2020;14:1680–1686.
12. Benfaremo D, Luchetti MM, Di Carlo M, et al; GRADES-IBD Study Group. Multicenter validation of the DETAIL Questionnaire for the screening of spondyloarthritis in patients with inflammatory bowel diseases. J Rheumatol 2021;48:179–187.
13. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–381.
14. von Elm E, Altman DG, Egger M, et al; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370:1453–1457.
15. Grimes DA, Schulz KF. Uses and abuses of screening tests. Lancet 2002;359:881–884.
16. Song K, Satsangi J, Coates LC. Arthritis complicating inflammatory bowel disease- the future is now. Lancet Rheumatol 2024;6:e805–e810.
17. Falloon K, Cohen B, Ananthakrishnan AN, et al. A United States expert consensus to standardise definitions, follow-up, and treatment targets for extra-intestinal manifestations in inflammatory bowel disease. Aliment Pharmacol Ther 2022;55:1179–1191.
18. Weisman MH, Stens O, Kim HS, et al. Axial pain and arthritis in diagnosed inflammatory bowel disease: US National Health and Nutrition Examination Survey Data. Mayo Clin Proc Innov Qual Outcomes 2022;6:443–449.
19. Battafarano DF, Ditmyer M, Bolster MB, et al. 2015 American College of Rheumatology Workforce Study: supply and demand projections of adult rheumatology workforce, 2015-2030. Arthritis Care Res (Hoboken) 2018;70:617–626.
20. Haberman R, Perez-Chada LM, Merola JF, et al. Bridging the gaps in the care of psoriasis and psoriatic arthritis: the role of combined clinics. Curr Rheumatol Rep 2018;20:76.
21. Sayers S, Lam D, Shah Q, et al. Impact on patient outcomes of spondyloarthritis-inflammatory bowel disease multi-disciplinary meetings. Rheumatology (Oxford) 2024;keae116.

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 785–791
DOI 10.1002/acr.25491

American College of Rheumatology
Empowering Rheumatology Professionals

Do the Provisional Paediatric Rheumatology International Trials Organisation Enthesitis/Spondylitis-Related Juvenile Idiopathic Arthritis Criteria Capture Youth With Axial Spondyloarthritis?

Pamela F. Weiss,[1] Timothy G. Brandon,[2] Amita Aggarwal,[3] Ruben Burgos-Vargas,[4] Robert A. Colbert,[5] Gerd Horneff,[6] Ronald M. Laxer,[7] Kirsten Minden,[8] Angelo Ravelli,[9] Nicolino Ruperto,[10] Judith A. Smith,[11] Matthew L. Stoll,[12] Shirley M. Tse,[13] Filip Van den Bosch,[14] Walter P. Maksymowych,[15] Robert G. Lambert,[16] David M. Biko,[1] Nancy A. Chauvin,[17] Michael L. Francavilla,[18] Jacob L. Jaremko,[16] Nele Herregods,[19] Ozgur Kasapcopur,[20] Mehmet Yildiz,[20] Hemalatha Srinivasalu,[21] Alison M. Hendry,[22] and Rik Joos[23]

Objective. The Paediatric Rheumatology International Trials Organisation (PRINTO) recently undertook an effort to better harmonize the pediatric and adult arthritis criteria. These provisional criteria are being refined for optimal performance. We aimed to investigate differences between patients who did and did not fulfill these PRINTO criteria among youth diagnosed with juvenile spondyloarthritis (SpA) that met axial juvenile SpA (axJSpA) classification criteria.

Methods. This was a retrospective cross-sectional sample of youth diagnosed with juvenile SpA who met the axJSpA classification criteria. Demographics, clinical manifestations, and physician and patient-reported outcomes were abstracted from medical records. Magnetic resonance imaging (MRI) scans underwent central imaging review by at least two central raters. Differences between groups were compared using Wilcoxon signed-rank test or chi-square test, as appropriate.

Results. Of 158 patients who met axJSpA criteria, 107 patients (68%) met the PRINTO provisional criteria for enthesitis/spondylitis-related arthritis. A total of 41 patients (26%) did not fulfill any of the three major PRINTO criteria due to lack of peripheral disease manifestations. Demographics, prevalence of inflammatory or structural lesions on MRI, family history of SpA, and duration of pain were not statistically different between those who did and did not meet PRINTO criteria. Those who fulfilled the PRINTO criteria had significantly more peripheral arthritis, enthesitis, and HLA-B27 positivity but reported less sacral/buttock pain.

Conclusion. Phenotypic differences of children with axJSpA between those who were and were not classified by the PRINTO criteria were primarily due to peripheral disease manifestations and HLA-B27 positivity. Modification of the PRINTO provisional criteria may facilitate capture of youth with primarily axial disease.

INTRODUCTION

Most children with spondyloarthritis (SpA) are classified as having enthesitis-related arthritis by the International League of Associations for Rheumatology (ILAR) criteria[1] and will be classified as having enthesitis/spondylitis-related arthritis by the provisional Paediatric Rheumatology International Trials Organisation (PRINTO) juvenile idiopathic arthritis (JIA) criteria.[2]

Supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), NIH (grant 1R01-AR-074098 to Dr Weiss); Spondylitis Association of America (to Dr Weiss); and National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grant ZIA-AR-041184 to Dr Colbert).

[1]Pamela F. Weiss, MD, MSCE, David M. Biko, MD: Children's Hospital of Philadelphia and Perelman School of Medicine, University of Pennsylvania; [2]Timothy G. Brandon, MPH: Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; [3]Amita Aggarwal, MD, DM: Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow, India; [4]Ruben Burgos-Vargas, MD: Hospital General de Mexico Dr Eduardo Ligeaga, Mexico City, Mexico; [5]Robert A. Colbert, MD, PhD: National Institute of Arthritis, Musculoskeletal, and Skin Diseases, NIH, Bethesda, Maryland; [6]Gerd Horneff, MD, Asklepios Klinik Sankt Augustin, Sankt Augustin, Germany, and University Hospital Cologne, Cologne, Germany; [7]Ronald M. Laxer, MDCM, FRCPC: University of Toronto, The Hospital for Sick Children, and St. Michael's Hospital, Toronto, Ontario, Canada; [8]Kirsten Minden, MD: Charité – Universitätsmedizin Berlin and Deutsches Rheuma-Forschungszentrum, Berlin, Germany; [9]Angelo Ravelli, MD: IRCCS Istituto Giannina Gaslini and Università degli Studi di Genova, Genoa, Italy; [10]Nicolino Ruperto, MD, MPH: Università Milano Bicocca, Milano, Italy, and IRCCS Fondazione San Gerardo dei Tintori, Monza, Italy; [11]Judith A. Smith, MD, PhD, University of Wisconsin–Madison School of Medicine and Public Health; [12]Matthew L. Stoll, MD, PhD, MSCS: University of Alabama at Birmingham; [13]Shirley M. Tse, MD, FRCPC: The Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada; [14]Filip Van den Bosch,

SIGNIFICANCE & INNOVATIONS

- Recently, classification criteria for youth with axial juvenile spondyloarthritis (axJSpA) were developed and validated.
- Many youths who fulfill the axJSpA criteria were not classifiable by the International League of Associations for Rheumatology enthesitis–related arthritis (28%) criteria or the Paediatric Rheumatology International Trials Organisation (PRINTO) provisional juvenile idiopathic arthritis classification criteria for enthesitis/spondylitis-related arthritis (32%).
- We propose minor modifications to the provisional PRINTO criteria to facilitate capture of youth with juvenile spondyloarthritis manifested primarily by axial disease.

Recently, classification criteria for youth diagnosed clinically with axial juvenile SpA (axJSpA) were developed and validated.[3,4] The axJSpA criteria are comprised of five clinical and two imaging domains. Each domain consists of two to four levels, and levels within each domain are mutually exclusive. Patients with a score of 55 or higher are classified as having axJSpA. In the validation cohort, the axJSpA criteria had a specificity of 97.5% (95% confidence interval [CI] 91.4–99.7), sensitivity of 64.3% (95% CI 54.9–73.1), and area under the receiver operating characteristic curve of 0.81 (95% CI 0.76–0.86) using clinical SpA expert consensus as the reference standard. The development of axial criteria is in accordance with adult axSpA criteria, which define distinct populations of adults with peripheral SpA and nonradiographic axSpA and radiographic SpA.[5]

The shortfalls of the ILAR JIA criteria[1] have been debated,[6] and PRINTO undertook an effort to better harmonize the pediatric and adult criteria for use in both routine clinical care and research. In 2019, the PRINTO provisional criteria for JIA were published with four categories of disease: systemic JIA, rheumatoid factor positive JIA, enthesitis/spondylitis-related JIA, and early-onset anti-nuclear antibody–positive JIA.[2] However, these new provisional criteria are not yet validated, and some refinement may be necessary for optimal performance. Recent work from the Canadian Research in Arthritis in Canadian Children, Emphasizing Outcomes cohort demonstrated that in comparison to the ILAR JIA criteria, a larger proportion of patients with JIA are unclassifiable by the PRINTO criteria. Of 1,228 Canadian children with juvenile arthritis, 12% and 63.3% were unclassifiable using the ILAR or PRINTO JIA criteria, respectively.[7]

In order to meet PRINTO criteria for spondylitis/enthesitis-related arthritis, a child must meet at least one of the following major criteria: (1) peripheral arthritis and enthesitis, (2) arthritis or enthesitis, plus at least three months of inflammatory back pain and sacroiliitis on imaging, or (3) arthritis or enthesitis, plus two of the following: (a) sacroiliac joint (SIJ) tenderness, (b) inflammatory back pain, (c) presence of HLA-B27 antigen, (d) acute (symptomatic) anterior uveitis, or (e) history of SpA in a first-degree relative. Peripheral arthritis, if present, should persist for at least six weeks. For major criterion 2, "arthritis" refers only to peripheral arthritis (not sacroiliitis), and inflammatory back pain was defined according to the Assessment of Spondyoarthritis International Society (ASAS) definition,[8] which is frequently used in adults with axial SpA but has not been validated in youth. Inflammatory back pain according to ASAS is defined by fulfilling at least four of the following criteria: (1) improvement with exercise, (2) pain at night, (3) insidious onset, (4) age at onset <40 years, or (5) no improvement with rest. In adults, these criteria have a sensitivity and specificity of 77% and 91.7%, respectively.[8] Imaging criteria for sacroiliitis can be on radiography or magnetic resonance imaging (MRI). The objectives of this project were to (1) compare the clinical and imaging characteristics of youth who fulfill the recently published axJSpA criteria between those who do and do not also fulfill PRINTO provisional criteria for enthesitis/spondylitis-related JIA, and (2) explore what modifications to the PRINTO criteria would improve capture of youth with juvenile SpA (JSpA) manifested primarily by axial disease.

PATIENTS AND METHODS

Study design. This study was reviewed by the Children's Hospital of Philadelphia (CHOP) Institutional Review Board (IRB), and the IRB determined the procedures met the exemption criteria per 45 Code of Federal Regulations 46.104(d) 4(iii) (IRB 19-016078). This was a retrospective cross-sectional study. This study was a subanalysis of data used in the validation of the classification criteria for patients with axJSpA.[3] The source of patients was an international cross-sectional sample of youth with clinically diagnosed JSpA between 2006 and 2021. Patients originated from seven centers in North America (Bethesda, Maryland;

MD, PhD: VIB-UGent Center for Inflammation Research and Ghent University Hospital, Ghent, Belgium; [15]Walter P. Maksymowych, MD, FRCP: University of Alberta and CARE Arthritis, Edmonton, Alberta, Canada; [16]Robert G. Lambert, MB BCh, FRCR, FRCPC, Jacob L. Jaremko, MD, PhD, FRCPC: University of Alberta, Edmonton, Alberta, Canada; [17]Nancy A. Chauvin, MD, Penn State Health Milton S. Hershey Children's Hospital; [18]Michael L. Francavilla, MD: Whiddon College of Medicine, University of South Alabama, Mobile; [19]Nele Herregods, MD, PhD: Ghent University Hospital, Ghent, Belgium; [20]Ozgur Kasapcopur, MD, Mehmet Yildiz, MD: Cerrahpasa Medical School, Istanbul University–Cerrahpasa, Istanbul, Turkey; [21]Hemalatha Srinivasalu, MD: Children's National Hospital, George Washington University School of Medicine, Washington, District of Columbia, and National Institute of Arthritis, Musculoskeletal and Skin Diseases, NIH, Bethesda, Maryland; [22]Alison M. Hendry, PGDipHealMgt: Counties Manukau District Health Board, Auckland, New Zealand; [23]Rik Joos, MD: Gent University Hospital, Gent, Belgium, and Ziekenhuisnetwerk Antwerpen, Antwerp, Belgium.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25491.

Address correspondence via email to Pamela F. Weiss, MD, MSCE, at weisspa@chop.edu.

Submitted for publication July 22, 2024; accepted in revised form December 6, 2024.

Birmingham, Alabama; Cincinnati, Ohio; Columbus, Ohio; Madison, Wisconsin; Philadelphia, Pennsylvania; and Washington, District of Columbia), two centers in South America (Rio de Janeiro, Brazil; Sao Paulo, Brazil), and five centers from Europe/Asia (Gent, Belgium; Garmisch-Partenkirchen, Germany; Sankt Augustin, Germany; Delhi, India; and Istanbul, Turkey). In addition to a clinical diagnosis of JSpA, all patients had axial symptom onset before the age of 18 years, had MRI as part of a diagnostic evaluation for axial disease, and fulfilled the axJSpA criteria.[4]

Clinical features. Clinical data (demographics, disease characteristics, physical examination findings, and patient- and physician-reported metrics) were abstracted from the patients' medical records and collected on a standardized electronic case report form using REDCap electronic data capture tools hosted at the CHOP.[9] All PRINTO provisional criteria components were collected as present/absent/unknown, and the remaining patient characteristics of interest were collected with as much granularity as was anticipated to be feasible through retrospective chart abstraction. As such, information related to a patient's history of inflammatory back pain was collected as present/absent/unknown at the interpretation of the submitting physician, and the individual inflammatory back pain criteria components from the ASAS definition were also collected.[8] To be included in the study, all patients had to have MRI of the pelvis performed, and the scans underwent central imaging review. Digital Imaging and Communications in Medicine MRI files were transferred using a secure file-sharing platform. Results of the imaging have been previously reported as part of the axJSpA classification criteria development.[10,11] The PRINTO criteria define sacroiliitis on MRI as follows: "1. Bone marrow edema (BMO) on a T2-weighted sequence (required criteria) sensitive for free water (such as short tau inversion recovery (STIR) or T2FS) or bone marrow contrast enhancement on a T1-weighted sequence (such as T1FS post Gadolinium). 2. Inflammation must be clearly present and located in a typical anatomical area (subchondral bone). 3. MRI appearance must be highly suggestive of SpA."[2] Each imaging team member's assessment of the MRI was evaluated for the presence of bone marrow edema, high-confidence inflammatory lesions (≥3 of 5) are compatible with active lesions seen in patients with axial SpA, and high confidence (≥3 of 5) that the findings on the MRI of the SIJs are indicative of SpA. Imaging was rated independently and blind to clinical details by at least two central imaging team members, and a third rater adjudicated patients for whom there was disagreement on the global assessment of the presence/absence of lesions typical of axial SpA. Unequivocal evidence of inflammatory lesions on MRI typical of axJSpA was defined as bone marrow edema in at least three SIJ quadrants across all SIJ MRI slices. Unequivocal evidence of structural lesion(s) on MRI typical of axJSpA was defined as erosion in at least three quadrants or sclerosis or fat lesion in at least two SIJ quadrants or backfill or ankylosis in at least two joint halves across all SIJ MRI slices.[10] In the absence of pelvic MRI, unequivocal evidence of structural lesions on radiograph typical of axJSpA was defined as follows: "unequivocal lesion (erosion, sclerosis, or ankylosis [partial or complete]) that must include at least one iliac bone; When sclerosis is present in isolation, if measurable, it should extend ≥5mm from the joint surface; The decision may be influenced by the presence of other lesions, which in themselves do not suffice to meet the criterion."[11]

Outcome and analysis. The primary outcome was fulfillment of the PRINTO provisional spondylitis/enthesitis-related JIA criteria. Patient demographics, clinical manifestations, and physician- and patient-reported outcomes were evaluated using standard descriptive statistics. Differences between groups were compared using Wilcoxon rank-sum test or chi-square test as appropriate.

RESULTS

Among 521 children and adolescents diagnosed with JSpA by a clinician, 158 fulfilled the axJSpA classification criteria, but only 113 of them (72%) met ILAR criteria for enthesitis-related arthritis, and 107 (68%) met PRINTO provisional criteria for enthesitis/spondylitis-related JIA. For patients who met axJSpA criteria, 109 patients (69%) were male, the median age was 15 years (interquartile range 12.6–16.8 years), and 63% were HLA-B27 positive. Table 1 compares the clinical features of patients who did and did not fulfill the provisional PRINTO criteria. Demographics, family history of SpA, and location of back pain were not statistically different between the two groups. More patients who fulfilled the PRINTO criteria had peripheral arthritis, enthesitis, acute anterior uveitis, a polyarticular disease course, pain with deep palpation or maneuver of the SIJ, and morning stiffness lasting 15 minutes or longer. HLA-B27 positivity, which is a minor PRINTO criterion, was also significantly higher in those who fulfilled the PRINTO criteria. More patients who did not meet the PRINTO criteria had complaints of sacral or buttock pain than those who met PRINTO criteria. Table 2 compares the MRI features at the SIJ of patients who did and did not fulfill the provisional PRINTO criteria. There were no significant differences in the prevalence of inflammatory or structural lesions on MRI between the groups.

The proportion of patients meeting each of the major and minor PRINTO criteria are shown in Table 3. A total of 41 patients (26%) classified as having axJSpA did not have peripheral disease manifestations of arthritis or enthesitis, thereby making fulfillment of any of the three major PRINTO criteria impossible. Of these patients, all 41 (100%) had objective evidence of unequivocal inflammatory and/or structural changes typical of axJSpA on pelvic MRI. Of the 10 patients who did have arthritis or enthesitis but still did not fulfill a major criterion, none fulfilled the minor criteria for inflammatory back pain of at least three months as reported by

Table 1. Clinical features of patients with axial disease*

Characteristic	All		PRINTO (−)		PRINTO (+)		*P* value
	N	n (%)	N	n (%)	N	n (%)	
Age at reference date, Median (IQR), y	158	15.0 (12.6–16.8)	51	14.4 (12.0–16.3)	107	15.4 (12.9–17.0)	0.17
Male	158	109 (69.0)	51	32 (62.7)	107	77 (72.0)	0.24
Family history of HLA-B27 disease	142	34 (23.9)	50	11 (22.0)	92	23 (25.0)	0.70
Pain							
Sacral/buttock	158	74 (46.8)	51	30 (58.8)	107	44 (41.1)	**0.04**
Hip/groin	158	98 (62.0)	51	31 (60.8)	107	67 (62.6)	0.82
Insidious onset	139	119 (85.6)	41	32 (78.0)	98	87 (88.8)	0.10
Duration							
≥6 weeks	152	137 (90.1)	50	47 (94.0)	102	90 (88.2)	0.26
≥12 weeks	152	96 (63.2)	50	37 (74.0)	102	59 (57.8)	0.05
Stiffness ≥15 minutes	109	67 (61.5)	30	13 (43.3)	79	54 (68.4)	**0.02**
Clinical characteristics							
SI pain on examination[a]	158	81 (51.3)	51	19 (37.3)	107	62 (57.9)	**0.01**
History of peripheral arthritis	158	86 (54.4)	51	6 (11.8)	107	80 (74.8)	**0.00**
History of enthesitis	158	68 (43.0)	51	4 (7.8)	107	64 (59.8)	**0.00**
HLA-B27	149	94 (63.1)	49	24 (49.0)	100	70 (70.0)	**0.01**
Polyarticular arthritis	158	18 (11.4)	51	2 (3.9)	107	16 (15.0)	**0.05**
Acute anterior uveitis	158	9 (5.7)	51	0 (0.0)	107	9 (8.4)	**0.03**

* Bolded values represent *P* values below 0.05. IQR, interquartile range; PRINTO, Paediatric Rheumatology International Trials Organisation; SI, sacroiliac.
[a] Pain with deep palpation or flexion abduction and external rotation/Mennell's sign/Gaenslen's maneuver was included.

the physician submitting the case report. None of these 10 patients met the ASAS definition either, primarily because most of the components of the criteria were not available in the patient chart. Nine of the 10 patients reported lumbar, sacral, buttock, hip, or groin pain of at least 12 weeks' duration.

The second major criterion of the provisional PRINTO enthesitis/spondylitis-related arthritis is primarily geared toward inclusion of children with axial disease. However, only 65 children (41.1%) who were classified as having axJSpA fulfilled this criterion. A total of 137 youth (86.7%) classified as having axJSpA had low-back/buttock/hip/groin pain for ≥6 weeks, 96 (60.8%) had low-back/buttock/hip/groin pain for ≥12 weeks, 93 (58.9%) had inflammatory back pain for ≥12 weeks (as defined by the treating physician), and 28 (17.7%) had the necessary component data available and met ASAS criteria for inflammatory back pain for ≥12 weeks. A total of 132 patients with axJSpA (83.5%)

Table 2. Imaging features of patients with axial disease*

Characteristic	All		PRINTO (−)		PRINTO (+)		*P* value
	N	n (%)	N	n (%)	N	n (%)	
Imaging experts' MRI findings: inflammatory lesions							
Inflammation in subchondral bone marrow	158	141 (89.2)	51	44 (86.3)	107	97 (90.7)	0.41
Inflammation at site of an erosion cavity	158	89 (56.3)	51	30 (58.8)	107	59 (55.1)	0.66
Inflammation in SIJ capsule	158	31 (19.6)	51	10 (19.6)	107	21 (19.6)	1.00
Joint space enhancement on contrast[a]	74	29 (39.2)	27	12 (44.4)	47	17 (36.2)	0.48
Joint fluid	158	16 (10.1)	51	4 (7.8)	107	12 (11.2)	0.51
Enthesitis outside SIJ	158	34 (21.5)	51	11 (21.6)	107	23 (21.5)	0.99
Unequivocal inflammatory lesions[b]	156	139 (89.1)	50	44 (88.0)	106	95 (89.6)	0.76
Imaging experts' MRI findings: structural lesions							
Sclerosis	158	70 (44.3)	51	24 (47.1)	107	46 (43.0)	0.63
Erosion	158	132 (83.5)	51	43 (84.3)	107	89 (83.2)	0.86
Fatty lesion	158	19 (12.0)	51	3 (5.9)	107	16 (15.0)	0.10
Fat metaplasia in an erosion cavity	158	13 (8.2)	51	2 (3.9)	107	11 (10.3)	0.17
Ankylosis	158	4 (2.5)	51	0 (0.0)	107	4 (3.7)	0.16
Unequivocal structural lesions[c]	155	136 (87.7)	50	46 (92.0)	105	90 (85.7)	0.26

* MRI, magnetic resonance imaging; PRINTO, Paediatric Rheumatology International Trials Organisation; SIJ, sacroiliac joint.
[a] Only patients who had MRI with contrast were included.
[b] Unequivocal evidence of inflammatory lesions typical of axial disease was bone marrow edema in at least three SIJ quadrants across all SIJ MRI slices.
[c] Unequivocal evidence of structural lesion(s) typical of axial disease was erosion in at least three quadrants or sclerosis or fat lesion in at least two SIJ quadrants or backfill or ankylosis in at least two joint halves across all SIJ MRI slices.

Table 3. Proportion of patients with juvenile SpA and axial disease fulfilling each PRINTO enthesis/spondylitis-related juvenile idiopathic arthritis criterion*

Criterion	N	n (%)
Peripheral arthritis and enthesitis	158	37 (23.4)
Peripheral arthritis	158	86 (54.4)
Enthesitis	158	68 (43.0)
Arthritis or enthesitis, plus at least three months of inflammatory back pain and sacroiliitis on imaging	158	65 (41.1)
Peripheral arthritis or enthesitis	158	117 (74.1)
At least three months of inflammatory back pain	156	93 (59.6)
Sacroiliitis on imaging	158	158 (100.0)
Arthritis or enthesitis plus at least two of the following: SIJ tenderness, inflammatory back pain, presence of HLA-B27 antigen, acute (symptomatic) anterior uveitis, or history of SpA in a first-degree relative	158	95 (60.1)
Peripheral arthritis or enthesitis	158	117 (74.1)
SIJ tenderness	156	105 (67.3)
Inflammatory back pain	147	100 (68.0)
Presence of HLA-B27 antigen	149	94 (63.1)
Acute (symptomatic) anterior uveitis (ever)	158	9 (5.7)
History of SpA in a first-degree relative	147	26 (17.7)

* PRINTO, Paediatric Rheumatology International Trials Organisation; SIJ, sacroiliac joint; SpA, spondyloarthritis.

met the definition of imaging typical of axial disease in the PRINTO criteria.[12]

We tested how modifications to the second major PRINTO enthesitis/spondylitis-related JIA criterion impacted the proportion of children with axJSpA who could be captured by the criteria (Table 4). The proportion of children classified with axJSpA captured by the second criterion increased by 31.1% and 45.6% by modifying the second criterion to "at least six weeks of low-back/buttock pain plus imaging typical of axJSpA" or "at least six weeks of low-back/buttock/hip/groin pain plus imaging typical of axJSpA," respectively. Both modifications facilitated capture of more than 93% of children classifiable as having axJSpA.

DISCUSSION

Based on the recently validated criteria for axial involvement in children with SpA,[3] current ILAR enthesitis-related arthritis and PRINTO enthesitis/spondylitis-related JIA classification criteria would miss between one-quarter and one-third of such patients, and about 20% of them would not fit in any recognized ILAR or PRINTO categorization. Further, patients fulfilling or not fulfilling the PRINTO criteria had indistinguishable pelvic MRI findings. Coverage of axJSpA could be improved by relaxing PRINTO criteria to allow for six weeks' duration of inflammatory pain in the back, buttocks, or hips and by not requiring the presence of peripheral arthritis or enthesitis.

The ILAR JIA classification criteria have faced a number of criticisms, one of which being that many children who primarily have axial disease do not fulfill the enthesitis-related, psoriatic, or undifferentiated arthritis ILAR criteria.[1,13] Facilitating classification has implications for inclusion in clinical trials and may also ultimately impact access to medications. The clinical and imaging manifestations of axial disease were not significantly different between those who did and did not fulfill the provisional PRINTO

Table 4. Iterations of PRINTO enthesitis/spondylitis-related arthritis criterion 2*

Criteria	Fulfilling criterion, n (%)	Classified as enthesitis/spondylitis-related JIA when added to criteria 1 and 3, n (%)
Provisional PRINTO criterion 2: arthritis or enthesitis, plus at least three months of inflammatory back pain[a] and sacroiliitis on imaging	65 (41.1)	107 (67.7)
Potential alternate criteria		
At least three months of inflammatory back pain[a] plus imaging typical of axial disease in patients with juvenile SpA[b]	93 (58.9)	135 (85.4)
At least three months of low-back/buttock pain plus imaging typical of axial disease in patients with juvenile SpA[b]	81 (51.3)	142 (89.9)
At least six weeks of low-back/buttock pain plus imaging typical of axial disease in patients with juvenile SpA[b]	114 (72.2)	147 (93.0)
At least six weeks of low-back/buttock/hip/groin pain plus imaging typical of axial disease in patients with juvenile SpA[b]	137 (86.7)	152 (96.2)

* JIA, juvenile idiopathic arthritis; MRI, magnetic resonance imaging; PRINTO, Paediatric Rheumatology International Trials Organisation; SIJ, sacroiliac joint; SpA, spondyloarthritis.
[a] Inflammatory back pain was designated by the submitting physician.
[b] Imaging typical of axial disease in patients with juvenile SpA was defined as at least one of the following: unequivocal evidence of inflammatory lesions on MRI typical of axial disease in patients with juvenile SpA (bone marrow edema in at least three SIJ quadrants across all SIJ MRI slices); unequivocal evidence of structural lesion(s) on MRI typical of axial disease in patients with juvenile SpA (erosion in at least three quadrants or sclerosis or fat lesion in at least two SIJ quadrants or backfill or ankylosis in at least two joint halves across all SIJ MRI slices); or in the absence of pelvic MRI, unequivocal evidence of structural lesions on radiography typical of axial disease in patients with juvenile SpA (erosion, sclerosis, or ankylosis [partial or complete]) that must include at least one iliac bone. When sclerosis is present in isolation, if measurable, it should extend ≥5 mm from the joint surface. The decision may be influenced by the presence of other lesions, which in themselves do not suffice to meet the criterion.

criteria. The phenotypic differences between those who were and were not classified by the provisional PRINTO criteria were primarily related to peripheral disease manifestations. Modification of the second major criterion of the PRINTO provisional criteria could greatly facilitate capture of most children with primarily axial disease.

There are several aspects of the provisional PRINTO criteria that present challenges to classifying children with axial disease. First, to fulfill the second major criterion, children must have peripheral disease—either arthritis or enthesitis. In the cohort of children classified as having axJSpA presented herein, approximately one-fourth did not have any peripheral disease manifestations. By eliminating the peripheral disease requirement in this criterion, an additional 17.7% of the cohort with axJSpA was classifiable by the provisional PRINTO criteria. Second, to fulfill the second major criterion, children must have a back pain duration of three months or longer. The international clinical expert panel involved in the development and validation of the axJSpA criteria achieved >80% agreement that duration of pain for at least six weeks was sufficient; additionally, the definition of chronic (peripheral) arthritis is a duration of at least six weeks and that cohesiveness in definitions may improve the utility of the criteria. Third, the provisional PRINTO criteria only consider "back" pain. In the development of the axJSpA criteria, many symptomatic children with axJSpA did not describe their pain as originating in the back but instead localized pain to the buttocks, groin, and hip. Last, to fulfill the second PRINTO criterion, back pain must meet the ASAS defintion of inflammatory back pain.[8] These criteria have not been validated in children, and applying the adult criteria for inflammatory back pain has historically had low sensitivity and specificity in this population.[14] Only 73% of the patients in this cohort had enough data to assess for the ASAS definition, and of those, 41.7% met the inflammatory back pain definition. The axial SpA criteria for adult patients incorporate chronic low-back pain of at least three months' duration as an essential requirement.[15] However, the back pain does not need to be inflammatory. Inflammatory back pain is one of the clinical SpA variables that may contribute to classification of patients as having axSpA but is not an essential variable because its inclusion does not enhance the performance of the criteria.

There are several limitations to this study. The study was cross-sectional and included data from existing patients. Therefore, some elements of the standardized case report form were unavailable. However, data missingness was minimal, and details that were missing reflect aspects of the history and/or examination that perhaps are not uniformly valued as being clinically important in the assessment of axial disease in children. There were no missing data for imaging, and an added strength was that all imaging was interpreted by a central imaging team. Another limitation is that the proposed modifications to the second PRINTO major criteria may not work as well when applied to the broader population with JSpA. Only children with a clinical JSpA diagnosis and who met axJSpA classification criteria were included in this analysis. However, the axJSpA criteria were developed leveraging a group of children with a clinical diagnosis of JSpA and suspected axial disease. The performance of the modification will ultimately need to be tested in a population inclusive of children with JSpA with and without axial disease.

In conclusion, one-third of children diagnosed clinically with JSpA who met axJSpA classification criteria remain unclassifiable by the provisional PRINTO criteria. There were no imaging differences between those who did and did not meet PRINTO criteria. The phenotypic differences of children with axJSpA between those who are and are not classified by the PRINTO criteria were primarily limited to peripheral disease manifestations and HLA-B27 positivity. Modification of the second major PRINTO provisional criterion, which is focused primarily on those with axial manifestations, may facilitate capture of youth with JSpA manifested primarily by axial disease.

ACKNOWLEDGMENTS

The authors would like to warmly remember and acknowledge Dr Naden for his methodologic expertise and guidance for the first several stages of this project. Dr Naden was able to contribute his knowledge from the other American College of Rheumatology Classification Projects (systemic lupus erythematosus, IgG4, and gout) to guide and inform the preliminary methodology for this project. Dr Naden sadly passed away in the summer of 2020, at which time the role of expert methodologist was filled by author AMH, for whom the other authors are also very grateful.

The authors thank Joel Paschke (CARE Arthritis) for preparing the imaging review modules for web-based assessment and data entry and the following collaborators for their contributions of patient clinical data and MRI scans: Matthew L. Stoll and Livie Timmerman (University of Alabama at Birmingham); Gerd Horneff, Giulia Armaroli, Ariane Klein, and Rebekka Heidebrecht (Asklepios Clinic Sankt Augustin); Hemalatha Srinivasalu (National Institute of Arthritis and Musculoskeletal and Skin Diseases [NIAMS], National Institutes of Health [NIH], and Children's National Hospital); Sarah Litt (Children's National Hospital); Manuk Manukyan (NIAMS, NIH); Judith A. Smith, Thomas P. Callaci, and John W. Garrett (University of Wisconsin–Madison); Joke Dehoorne (Gent University Hospital); Nadia Aikawa (University of Sao Paulo); Flavio Sztajnbok (University Hospital Pedro Ernesto); Edward Oberle (Nationwide Children's Hospital); Daniel Lovell and Megan Quinlan-Waters (Cincinnati Children's Hospital Medical Center); Sujata Sawhney and Devang Pandya (Sir Ganga Ram Hospital); Benson Ogunjimi (Antwerp University Hospital and Ziekenhuisnetwerk Antwerpen); Ilaria Pagnini and Edoardo Marrani (Meyer Children's Hospital); and Boris Hugle (German Center for Pediatric and Adolescent Rheumatology).

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Weiss confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Petty RE, Southwood TR, Manners P, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31(2): 390–392.
2. Martini A, Ravelli A, Avcin T, et al. Toward new classification criteria for juvenile idiopathic arthritis: first steps, Pediatric Rheumatology International Trials Organization international consensus. J Rheumatol 2019;46(2):190–197.
3. Weiss PF, Brandon TG, Aggarwal A, et al. Classification criteria for axial disease in youth with juvenile spondyloarthritis. Arthritis Rheumatol 2024;76:1797–1808.
4. Weiss P, Brandon T, Aggarwal A, et al. Development, refinement and weighting of candidate criteria for axial disease in juvenile spondyloarthritis: an international collaboration. Arthritis Rheumatol 2022;74-(suppl 9).
5. Dubreuil M, Deodhar AA. Axial spondyloarthritis classification criteria: the debate continues. Curr Opin Rheumatol 2017;29(4):317–322.
6. Beukelman T, Nigrovic PA. Juvenile idiopathic arthritis: an idea whose time has gone? J Rheumatol 2019;46(2):124–126.
7. Lee JJY, Eng SWM, Guzman J, et al. A comparison of International League of Associations for Rheumatology and Pediatric Rheumatology International Trials Organization classification systems for juvenile idiopathic arthritis among children in a canadian arthritis cohort. Arthritis Rheumatol 2022;74(8):1409–1419.
8. Sieper J, van der Heijde D, Landewe R, et al. New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). Ann Rheum Dis 2009;68(6):784–788.
9. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377–381.
10. Weiss PF, Brandon TG, Lambert RG, et al. Data-driven magnetic resonance imaging definitions for active and structural sacroiliac joint lesions in juvenile spondyloarthritis typical of axial disease: a cross-sectional international study. Arthritis Care Res (Hoboken) 2023;75: 1220–1227.
11. Weiss PF, Brandon TG, Lambert RG, et al. Consensus-driven definition for unequivocal sacroiliitis on radiographs in juvenile spondyloarthritis. J Rheumatol 2023;50:1173–1177.
12. Lambert RGW, Bakker PAC, van der Heijde D, et al. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis 2016;75(11): 1958–1963.
13. Colbert RA. Classification of juvenile spondyloarthritis: enthesitis-related arthritis and beyond. Nat Rev Rheumatol 2010;6(8):477–485.
14. Weiss PF, Colbert RA. Reply. Arthritis Care Res (Hoboken) 2016; 68(8):1213–1214.
15. Rudwaleit M, van der Heijde D, Landewe R, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68(6):777–783.

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 792–800
DOI 10.1002/acr.25501

American College of Rheumatology
Empowering Rheumatology Professionals

Priority Musculoskeletal Health Research Questions for People With Generalized Joint Hypermobility: An International Delphi Study

Sara Habibian,[1] Verity Pacey,[2] Cliffton Chan,[1] Alan J. Hakim,[3] and Cylie M. Williams[4]

Objective. This study aimed to identify the top 10 international research priorities for musculoskeletal health of people with generalized joint hypermobility.

Methods. A three-round Delphi method using an online survey was implemented. Three participant stakeholder groups were eligible for inclusion: (1) people with lived experience of joint hypermobility or their carers, (2) health care professionals, and (3) researchers with experience working with individuals with hypermobility. Participants provided up to three priority research questions in round 1. In round 2, participants prioritized 10 research questions from the unique questions proposed in round 1. In round 3, participants were presented with the top 10 questions from the overall cohort and for their stakeholder group(s) and asked to rank these in order of importance.

Results. Round 1 commenced with 396 participants who provided 958 individual questions, which reduced to 210 unique questions following data cleaning. There were 257 participants (65% of 396) in round 2, and 249 participants (63% of 396, lived experience n = 230, health care professionals n = 73, and researchers n = 21) in round 3. The overall top-ranked question was, "How can we prevent disability, pain, and poor quality of life associated with the musculoskeletal comorbidities of symptomatic generalized joint hypermobility?" Specific stakeholder group priority research questions varied. People with lived experience prioritized treatment questions, whereas health care professionals and researchers prioritized service-impact and utilization research questions.

Conclusion. Priority research questions relating to musculoskeletal health of people with generalized joint hypermobility have been internationally identified. These questions provide a future focus for meaningful and necessary research in this field.

INTRODUCTION

Joint hypermobility is common in the general population, with the prevalence ranging between 5% to 40% in children and 10% to 20% of adults.[1] Excessive movement of joints beyond their normal range is the descriptor of joint hypermobility. Associated concerns for individuals with joint hypermobility include joint and soft tissue injury and health issues.[1,2] In certain groups, joint hypermobility can become symptomatic and present as a feature of heritable connective tissue disorders. Common heritable connective tissue disorders are hypermobility spectrum disorder and hypermobile Ehlers-Danlos syndrome (EDS), with an estimated combined prevalence of 19 in 10,000 people.[3] Musculoskeletal problems are the most common concern and one of the most important difficulties faced by and prioritized by this population because of the significant impact on long-term health and quality of life.[4–7]

People with symptomatic generalized joint hypermobility experience various musculoskeletal issues, including but not limited to acute and chronic pain and fatigue, joint instability, soft tissue injury, muscle weakness, balance issues, joint proprioception difficulties, movement dysfunction, and pelvic floor dysfunction.[7] Multiple population studies involving adults and symptomatic children have demonstrated the breadth of these issues across the

Supported by the International Macquarie University Research Excellence Scholarship (scholarship allocation number 20235605).

[1]Sara Habibian, MSci, MRes, Cliffton Chan, BPhysio(Hons1), PhD: Macquarie University, Sydney, New South Wales, Australia; [2]Verity Pacey, BAppSci(Phty), PhD: Macquarie University, Sydney, New South Wales, and Monash University, Melbourne, Victoria, Australia; [3]Alan J. Hakim, MD: The Ehlers-Danlos Society, London, United Kingdom; [4]Cylie M. Williams, BAppSci(Pod), PhD: Monash University, Melbourne, Victoria, Australia.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25501).

Author disclosures, a graphical abstract, and a video abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25501.

Address correspondence via email to Verity Pacey, BAppSci(Phty), PhD, at verity.pacey@mq.edu.au.

Submitted for publication June 23, 2024; accepted in revised form January 14, 2025.

SIGNIFICANCE & INNOVATIONS

- Preventing disability, pain, and poor quality of life associated with the musculoskeletal comorbidities of symptomatic generalized joint hypermobility is the top identified research priority.
- Other highly ranked questions related to disease etiology, disease prevention, diagnosis, prognosis, treatment, meaning, and service-impact.
- People with lived experience prioritized treatment research questions, whereas health care professionals and researchers prioritized service-impact and utilization research questions.

lifespan.[5,8–11] A wide range of therapeutic interventions have been suggested to improve musculoskeletal health and the overall quality of life of people with symptomatic generalized joint hypermobility.[12–14] However, there remains a lack of consistency in the research methodology and findings addressing intervention approaches for individuals with generalized joint hypermobility. This creates confusion and makes decision-making difficult for patients and their health care providers, despite the considerable number of emerging interventional studies in this area.[15,16]

In optimizing care, it is valuable to determine the most concerning problems from different stakeholder perspectives including people with lived experience of a condition, health care professionals, and researchers. This guides and prioritizes further research aimed at informing key areas required to minimize and prevent musculoskeletal ill-health and associated reduced quality of life. To the best of our knowledge, there has been no investigation that has formally identified the key research questions deemed relevant by stakeholders in the area of musculoskeletal health related to generalized joint hypermobility. This study sets out a top 10 list of research priorities for consumers (people with lived experience), health care professionals, and researchers, aiming to guide the prioritization of resources in this field.

MATERIALS AND METHODS

Design. The study was designed as a three-round modified Delphi online survey to generate the top 10 priority research questions according to three stakeholder groups. The Delphi technique was employed to purposefully gain international consensus from multiple perspectives, allowing equal input from all participants with anonymity.[17,18] This study was approved by Monash University Human Research Ethics Committee (MUHREC37840). Reporting of results was informed using the Conducting and Reporting Delphi Studies checklist.[19]

Participants. People living with symptomatic generalized joint hypermobility or their carer (consumers), health care professionals, and researchers were invited to complete the three rounds. All participants were required to consent to provision of their email for tracking completion across rounds, being above 18 years of age, and provided informed consent following reading the information and consent forms. There were no reimbursements for participation, and withdrawal was possible at any time through nonresponse.

Recruitment. Participants were recruited through direct email, links on websites, and social media posts (eg, LinkedIn, Facebook, Instagram, and X) inviting participation by organizations based within the United Kingdom and United States of America who play an advocacy role in musculoskeletal health internationally (eg, The Ehlers-Danlos Society, working groups of The International Consortium on EDS and Hypermobility Spectrum Disorders, and Pediatric Global Musculoskeletal Taskforce Working Group). Snowballing recruitment was encouraged in recruitment requests to professional networks and on social media. Recruitment commenced for round 1 in June 2023, with round 3 closing in November 2023.

Procedures. The survey was administered using the online platform Qualtrics (Provo, Utah). Round 1 was open for five weeks, with subsequent rounds open for two to four weeks. At the commencement of rounds 2 and 3, participants were emailed instructions for completion and reminded weekly of the closing date. At the start of each round, participants were provided additional information to support their responses (eg, terminology definitions, Supplementary Data S1). Before commencement of the study, the survey was piloted within the research team, which included an individual affected with symptomatic generalized joint hypermobility, a carer of an individual affected with symptomatic generalized joint hypermobility, and health professionals from three countries and three different professions.

Round 1 design and analysis. Round 1 contained questions relating to demographic information including sex, age, country of residence, and the stakeholder group(s) participants identified with (consumer, health care professional, or researcher). Sex was self-reported from one of four categories (female, male, nonbinary, or other), with participants able to provide an open-ended response if they chose other. Additional information was collected based on the responses. For example, health care professionals were asked to provide profession, country of work, and working sector(s) (public or private). Survey logic functions were used to raise the additional questions with the relevant participants. Participants were asked, "What questions about musculoskeletal health and generalized joint hypermobility would you most like to see answered by research?" Participants had the option of providing up to three questions of any type, such as treatment, diagnosis, and prognosis. The Qualtrics language function was used to allow participants to complete the survey in English,

Persian, or Chinese, and translations were undertaken by the study authors who are native speakers of these languages (SH, CC).

Responses were exported into Microsoft Excel (version 23, Microsoft Corporation). Each response was independently analyzed by a combination of two authors (SH and VP, CC or CMW). Questions unrelated to generalized joint hypermobility and musculoskeletal health were excluded. All questions were checked to ensure that they had not been previously answered through a published systematic review with meta-analysis in a peer-reviewed journal. Categorization adapted from the original population, intervention, control, and outcomes (PICO) was used to organize questions based on the phrasing by participants.[20,21] This classification system included questions relating to etiology, diagnosis, prevention, treatment, prognosis, meaning, and service-impact and utilization. All excluded questions and all categorization allocations were independently reviewed by each of the authors pairs, and disagreement was resolved by consensus through a third author.

Included questions were distributed among three pairs of authors (SH and either VP, CC, or CMW) to consolidate and reformat for consistency based on specific PICO categories. Several principles were followed while ensuring minimal editing of the questions with the aim of (1) increasing the clarity of the main intent of the question, (2) using culturally sensitive or person-centered language, and (3) ensuring the final output conformed

Table 1. Summary of participants' characteristics

Characteristics	Round 1	Round 2	Round 3
Total sample, n (%)	396 (100.0)	257 (64.8)	249 (62.8)
Age (yr), n (%)			
Younger than 20	10 (2.5)	7 (2.7)	8 (3.2)
20–29	50 (12.6)	25 (9.7)	28 (11.3)
30–39	109 (27.5)	82 (32)	69 (27.7)
40–49	93 (23.5)	54 (21.0)	57 (22.9)
50–59	88 (22.2)	57 (22.2)	55 (22.1)
60–69	32 (8.1)	25 (9.7)	25 (10.0)
70–79	13 (3.3)	7 (2.7)	7 (2.8)
80 years or older	1 (0.3)	–	–
Sex, n (%)			
Female	353 (89.1)	228 (88.7)	221 (88.8)
Male	26 (6.6)	19 (7.4)	18 (7.2)
Nonbinary	8 (2.0)	4 (1.5)	5 (2.0)
Preferred to describe	9 (2.3)	6 (2.3)	5 (2.0)
Stakeholder groups, n (%)			
Consumers: people living with generalized joint hypermobility or carer	359 (72.4)	230 (69.0)	230 (71.0)
Health care professionals	105 (21.2)	78 (23.0)	73 (22.5)
Researchers	32 (6.4)	27 (8.0)	21 (6.5)
Geographic diversity, n (%)			
North America	231 (58.3)	141 (54.9)	137 (55.0)
Australia	90 (22.7)	60 (23.3)	62 (24.9)
Europe	63 (15.9)	47 (18.3)	41 (16.5)
Asia	9 (2.3)	7 (2.7)	7 (2.8)
Africa	2 (0.5)	2 (0.8)	2 (0.8)
South America	1 (0.3)	–	–
Health care professional group breakdown, n (%)	105	78	67
Physical therapist	53 (50.5)	43 (55.0)	39 (58.1)
Physician	18 (17.1)	13 (16.7)	11 (16.4)
Occupational therapist	5 (4.8)	2 (2.6)	2 (3.0)
Podiatrist	5 (4.8)	4 (5.1)	2 (3.0)
Nurse	4 (3.8)	2 (2.6)	4 (6.0)
Dietitian/nutritionist	4 (3.8)	3 (3.8)	–
Psychologist	4 (3.8)	2 (2.6)	2 (3.0)
Surgeon	2 (1.9)	1 (1.3)	1 (1.5)
Chiropractor	2 (1.9)	1 (1.3)	1 (1.5)
Other[a]	8 (7.6)	7 (9.0)	5 (7.5)
Health care sector, n (%)			
Private	65 (61.9)	50 (64.1)	42 (62.7)
Public	34 (32.4)	24 (30.8)	21 (31.3)
Other[b]	6 (5.7)	4 (5.1)	4 (6.0)

[a] Includes sports therapist (n = 1), medical writer (n = 1), speech therapist (n = 1), board-certified patient advocate (n = 1), craniosacral therapist (n = 1), psychotherapist (n = 1), mental health counselor (n = 1), and genetic counseling (n = 1).
[b] Includes social aid, waiver programs for adults with disabilities, and university-affiliated hospital.

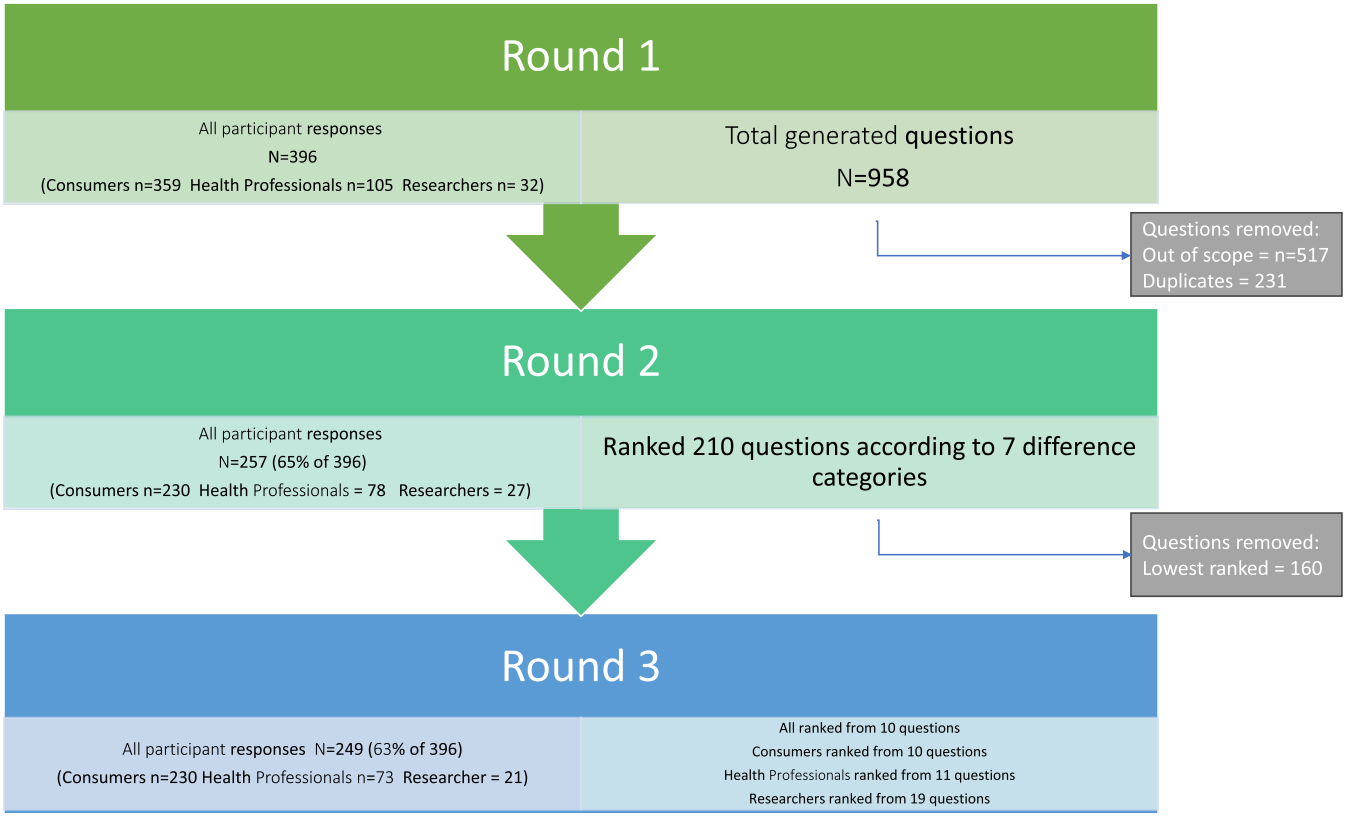

Figure 1. Flow of participants and research questions through the Delphi rounds.

to a question format. Author pairs independently edited the reformatted questions and then reviewed the outputs as a group. Any disagreement in editing was resolved by consensus at the group discussions. The final questions from round 1 were presented in their respective themes to the participants in round 2.

Round 2 design and analysis. Participants were presented with the full list of questions and instructed to choose as many questions as they deemed significant from the entire list. To reduce selection bias, the randomized function within the survey software was used to present the questions in varied order between participants. The survey flow then allowed participants to narrow their selections to a maximum of 10 questions. When requested, participants who contacted the research team reporting difficulties reviewing the high number of questions were supported via email by one of authors (CMW) to narrow down their choices by representing multiple smaller lists within the time period of round 2. The top 10 questions were identified by calculating the frequency of all participant responses. In addition, the top 10 questions for each group were identified in the same manner. When there was an exact equal frequency of responses to a question within the top 10, each of these questions were returned to participants in round 3.

Round 3 design and analysis. All participants, regardless of whether they had participated in round 2, were invited to take part in round 3 to rank the top 10 questions overall and to rank the top 10 questions related to their stakeholder group(s). Each question was allocated a score depending on the ranking, and total points calculated for each research question based on all responses. This method was repeated for questions within each group.

RESULTS

A total of 396 participants provided 958 questions in round 1. Table 1 shows the distribution of overall participant demographic data and a detailed breakdown of the health care professional group. Almost all responses provided by participants were in English (n = 394, 99.5%), with two responses in Persian (0.5%). Figure 1 summarizes each Delphi round leading to the identification of the top 10 questions.

The 958 questions were cleaned by amalgamating 231 questions due to repetition and the removal of 517 considered outside the scope of this research subject such as causes or treatments of nonmusculoskeletal symptoms (eg, autoimmune disorders, gastrointestinal conditions, and orthostatic intolerance). Participants were provided with information about how these questions were cleaned with examples (Supplementary Data S2). The final 210 questions were grouped by PICO category of prevention (n = 10, 5% of questions), diagnosis (n = 30, 14%), prognosis (n = 11, 5%), etiology (n = 36, 17%), treatment (n = 80, 38%),

Figure 2. The top 10 research questions identified by all participants. EDS, Ehlers-Danlos syndrome; GJH, generalized joint hypermobility; hEDS, hypermobile Ehlers-Danlos syndrome; HSD, hypermobility spectrum disorders. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25501/abstract.

meaning (n = 10, 5%), and service-impact and utilization (n = 33, 16%) (Supplementary Data S3).

The overall top 10 priority research questions of the cohort covered all seven PICO categories (Figure 2). The top 10 priority research questions differed between stakeholder groups. The priority research questions from the consumers only perspective (Table 2) related to treatment and prevention questions, whereas both health care professionals (Table 3) and researchers (Table 4) prioritized questions regarding service-impact and utilization (eg, clinical practice guidelines).

DISCUSSION

To our knowledge this is the first study to explore and prioritize research questions regarding musculoskeletal health concerns in people with generalized joint hypermobility. Participants prioritized a broad range of questions across seven categories adapted from PICO.[20,21] These findings provide an important foundation for funders and research groups to drive research agendas that consider different stakeholder perspectives. The number of priority questions and their breadth across the

Table 2. Top 10 research questions by consumers only after round 3 (n = 186)*

Rank	Research question	Modified PICO category
1	What is the most effective treatment for pain and fatigue associated with GJH and EDS?	Treatment
2	How can we prevent disability, pain and poor quality of life associated with the musculoskeletal comorbidities of symptomatic GJH?	Prevention
3	What is the most effective way to strengthen muscles without postexercise side effects in people with GJH?	Treatment
4	What are the clinical practice guidelines for GJH, HSD, and hEDS across the lifespan?	Service-impact and utilization
5	What are the most common musculoskeletal comorbidities in people with GJH across the lifespan?	Diagnosis
6	What factors predict transition from asymptomatic to symptomatic GJH?	Prognosis
7	Why do people with GJH commonly have neurologic problems?	Etiology
8	What are the most limiting aspects of hypermobility in daily life activities as perceived by individuals with symptomatic GJH?	Meaning
9	Why do health professionals demonstrate a strong aversion toward individuals with chronic pain and/or fatigue in conjunction with GJH?	Meaning
10	What is the most effective way to cope with severe muscle spasms for people with GJH?	Treatment

* EDS, Ehlers-Danlos syndrome; GJH, generalized joint hypermobility; hEDS, hypermobile Ehlers-Danlos syndrome; HSD, hypermobility spectrum disorders; PICO, population, intervention, control, and outcomes.

Table 3. Top 10 research questions by health care professionals after round 3 (n = 73)*

Rank	Research question	Modified PICO category
1	What are the clinical practice guidelines for GJH, HSD, and hEDS across the lifespan?	Service-impact and utilization
2	What are the clinical guidelines for developing strength, balance, stability, and proprioception in children with GJH?	Service-impact and utilization
3	What are the best screening tools to identify GJH?	Diagnosis
4	What is the role of the sensory system, specifically the proprioceptive system, in people with GJH?	Prognosis
5	How early can children be identified with GJH or diagnosed with HSD or hEDS?	Diagnosis
6	What is the impact of puberty on the musculoskeletal health of individuals with GJH?	Prognosis
7	What are the most common musculoskeletal comorbidities in people with GJH across the lifespan?	Diagnosis
8	What factors predict transition from asymptomatic to symptomatic GJH?	Prognosis
9	What is the prevalence of nutritional deficiencies that impact connective tissue formation and muscle function/recovery/repair in people with GJH?	Diagnosis
10	Do individuals with GJH have a slower rate of tissue healing than individuals without GJH?	Prognosis

* GJH, generalized joint hypermobility; hEDS, hypermobile Ehlers-Danlos syndrome; HSD, hypermobility spectrum disorders; PICO, population, intervention, control, and outcomes.

categories also allows for a strong staged research program to be adopted, with many of the questions interlinked across the translational medicine and implementation science landscape from etiology to service delivery.

The highest overall ranked question was, "How can we prevent disability, pain, and poor quality of life associated with the musculoskeletal comorbidities of symptomatic generalized joint hypermobility?" The question was in the prevention category. There is currently insufficient evidence regarding preventive measures for people with generalized joint hypermobility. Although there is some evidence to support a wide range of treatment options for musculoskeletal symptoms including pain, feelings of instability, and joint injuries in adults and children,[1,22,23] there remains a lack of high-quality evidence that support high-value and sustainable management and prevention of symptomatic generalized joint hypermobility.[4,16] This view is supported by a qualitative study reporting parents feeling conflicted about whether physical activity should be promoted or avoided for the musculoskeletal health of their child with syndromic-related generalized joint hypermobility.[24] In addition, although emerging, there is a dearth of research on the older person with joint hypermobility–related musculoskeletal disease aside from reported early onset of osteoarthritis.[25,26] Long-term prospective longitudinal studies of individuals presenting with generalized joint hypermobility before any symptomatology presenting are needed to determine the optimal ways to prevent symptomatology and musculoskeletal comorbidities.

The phrasing of the first ranked question encourages broad exploration, addressing both preventive approaches and the need to accurately define musculoskeletal comorbidities associated with generalized joint hypermobility. This question also prompts consideration of the most appropriate disability and

Table 4. Top 10 research questions by researchers after round 3 (n = 21)*

Rank	Research question	Modified PICO category
1	What are the clinical practice guidelines for GJH, HSD, and hEDS across the lifespan?	Service-impact and utilization
2	Why do some individuals with GJH develop pain and others do not?	Etiology
3	What factors predict transition from asymptomatic to symptomatic GJH?	Prognosis
4	What is the best way to apply, prescribe, and individualize exercise interventions for people with GJH to safely improve their overall health?	Service-impact and utilization
5	What is the most effective treatment for pain and fatigue associated with GJH and EDS?	Treatment
6	How can we prevent disability, pain, and poor quality of life associated with the musculoskeletal comorbidities of symptomatic GJH?	Prevention
7	What core data sets are appropriate for children with hypermobility and HSD or hEDS?	Service-impact and utilization
8	What is the most effective way to strengthen muscles without postexercise side effects in people with GJH?	Treatment
9	What are the red-flags that are identifiable early in life to predict pain onset, injury, or dysfunction in people with GJH?	Diagnosis
10	What early preventive strategies can be employed to avoid future impact of GJH such as joint injuries, pain, or damage in individuals with hEDS?	Prevention

* EDS, Ehlers-Danlos syndrome; GJH, generalized joint hypermobility; hEDS, hypermobile Ehlers-Danlos syndrome; HSD, hypermobility spectrum disorders; PICO, population, intervention, control, and outcomes.

quality of life measurements in clinical practice and research. These important factors relating to the question require definition to also answer many of the other ranked questions focused on treatment (questions 2 and 10), service-impact and utilization (questions 3 and 4), or prognosis (question 6). Therefore, there is an urgency in answering this top-ranked question because its wide impact on satisfying all stakeholders.

Health care professionals primarily focused on questions relating to prognosis and diagnosis, whereas other groups had broader priorities. This may reflect health care professional difficulties when faced with uncertainty in clinical practice[27] or the variability and complexity in the presentation of symptoms among people seeking care who have generalized joint hypermobility.[28] A recent study of health care providers' awareness of the EDS indicated inadequate education may cause diagnostic delay,[29] whereas less disease complications have been demonstrated to be associated with early diagnosis.[30] A study of knowledge and confidence in the diagnosis of EDS and an exploration of the impact of an education program showed that baseline knowledge and confidence in this field is low across the health care professional disciplines.[31] This is important to consider as the presentation and reasons for seeking care related to generalized joint hypermobility may vary during the lifespan.[23] There is also scarce research regarding prognosis in the peer-reviewed literature, further reinforcing the need to determine prognostic certainty as a research priority.[28]

Our results demonstrated that researchers and consumers indicated a similar priority to the question categories that focused on prevention, treatment, prognosis, service-impact, and utilization. We postulate one reason for this may be that there has been an increase in qualitative research recently with researchers engaging in-depth with consumers.[7,32,33]

Only consumers ranked a question in the "meaning" category as a priority. This was unsurprising given the impact that generalized joint hypermobility has on their lives[27,34] and the difficulties and challenges they face while seeking health care.[33–35] Multiple studies to date acknowledge a perceived lack of understanding and knowledge of the impact and appropriate management of generalized joint hypermobility by health care professionals providing a barrier to patient experience of quality care.[36,37] Further studies should be cognizant of this disparity when designing and involving different stakeholders.

This study is limited by factors such as low ethnic diversity from countries where English is not the primary language and some attrition after each round of survey. Prioritization of questions should be cautiously approached when applied to non-English speaking countries. Although there was a decreasing participation rate between rounds when the ranking occurred, this is often observed in research designs when ranking of large volumes of information is required.[38] Despite this decrease, the study response rate of >60% is comparable to other large Delphi surveys.[39] Although caregivers of individuals with hypermobility were eligible for inclusion within the current study, participation was limited to adults over the age of 18 years, which did not provide an independent voice for children or adolescents to contribute to developing research priorities. Finally, participants were not provided with any training materials on writing research questions before participation in this Delphi study, and there was not any opportunity for deliberation between stakeholders. This required the research team to review and reformat questions and provide explanatory terms for specialized medical terminology to participants. However, our study inclusion strategy did allow a large number of participants of all educational and cultural backgrounds worldwide to participate within the subsequent rounds of the Delphi.

In conclusion, this study has identified priority research questions from people with lived experience of generalized joint hypermobility, health care providers, and researchers. Researchers and

funding bodies should consider these questions and the different expectations of stakeholders when prioritizing research decisions. In our opinion, initial future studies should focus on the three question categories of prevention, treatment and service-impact, and utilization.

ACKNOWLEDGMENTS

Open access publishing facilitated by Macquarie University, as part of the Wiley - Macquarie University agreement via the Council of Australian University Librarians.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Pacey confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Nicholson LL, Simmonds J, Pacey V, et al. International perspectives on joint hypermobility: a synthesis of current science to guide clinical and research directions. J Clin Rheumatol 2022;28(6):314–320.
2. Tinkle B, Castori M, Berglund B, et al. Hypermobile Ehlers-Danlos syndrome (a.k.a. Ehlers-Danlos syndrome type III and Ehlers-Danlos-syndrome hypermobility type): clinical description and natural history. Am J Med Genet C Semin Med Genet 2017;175(1):48–69.
3. Ehlers-Danlos Society. Are the Ehlers-Danlos syndromes and hypermobility spectrum disorders rare or common? Accessed May 20, 2024. https://www.ehlers-danlos.com/is-eds-rare-or-common/.
4. Hakimi A, Bergoin C, Mucci P. What are the most important symptoms to assess in hypermobile Ehlers-Danlos syndrome? A questionnaire study based on the Delphi technique. Disabil Rehabil 2022;44(26):8325–8331.
5. Teran-Wodzinski P, Kumar A. Clinical characteristics of patients with hypermobile type Ehlers-Danlos syndrome (hEDS) and generalized hypermobility spectrum disorders (G-HSD): an online survey. Rheumatol Int 2023;43(10):1935–1945.
6. Garreth Brittain M, Flanagan S, Foreman L, et al. Physical therapy interventions in generalized hypermobility spectrum disorder and hypermobile Ehlers-Danlos syndrome: a scoping review. Disabil Rehabil 2024;46(10):1936–1953.
7. Albayrak İ, Yilmaz H, Akkurt HE, et al. Is pain the only symptom in patients with benign joint hypermobility syndrome? Clin Rheumatol 2015;34(9):1613–1619.
8. Jeong H-J, Engel JM, Wilwert O, et al. Pain characteristics and symptom management in children with hypermobile Ehlers-Danlos syndrome and hypermobility spectrum disorder. Phys Occup Ther Pediatr 2023;43(5):630–643.
9. Tran ST, Jagpal A, Koven ML, et al. Symptom complaints and impact on functioning in youth with hypermobile Ehlers-Danlos syndrome. J Child Health Care 2020;24(3):444–457.
10. Morris SL, O'Sullivan PB, Murray KJ, et al. Hypermobility and musculoskeletal pain in adolescents. J Pediatr 2017;181:213–221.e1.
11. Pacey V, Tofts L, Adams RD, et al. Quality of life prediction in children with joint hypermobility syndrome. J Paediatr Child Health 2015;51(7):689–695.
12. Smith TO, Bacon H, Jerman E, et al. Physiotherapy and occupational therapy interventions for people with benign joint hypermobility syndrome: a systematic review of clinical trials. Disabil Rehabil 2014;36(10):797–803.
13. Bale P, Easton V, Bacon H, et al. The effectiveness of a multidisciplinary intervention strategy for the treatment of symptomatic joint hypermobility in childhood: a randomised, single Centre parallel group trial (The Bendy Study). Pediatr Rheumatol Online J 2019;17(1):2.
14. Van Meulenbroek T, Conijn AEA, Huijnen IPJ, et al. Multidisciplinary treatment for hypermobile adolescents with chronic musculoskeletal pain. J Rehabil Med Clin Commun 2020;3(1):1000033.
15. Peterson B, Coda A, Pacey V, et al. Physical and mechanical therapies for lower limb symptoms in children with hypermobility spectrum disorder and hypermobile Ehlers-Danlos syndrome: a systematic review. J Foot Ankle Res 2018;11(1):59.
16. Palmer S, Davey I, Oliver L, et al. The effectiveness of conservative interventions for the management of syndromic hypermobility: a systematic literature review. Clin Rheumatol 2021;40(3):1113–1129.
17. Niederberger M, Köberich S, members of the DeWiss Network. Coming to consensus: the Delphi technique. Eur J Cardiovasc Nurs 2021;20(7):692–695.
18. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. PLoS Med 2011;8(1):e1000393.
19. Jünger S, Payne SA, Brine J, et al. Guidance on Conducting and Reporting Delphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med 2017;31(8):684–706.
20. Snowball R. Using the clinical question to teach search strategy: fostering transferable conceptual skills in user education by active learning. Health Libr Rev 1997;14(3):167–172.
21. Villanueva EV, Burrows EA, Fennessy PA, et al. Improving question formulation for use in evidence appraisal in a tertiary care setting: a randomised controlled trial [ISRCTN66375463]. BMC Med Inform Decis Mak 2001;1(1):4.
22. Nicholson LL, Chan C, Tofts L, et al. Hypermobility syndromes in children and adolescents: assessment, diagnosis and multidisciplinary management. Aust J Gen Pract 2022;51(6):409–414.
23. Song B, Yeh P, Nguyen D, et al. Ehlers-Danlos syndrome: an analysis of the current treatment options. Pain Physician 2020;23(4):429–438.
24. Berglund B, Nordström G, Lützén K. Living a restricted life with Ehlers-Danlos syndrome (EDS). Int J Nurs Stud 2000;37(2):111–118.
25. Gullo TR, Golightly YM, Flowers P, et al. Joint hypermobility is not positively associated with prevalent multiple joint osteoarthritis: a cross-sectional study of older adults. BMC Musculoskelet Disord 2019;20(1):165.
26. Anderson LK, Lane KR. Clinical trajectory of hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorders in older adults. J Am Assoc Nurse Pract 2023;35(10):605–612.
27. Clark NL, Johnson M, Rangan A, et al. The biopsychosocial impact of hypermobility spectrum disorders in adults: a scoping review. Rheumatol Int 2023;43(6):985–1014.
28. Feldman ECH, Homan KJ, Williams SE, et al. A narrative review of the literature on illness uncertainty in hypermobile ehlers-danlos syndrome: implications for research and clinical practice. Pediatr Rheumatol Online J 2023;21(1):121.
29. Jones JT, Black WR. Provider knowledge and experience in care, management, and education of pediatric Ehlers-Danlos syndrome. *Glob Pediatr* Health 2022;9:2333794X221112841.

30. Kalisch L, Hamonet C, Bourdon C, et al. Predictors of pain and mobility disability in the hypermobile Ehlers-Danlos syndrome. Disabil Rehabil 2020;42(25):3679–3686.
31. Schubart JR, Bascom R, Francomano CA, et al. Initial description and evaluation of EDS ECHO: an international effort to improve care for people with the Ehlers-Danlos syndromes and hypermobility spectrum disorders. Am J Med Genet C Semin Med Genet 2021;187(4): 609–615.
32. Bennett SE, Walsh N, Moss T, et al. Developing a self-management intervention to manage hypermobility spectrum disorders (HSD) and hypermobile Ehlers-Danlos syndrome (hEDS): an analysis informed by behaviour change theory. Disabil Rehabil 2022;44(18):5231–5240.
33. Simmonds JV, Herbland A, Hakim A, et al. Exercise beliefs and behaviours of individuals with joint hypermobility syndrome/Ehlers-Danlos syndrome - hypermobility type. Disabil Rehabil 2019;41(4): 445–455.
34. Scheper MC, Juul-Kristensen B, Rombaut L, et al. Disability in adolescents and adults diagnosed with hypermobility-related disorders: a meta-analysis. Arch Phys Med Rehabil 2016;97(12):2174–2187.
35. Halverson CME, Penwell HL, Francomano CA. Clinician-associated traumatization from difficult medical encounters: results from a qualitative interview study on the Ehlers-Danlos Syndromes. SSM Qual Res Health 2023;3:100237.
36. Bennett SE, Walsh N, Moss T, et al. The lived experience of joint hypermobility and Ehlers-Danlos syndromes: a systematic review and thematic synthesis. Phys Ther Rev 2019;24(1–2):12–28.
37. Langhinrichsen-Rohling J, Lewis CL, McCabe S, et al. They've been BITTEN: reports of institutional and provider betrayal and links with Ehlers-Danlos Syndrome patients' current symptoms, unmet needs and healthcare expectations. Ther Adv Rare Dis 2021;2: 26330040211022033.
38. McHale S, Cowie A, Brown S, et al. Research priorities relating to the delivery of cardiovascular prevention and rehabilitation programmes: results of a modified Delphi process. Open Heart 2023;10(1): e002248.
39. O'Connor L, McAuliffe E, Casey M, et al. Operationalising a modified Delphi study to progress quality care process nursing metrics for acute care. J Res Nurs 2022;27(7):655–676.

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 801–810
DOI 10.1002/acr.25487

"It's Just Good Science": A Qualitative Study Exploring Equity, Diversity, and Inclusion in Canadian Arthritis Research

Megan M. Thomas,[1] Mark Harrison,[2] Cheryl Barnabe,[3] Charlene E. Ronquillo,[4] J. Antonio Avina-Zubieta,[1] Anna Samson,[5] Michael Kuluva,[6] Natasha Trehan,[7] and Mary A. De Vera[1]

Objective. Despite knowledge that health outcomes vary according to patient characteristics, identity, and geography, including underrepresented populations in arthritis research remains a challenge. We conducted interviews to explore how researchers in arthritis have used equity, diversity, and inclusion (EDI) principles to inform their research.

Methods. Semistructured interviews were conducted with individuals who 1) have experience conducting arthritis research studies, 2) reside in and/or conduct their research in Canada, and 3) speak English or French. Participants were recruited using purposive and respondent-driven sampling. Interviews were conducted over video call and audio recordings were transcribed. Template analysis was applied to interview transcripts to explore participant experiences and perceptions of EDI in arthritis research.

Results. Participants (n = 22) identified that a lack of representation in arthritis research translates to the inability to provide comprehensive care. Participants emphasized considering EDI early in all arthritis research to effectively affect a study. Themes were categorized as benefits, barriers, and facilitators. The perceived benefits were the ability to generate knowledge and reduce health disparities. Barriers included mistrust from historically exploited populations, unintended consequences, lack of access to research opportunities, and logistical challenges. Facilitators included building community partnerships, curating diverse research teams, incentivizing researchers and funder support, and fostering humility in research environments.

Conclusion. Improving representation in research is needed to improve health outcomes for diverse groups of people living with arthritis. Identified barriers to EDI in research must be addressed and partnerships and supports must be facilitated to achieve more representation in arthritis research within Canada.

INTRODUCTION

Health outcomes for people with arthritis are influenced by many elements, including biologic (e.g., genetics), social, and environmental factors, referred to as the social determinants of health (SDOH).[1] Disparities in SDOH also are associated with exclusion and low participation from particular communities in underrepresented patients in research. Specifically, within rheumatology, studies have overrepresentation of White, middle-aged, female participants.[2–4] Although most studies will summarize participant demographics and the inclusion of age and sex is standard, reporting of additional demographic factors is rare and not standardized across studies, and there is typically no stratification of effect across groups, making it challenging to compare findings. The final, perhaps most pressing issue is that current recruitment and reporting standards limit our ability to apply evidence to practice for communities facing health inequities.

In Canada, rural and remote, Indigenous, older adults with frailty, first-generation immigrant and refugee, low-income and vulnerably housed, and diverse gender and sex populations experience inequities in arthritis care, such as diagnostic delays

Supported by the Canadian Institutes of Health Research (planning grant PCS-183488).

[1]Megan M. Thomas, MSc, J. Antonio Avina-Zubieta, MD, PhD, FRCP, Mary A. De Vera, PhD: The University of British Columbia and Arthritis Research Canada, Vancouver, British Columbia, Canada; [2]Mark Harrison, PhD: The University of British Columbia, Arthritis Research Canada, and St Paul's Hospital, Vancouver, British Columbia, Canada; [3]Cheryl Barnabe, MD, MSc, FRCPC: Arthritis Research Canada, Vancouver, British Columbia, Canada, and University of Calgary and McCaig Institute for Bone and Joint Health, Calgary, Alberta, Canada; [4]Charlene E. Ronquillo, RN, PhD: The University of British Columbia Okanagan, Kelowna, British Columbia, Canada; [5]Anna Samson, BA: Canadian Arthritis Patient Alliance, Ottawa, Ontario, Canada; [6]Michael Kuluva, BA: Creaky Joints, Vancouver, British Columbia, Canada; [7]Natasha Trehan: University of Ottawa, Ottawa, Ontario, Canada.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25487).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25487.

Address correspondence via email to Mary A. De Vera, PhD, at mdevera@mail.ubc.ca.

Submitted for publication July 5, 2024; accepted in revised form November 27, 2024.

SIGNIFICANCE & INNOVATIONS

- To our knowledge, this is the first study in Canada aiming to understand researcher perspectives on equity, diversity, and inclusion (EDI) in arthritis research and propose mitigation strategies for identified barriers.
- Qualitative findings showed that, although the benefits of using an EDI lens are evident to researchers, barriers remain toward improving representation in Canadian arthritis research, namely, a lack of resources and perceptions from underrepresented communities. Balancing intentionality with feasibility is crucial, and researchers may need to target specific factors (e.g., race, sex, gender, socioeconomic status, etc.) for which representation is important for the research question.
- The need for inclusivity and diversity in research has gone from recognition to action, and Canada is positioned to be a leader in this. To this aim, our findings suggest researchers may benefit from comprehensive guidance, both from rheumatology societies as well as funding agencies, on applying EDI to their own work.

and suboptimal access to interventions, namely, medications and surgery.[5] However, these populations continue to be underrepresented in research. In the growing movement toward improving consideration of equity, diversity, and inclusion (EDI) in research to reduce health disparities, guidance from those who have engaged participants from diverse communities is beneficial for meeting EDI mandates while also looking to favorably affect outcomes. To add to existing efforts, it is important to consider ways to improve representation when conducting research that aims to be generalizable to the population of those living with arthritis in Canada. Therefore, our objective was to explore how arthritis researchers in Canada have used EDI to inform their research.

MATERIALS AND METHODS

Study design. This qualitative study was conducted under a pragmatic paradigm, a philosophy suggesting that knowledge is based in experiences and that what is considered to be the "truth" is what is most useful.[6,7] The perspectives of researchers in rheumatology (both MD and PhD research scientists) related to EDI in research within Canada were solicited via semistructured interviews, an approach that aligns with the pragmatic paradigm, and we followed the consolidated criteria for reporting qualitative research (Supplementary Table 1). This study received institutional ethics approval (BREB# H21-03200).

Eligibility criteria and recruitment. Participants were eligible if they 1) had experience conducting research studies (e.g., clinical, observational, qualitative, health economics, etc.) in arthritis; 2) resided in and/or conducted their research in Canada; and 3) spoke English or French. Participants were identified using a purposive and respondent-driven (snowball) sampling approach with aim of representation across genders, career stage, and province of residence. We began recruitment by asking members of the research team for recommendations and then used snowball sampling to achieve our target sample. Recruitment via email occurred between May and November 2023. All study participants provided informed consent for participation.

Data collection. We conducted and recorded one-on-one semistructured interviews lasting ~60 minutes via video call using the Zoom platform. All interviews were conducted by MMT, a woman of color and graduate student researcher (author positionality statements in Supplementary Table 2). Participants were informed about the study goal and the interviewer's role within the project. The interview guide was developed by all members of the research team to explore the role of EDI in inflammatory arthritis research, with semistructured questions to allow for investigating further topics (Supplementary Table 3). We conducted two pilot interviews with members of the research team to refine the interview guide. Pilot interview findings were incorporated in the final analysis; participants were aware their interviews could be included as part of the study, and informed consent was obtained for this purpose. We interviewed participants until data saturation (new data repeat previous expressions in data) and inductive thematic saturation (no identification of new codes or themes) were achieved.[8] All data were handled by MMT and MADV. The interviews were audio-recorded, transcribed verbatim, and coded for overall analysis. Raw transcripts were created using Sonix, an online transcription service (https://sonix.ai), then reviewed for accuracy by MMT.

Analysis. A codebook approach to thematic analysis was used, specifically template analysis, which allows for deductive and inductive coding and is aligned with the pragmatic paradigm.[9] The data were analyzed by MADV, MH, and MMT using a combination of deductive and inductive coding, organized with NVivo software. Inductive coding began with identifying and collating raw codes (preliminary themes using participants' exact words) derived directly from the data to develop a single coding scheme.[10] Through classifying, constant comparison, and memo writing, we transformed codes into subthemes (middle-order themes to transform basic themes into similar concepts) and eventually themes to capture global patterns in the data. A combination of latent and semantic coding was used to capture surface-level meanings as well as underlying assumptions of what was being expressed by participants. To ensure dependability of the coding after the generation of initial themes, MADV, MH, and MMT discussed the initial coding framework over several meetings. Once satisfied with the coding framework, MMT coded all transcripts using strategies to ensure trustworthiness, including

reflexive journaling, and discussing themes with the research team.[11] After analyses, a narrative summary of the main findings was developed.

RESULTS

We interviewed 22 arthritis researchers (45.5% women and 54.5% men) in Canada. Demographic details of participants are in Table 1. For each of the three themes, reflective quotations highlighted benefits, barriers, and facilitators for prioritizing EDI in rheumatology research. Additional supporting quotations can be found in Table 2.

Theme 1: benefits of prioritizing EDI in rheumatology research. All participants agreed that prioritizing EDI in rheumatology research is beneficial. From their sentiments, two subthemes were constructed.

Increasing knowledge to improve generalizability and quality of care. Researchers highlighted that prioritizing EDI could provide us with more information on how arthritis affects patients differently to provide better care.

Table 1. Demographic characteristics of participants

Demographic information	n (%), n = 22
Province of residence	
Alberta	3 (13.6)
British Columbia	10 (45.5)
Manitoba	1 (4.5)
Nova Scotia	1(4.5)
Ontario	3 (13.6)
Quebec	4 (18.2)
Gender	
Male	12 (54.5)
Female	10 (45.5)
Current employment	
Clinician scientist	15 (68.2)
Researcher	7 (31.8)
Career stage	
Early (<5 y)	2 (9.1)
Middle (5–15 y)	8 (36.4)
Late (>15 y)	12 (54.5)
Type of inflammatory arthritis studied (select all that apply)	20 (90.9)
Rheumatoid arthritis	2 (9.1)
Juvenile idiopathic arthritis	2 (9.1)
Ankylosing spondylitis	5 (22.7)
Systemic lupus erythematosus	6 (27.3)
Psoriatic arthritis	5 (22.7)
Gout	3 (13.6)
Other	6 (2.73)
Type of research primarily conducted (select all that apply)	
Experimental studies/clinical trials	10 (45.5)
Observational studies	21 (95.5)
Qualitative research	5 (22.7)
Health economics research	6 (27.3)
Other	2 (9.1)

> "I think overall it's probably very underappreciated how important these factors are […], in patients with rheumatoid arthritis or psoriatic arthritis or ankylosing spondylitis who are of a non-[White] background. It's very difficult for us to advise. Similarly for patients who are pregnant, these people are actively excluded from studies."—*Participant 20*

Having more representation among study participants was perceived to improve the opportunity to find answers on how and why patient outcomes differ, as well as factors that limit access to care for patients.

> "[A benefit of including EDI in research is] getting answers to why some people don't do as well as others. We have multiple examples, not only in arthritis, that people that have poor access, people with certain type of barriers to their health care don't do well and some of these are not very well understood."—*Participant 10*

Furthermore, by having more representative study samples, participants felt that the generalizability and clinical impact of research could be improved.

> "If we don't include [EDI], we can't generalize our findings because our findings are only generalizable to the people who are included in the research."—*Participant 17*

Overall, participants agreed on the value of considering EDI in arthritis research, highlighting that it goes beyond an "ethical responsibility" and improves our ability to provide better care, with one participant stating that "it's just good science."

Reducing health disparities. Many participants spoke to the ultimate goal of research being to reduce health disparities, i.e., unfair differences in health outcomes, and highlighted this as a key benefit of considering EDI.

> "I think really the goal of all of our research should be to reduce health disparities. […] the benefit of considering [EDI] and including it in research is actually trying to address those disparities and at least make sure that we're not making these things worse."—*Participant 1*

Table 2. Additional quotations supporting thematic analysis*

Global theme	Subtheme	Quotations
Benefits of prioritizing EDI in rheumatology research	Increased knowledge to improve generalizability and quality of care	"*I think if we can look at a disease through the lens of the interactions with racial background or interactions with sex, we start to learn more about the disease and how it functions. And I think that will lead to better outcomes and better interventions for everybody if we have a better understanding of it.*"—Participant 20
	Reducing health disparities	"*Because if you're going to create clinical insights and recommendations or guidelines and those guidelines are only relevant to 50% of the population, then there's a whole subset that aren't being included and that can lead to obviously the exclusion is huge.*"—Participant 2
		"*If we're only including the dominant group in our research, only doing research that benefits the dominant group and only disseminate research to the dominant group, it's one of the system-level inequities and discrimination in our society because we're not actually doing research that benefits everybody in society.*"—Participant 14
Barriers to prioritizing EDI in rheumatology research	Mistrust from historically exploited populations	"*I have to approach every patient in my clinical care, assuming they've had prior bad experiences. And I have some making up to do for experiences that the patients have had with other health care providers until we build that relationship of trust.*"—Participant 14
		"*Yes, we want to be inclusive, but we also have to be respectful that different people have different backgrounds, that for them, research may not necessarily have the same meaning as it does to us.*"—Participant 18
	Lack of access to research opportunities	"*I have to constantly remind a lot of people, my staff, my colleagues, that people participating in research are doing us the favor, because it often doesn't feel that way [...] You know, we're doing the interviews in French and English between 9 and 5 because that's what's easiest for us.*"—Participant 12
		"*So I think the other thing too, again as I mentioned, because we work in this multiethnic area, many of these people of these different groups have lower socioeconomic status. And to them, research is not meaningful. They have to get on with their job [...] they don't have this luxury of participating in research.*"—Participant 18
	Logistical challenges	"*It always comes back to more resources [...] And if you're you have limited resources, you're going to take the, you know, more low-lying fruit.*"—Participant 18
	Unintended consequences	"*I totally can appreciate the concern that if we modify criteria or change recruitment practices to try and include more vulnerable populations are we actually then almost like exploiting those populations for the benefit of others, of more privileged individuals?*"—Participant 1
		"*I think especially in like clinical care settings or even in teams that aren't used to taking a health equity perspective, if you don't know how to collect and analyze certain types of data, then you could do it wrong and end up having an unanticipated consequence on the findings.*"—Participant 2
Facilitators to support EDI prioritization in rheumatology research	Building community partnerships	"*So, we needed to always have champions. So when we go into the Black community, we would partner with leaders in the community and we would talk to them about how to prepare our materials, how to hold informational sessions. They would help us to understand what would be important to their constituents, what kind of reassurances they would be looking for. And just, you know how to create a partnership, a meaningful partnership.*"—Participant 12
	Curating diverse research teams	"*If you have a team that's diverse and it shows maybe more openness, it may be easier to have people agree to participate in your in your work.*"—Participant 10
	Incentivizing researchers and funder support	"*We're not going to see the degree of change that we want to see unless there's a sort of a mandatory aspect to some of these things.*"—Participant 1
		"*[We need] better understanding from funders that we don't develop trust overnight with communities. And that we need some time and some resources to be able to do it.*"—Participant 12
	Fostering humility in research environments	"*But at the same time, I can also appreciate how some people would really be hesitant and worried about making mistakes in this realm again, because of the current climate of cancel culture that we that we live in. And I can see how [...] I guess, just maybe not feeling like it's a safe space to make mistakes in*"—Participant 1
		"*If you get it wrong, then you have the chance to correct it. If you don't do it, then we are going to be in the dark all this time.*"—Participant 6

* EDI, equity, diversity, and inclusion.

As well, participants expressed that, by using an EDI lens to design research, there is an opportunity to have an impact on factors we are able to change as opposed to focusing on genetic factors alone, which is critical for chronic conditions such as arthritis.

> "You could have a study design, addressing that specifically and getting answers on some of these things that are modifiable, right? You may not be able to change your genes, but you might be able to change your socioeconomic factors or write, report or find some interventions, right?"—*Participant 10*

Theme 2: barriers to prioritizing EDI in rheumatology research. From the experiences of researchers, four barriers were identified as affecting current approaches to prioritizing EDI in Canadian rheumatology research. These barriers are described below and include: Mistrust from historically exploited populations, Lack of access to research opportunities, Logistical challenges, and Unintended consequences

Mistrust from historically exploited populations. Harmful stereotypes, racism, and exclusion in health care and research settings have led to mistrust about the intentions of researchers and care providers within certain communities. Participants specifically highlighted the need to consider underrepresented populations when considering this barrier. This mistrust was also perceived to lead to conflicting priorities between patient participants and researchers/providers because research may have negative connotations for many communities.

> "So, for instance, like the Black community, it's been documented that they distrust the medical community, especially if the […] clinical provider is not in their community. I think that creates a lot of mistrust and that stems from historical kind of roots."—*Participant 2*

Particularly when considering collecting information from participants, mistrust should be addressed by providing rationale for why this information is required or important for researchers to know.

> "I know in my culture there can be a little bit of distrust when it comes to others trying to gather information […] It's like, 'why do you want this information?'"—*Participant 16*

Participants also spoke to the need for researchers and providers to consider how unconscious biases may be influencing their own work and focus on rebuilding trust with their patients in safe, trauma-informed environments.

> "One key way is to just challenge whenever I have assumptions because usually that's where your unconscious biases creep out. So never assume that a patient can't afford a treatment […] Don't assume that they don't want drugs or are drug-averse because they're from a group that traditionally has been labeled as not liking drugs."—*Participant 14*

Lack of access to research opportunities. When speaking about lack of access, participants noted this referred to not only who has the opportunity to participate in research, specifically considering participant burden, but also who hears about the opportunity to participate in research.

> "Yeah, well, we think we include everybody, but we don't acknowledge that for some certain part of the population, it's harder to be part of the study, to be part of research."—*Participant 3*

With respect to participant burden, researchers spoke to the challenge of being able to appropriately compensate patients for the time taken to participate in the study, as well as additional financial and nonfinancial costs (such as child care, transportation, taking time off work to participate, living in rural or remote settings, etc.).

> "If you think about a clinical trial, like a traditional pharmaceutical company sponsored clinical trial, they typically require a lot of monitoring, you have to go to and you have to show up in-person for a visit every 2 to 4 weeks, and it's quite intensive. So, if you live in a remote place or you work in a job which doesn't allow you the ability to leave the job to go to these visits, you might not get access to whatever medication is offered in that trial, right?"—*Participant 17*

Logistical challenges. All participants spoke to logistical challenges for researchers to accommodate participant needs,

including sufficient funding, effort, time, and resources, as a key barrier for their own ability to consider EDI in research. Actively aiming toward having better representation in research was perceived to be expensive, in both tangible and intangible ways.

> "Doing well is expensive. You know, expensive not just for money, but expensive in terms of time and energy."—*Participant 8*

Language was highlighted as a key logistical challenge preventing many patients from participating in research, and participants spoke to how available translation resources are not always comprehensive enough to address this.

> "Speaking English is often a criterion for many research studies. So, I think people […] maybe don't feel welcomed or interested in participating."—*Participant 7*

Furthermore, careful consideration is needed for how researchers are collecting and analyzing information, both in terms of rigor and ethical ramifications. Data collection requires considering how researchers are able to balance potentially competing objectives of recruiting more difficult-to-reach patient participants (e.g., data ownership, data protection, etc.) with concerns about achieving statistical power and being able to categorize and present data. Related to this, participants spoke to how to collect data appropriately if they perceive having a lack of expertise in EDI.

> "How do you collect that data? So, you know, if you don't work with a sex and gender or an EDI specialist, what are the questions?"—*Participant 18*

There is also a need to consider what analyses are possible when subgroups are small.

> "If you don't have enough representation in your sample, then you can't really do these analyses and these subgroups because your samples are too small."—*Participant 14*

Unintended consequences. The final barrier participants identified was a fear of unintended consequences that could worsen health disparities. This stemmed from the issue of "*drive-by research*," wherein researchers conduct studies in underserved communities, uncover problems, and fail to solve them. Specifically, researchers spoke to not feeling knowledgeable enough to work with underrepresented communities, despite having the intention.

> "I don't feel I'm knowledgeable enough to be able to conduct Indigenous research. So […] some researchers are just afraid that they do something wrong […] they don't know the culture."—*Participant 13*

Several participants spoke to the need to consider the nuance of having an absence of data on the recommendations made for clinical practice in certain populations, as this could lead to incorrectly interpreting data.

> "It would be great if we did have more evidence in these populations, but we shouldn't assume that there's going to be biologic differences in how the drugs work unless we have evidence to support that. […] I mean, it's almost like are you worsening inequities by making that assumption?"—*Participant 4*

However, other participants highlighted that the fear of unintended consequences should not be used as a reason to not consider EDI in research.

> "I think that might be a convenient excuse. I mean, I think if you approach it with a genuine spirit and you want to learn and you're open and you say, 'Look. Help me understand.'"—*Participant 12*

The overarching concern was how to balance increasing diversity in study samples while ensuring that one or a small number of patients are not being used to represent all patients with similar patient profiles (e.g., two patients of the same race and gender identity can have different health outcomes, preferences, life experiences, etc.).

> "I do think it's important that we recognize that we have diversity in our populations, but I actually also think it's a bit dangerous because people think the

voice of one person is enough for the entire group that they represent."—*Participant 17*

Although generalizability was viewed as a benefit overall, participants felt it was important to be intentional about who is being recruited and transparent about the evidence gained from research.

Theme 3: facilitators to support EDI prioritization in rheumatology research. To address existing barriers, researchers highlighted four facilitators to be leveraged when prioritizing EDI in Canadian rheumatology research.

Building community partnerships. The first facilitator identified by participants was the need to develop community partnerships, particularly having a "champion" or organization to help build relationships and bridge the gap between researchers and underrepresented communities. By focusing on building these relationships, it was perceived that researchers could establish trust with more difficult-to-reach populations and address the aforementioned barrier of mistrust from historically exploited groups.

"What we learned was that our champions were so much more powerful than anything that we could say or do, right? We had to earn their trust."—*Participant 12*

Participants also acknowledged that building these partnerships will take time, which can be a challenge for researchers but is a necessary step that could allow for more careful consideration of the way research is conducted.

"There's greater acknowledgment that things just take time, and we need to put more effort into community building and being part of the community a bit more, which is good because it forces researchers to take more thoughtful approaches"—*Participant 2*

In order to successfully build community partnerships, working with patients and/or advisory boards was perceived to be critical.

"[…] a big part of it is actually to talk to the people that we want to incorporate in that and understand how we can support them to include them better."—*Participant 3*

Curating diverse research teams. Participants identified that using an EDI lens in research should be integrated throughout all phases of studies (i.e., design, recruitment, data collection analysis, and knowledge translation) but that it begins with considering who is conducting the research. It was perceived that having more diverse research teams could encourage better representation in studies.

"It starts with not even the data collection, but who we're including in our research teams, right? So, do we have diverse individuals who are on the staff for informing the design of measures, design of the study, design of the grants? Do we have diverse patients or consumers who are involved in providing insights into lived experience?"—*Participant 2*

Participants felt that having diverse perspectives on research teams allows for better consideration of how to recruit patients from different communities, as well as identifying potential barriers to participation and solutions.

"So more diverse staff in our environment who can just alert us to things that we might be doing inadvertently that would discourage others—may discourage people from participating."—*Participant 12*

Incentivizing researchers and funder support. Because many of the logistical challenges identified were financially driven (i.e., requiring more time and resources to adequately address patient needs to participate in research), participants tended to consider financial support as critical for considering EDI in rheumatology research in general, as well as in their own studies. Participants spoke to the role of external decision-makers in improving consideration of EDI in research, specifically the need for recognition from funding organizations and regulatory agencies.

"If we can get funders to also appreciate that these principles mean that we need to be funding at higher levels, [and that] we need to be resourcing teams to be able to do this and expecting teams to be able to do it. It's one thing to expect research teams to incorporate EDI principles. It's another matter to actually resource it and fund it."—*Participant 20*

Furthermore, participants spoke to the need to incentivize researchers to better consider EDI in their work,

but that it may need to begin as a mandated approach to be effective.

Fostering humility in research environments. Participants spoke both directly and indirectly about the need to create safe spaces in research environments, which would require a sense of humility from researchers. Some participants highlighted how the language around EDI could be distracting from the importance, as many researchers value the importance but tend to feel challenges in trying keep up with changing terminology.

> "The first place my head goes is like, 'Oh my God, am I addressing them by the right pronouns?'"—*Participant 16*

However, the overarching consensus was the need to acknowledge research limitations with a mindset of being transparent and improving over time.

> "I will still say, it's better to try and get it wrong and then admit that you're wrong and that you didn't have the right people […] I think that's the matter of science, is that you try something, it doesn't work. Then we try it in a different way. We never advance science by not looking into it."—*Participant 6*

In order to combat the "current climate of cancel culture that we live in" (Participant 1), participants emphasized needing to balance between the importance of considering EDI and their ability to do so, which could require additional educational opportunities and training. Ultimately, the ability of researchers to be open and willing to learn and to create safe spaces to conduct research were perceived as key to being able to consider EDI in participants' research.

Thematic synthesis. From the interview data, we constructed a figure (Figure 1) that highlights how the barriers (Theme 2) participants identified corresponded to their suggested facilitators (Theme 3). Horizontal arrows depict the direct connections between barriers and facilitators, whereas the dashed lines indicate potential linkages between barriers and facilitators. Barriers and facilitators that are connected to one another are indicated by brackets. The barrier of mistrust from exploited communities corresponds to the need to rebuild trust and create safer research environments for patient participants, particularly from underrepresented communities, and this often requires including partners that communities can trust. The lack of access to research opportunities for underrepresented patients could be improved by having diverse research teams, which bring richer perspectives on how different patient populations may respond to the opportunity to participate in research (including recruitment strategies, patient barriers to participation, understanding cultural differences, etc.). Many of the logistical challenges mentioned by participants could primarily be resolved by having more resources to conduct research with EDI considerations, particularly around appropriate compensation for patient participants, but also for issues such as language barriers. Finally, the fear of unintended consequences that could potentially increase health disparities is linked to the need to foster humility in research environments and create safe spaces where researchers in Canada can consider EDI and be transparent about the limitations and be open to learning from others with different expertise.

DISCUSSION

This qualitative study was conducted to explore how researchers in rheumatology have applied EDI principles to inform

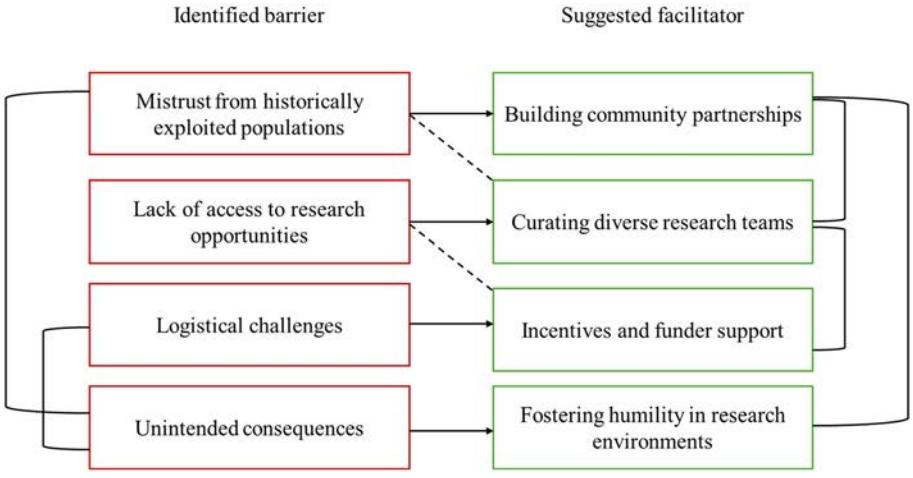

Figure 1. Identified barriers and corresponding facilitators for mitigation.

their research. Participants agreed that EDI is important to consider, but there is a need to balance intentionality with feasibility. Furthermore, to have the most impact, it is critical to consider EDI early and at all stages of the research process, from study team composition through to study design, recruitment, analysis, and knowledge translation.

The challenges of including underrepresented populations in research have been recognized in many different fields and specialties.[12–14] Known barriers to participation for underrepresented populations were aligned with our findings, such as lack of awareness about research opportunities, mistrust, lack of diversity among the research team leading to patients not feeling represented, research not being conducted in the community, and logistical issues.[14] Existing strategies to address these barriers were similar to suggestions from our study participants, including tailoring recruitment approaches to different communities (e.g., media campaigns, recruiting in places of worship, etc.), reducing/limiting indirect and direct costs to participation, transparency and education to improve awareness, translating research materials into multiple languages, and recruiting diverse research team members.[12–14]

Although several important barriers and facilitators were discussed by participants in this study, some relevant concepts were not explored in-depth. One barrier that was not explicitly discussed by participants in this study was the perception of funders and peers toward individual researcher efforts to promote representativeness in their work. This barrier has been recognized as hindering EDI prioritization in other fields and speaks to a larger issue for the research community to conducting research that will be funded versus conducting research that can create meaningful change.[15] The need for funding organizations to have rigid methods for proposed projects limits the ability to provide the flexibility and adaptability necessary for EDI prioritization in research. Addressing this barrier may involve increasing diversity in power positions such as members of review panels for different funding organizations.

Additionally, the role of other interested parties beyond the patient and provider (i.e., caregivers, family members and friends, peers, community advocates, religious and political leaders, etc.) were not discussed, which may suggest a need to expand researchers' scope of who needs to be involved when considering how to recruit and retain patient participants in studies. Using flexible, person-centered strategies may help mitigate barriers for participation and are important for improving representation in research teams, as well as in study participants.

Based on our findings and existing research,[12–14,16] comprehensive guidance for researchers is needed for ways to feasibly consider representation in their own work. It is important to note that it is likely not possible to have representation across all SDOH (e.g., sex, gender, age, race, socioeconomic status, etc.), so researchers may need to target specific factors for which representation is key for the research question. Intentionality is critical when using approaches to improving representation and should involve providing a rationale to participants explaining why particular data need to be collected. For example, if researchers are aiming to achieve a better range of racial and ethnic representation in their study, they should consider how to focus on these factors through each stage of the study. This may require having a more racially and ethnically diverse research team to provide insights into study design, providing translation services for participants who do not speak English, and targeted recruitment strategies to reach specific communities, among other considerations. These findings may lead to discussions with funding agencies (e.g., Canadian Institutes of Health Research) to help develop guidance for researchers in Canada. A possible output from future research is the development of modules to support and provide education on EDI in research to individuals wishing to apply to federal funding sources. Future research should also involve working with patients to better understand how to address barriers to their participation in research.

A strength of this research is the contribution of patient partners, rheumatologists, and researchers to the overall study design and interpretation of results. Following guidance on thematic analysis, the researchers most immersed in the interview content led the analysis, and code development was discussed and verified with others.[17] However, there was overrepresentation of participants from British Columbia, which may limit the applicability of findings to other provinces and/or territories with limited representation. Additionally, people who agreed to participate in the study may have been better informed about or more interested in discussing EDI than those who did not respond or those who declined to participate (n = 4) and may have been more likely to have experience considering EDI in their own research. Finally, we may not have captured the range of arthritis research in our recruitment, as we did not have basic scientists, translational researchers, or other related researchers as participants.

Altogether, our findings highlight the perceived value of EDI in Canadian arthritis research to improve health disparities and our ability to provide comprehensive, high-quality care for people living with arthritis. Considering EDI should be a priority not because of a societal responsibility to improve health disparities through research, but because "it's just good science."

ACKNOWLEDGMENTS

The authors would like to thank the participants of this study for their time and candor.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr De Vera confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the

affirmations regarding article submission (e.g., not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Barnabe C. Disparities in rheumatoid arthritis care and health service solutions to equity. Rheum Dis Clin North Am 2020;46(4):685–692.
2. Gheshlaghi N, Thomas M, Trehan N, et al. Reporting of determinants of health inequities and participant characteristics in randomized controlled trials of juvenile idiopathic arthritis in Canada: a scoping review. Pediatr Rheumatol Online J 2023;21(1):134.
3. Thomas M, Harrison M, De Vera MA. Reporting of determinants of health inequities in rheumatoid arthritis randomized controlled trials in Canada: a scoping review. Arthritis Care Res (Hoboken) 2023;75(1):101–114.
4. Thomas M, Verma V, Gheshlaghi N, et al. Reporting of determinants of health inequities and participant characteristics in randomized controlled trials of systemic lupus erythematosus in Canada: a scoping review. Lupus 2024;33(5):462–469.
5. Barnabe C. Promoting equity in rheumatoid arthritis through treatment guidelines: implementation guides for six populations at risk for inequitable outcomes in Canada. Canadian Institutes of Health Research: Musculoskeletal Health and Arthritis. 2020. Accessed [June 16, 2024]. https://cihr-irsc.gc.ca/e/51798.html
6. Kaushik V, Walsh CA. Pragmatism as a research paradigm and its implications for social work research. Soc Sci 2019;8(9):255.
7. Koehn DR. C. S. Peirce's "Illustrations of the Logic of Science" and the pragmatic justification of induction. Trans Charles S Peirce Soc 1973;9(3):157–174.
8. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893–1907.
9. King N, Brooks J. Thematic analysis in organizational research. In: Cassell C, Cunliffe AL, Grandy G, eds. The SAGE Handbook of Qualitative Business and Management Research Methods: Volume 2: Methods and Challenges. Sage Publications Ltd.; 2018.
10. Attride-Stirling J. Thematic networks: an analytic tool for qualitative research. Qual Res 2001;1(3):385–405.
11. Krefting L. Rigor in qualitative research: the assessment of trustworthiness. Am J Occup Ther 1991;45(3):214–222.
12. Lai GY, Gary TL, Tilburt J, et al. Effectiveness of strategies to recruit underrepresented populations into cancer clinical trials. Clin Trials 2006;3(2):133–41.
13. Waheed W, Hughes-Morley A, Woodham A, et al. Overcoming barriers to recruiting ethnic minorities to mental health research: a typology of recruitment strategies. BMC Psychiatry 2015;15:101.
14. Haley SJ, Southwick LE, Parikh NS, et al. Barriers and strategies for recruitment of racial and ethnic minorities: perspectives from neurological clinical research coordinators. J Racial Ethn Health Disparities 2017;4(6):1225–1236.
15. Bibbins-Domingo K, Helman A, eds. Improving Representation in Clinical Trials and Research: Building Research Equity for Women and Underrepresented Groups. The National Academies Press; 2022.
16. Sheikh SZ, Wanty NI, Stephens J, et al. The state of lupus clinical trials: minority participation needed. J Clin Med 17 2019;8(8):1245.
17. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77–101.

Arthritis Care & Research
Vol. 77, No. 6, June 2025, pp 811

American College of Rheumatology
Empowering Rheumatology Professionals

LETTER

DOI 10.1002/acr.25484

Risk of mortality after hip fractures in rheumatoid arthritis: comment on the article by Jones et al

To the Editor:

We read with great interest the article by Jones et al[1] investigating the incidence and mortality risk associated with hip fractures in patients with rheumatoid arthritis (RA). This study provides important insights into the epidemiology of RA-related fractures and highlights how advancements in RA management may have contributed to improved outcomes. However, as a health care professional who also lives with RA, I feel compelled to expand upon key findings and their implications for both clinical practice and health policy.

Jones et al reported that patients with RA had a 28% higher risk of hip fractures than age- and sex-matched controls, yet their postfracture mortality risk was comparable to the general population. These findings underscore the effectiveness of modern RA therapies in mitigating inflammation and managing osteoporosis. For instance, Pawar et al[2] demonstrated that biologic and targeted synthetic disease-modifying antirheumatic drugs do not significantly differ in the risk of nonvertebral fractures, providing reassurance to physicians regarding their safety profiles. However, the persistent fracture risk identified in the cohort of Jones et al, despite these therapeutic advances, underscores the necessity of addressing factors such as falls and comprehensive osteoporosis management.

A key aspect that deserves further exploration is the role of fall prevention in RA populations. As Brenton-Rule et al[3] noted, up to 50% of patients with RA experience falls within a year, significantly heightening their fracture risk. Multifaceted interventions, including home modifications and balance training, could complement pharmacologic strategies to reduce the overall fracture burden.

Although Jones et al observed no excess mortality risk post–hip fracture in patients with RA compared with controls, this finding diverges from data in Asian populations, where Lin et al[4] and Kwon et al[5] found higher mortality rates in RA cohorts following hip fractures. Such discrepancies might reflect geographic, ethnic, or health care system variations, as previously suggested by Kanis et al.[6] These findings warrant further multinational studies to elucidate contextual differences and inform tailored interventions.

Finally, the study by Jones et al highlights the critical importance of addressing osteoporosis-related bone loss in patients with RA receiving glucocorticoids. Saag et al[7] demonstrated that discontinuing denosumab in this population leads to a rebound in bone turnover and a return to baseline bone mineral density, emphasizing the need for consistent and long-term osteoporosis management strategies.

In conclusion, Jones et al have provided robust evidence of the persistent fracture risk in RA populations and the necessity for integrated management approaches. Future studies should aim to incorporate fall prevention strategies, explore regional disparities in outcomes, and assess the long-term effectiveness of novel therapeutics in mitigating fracture risk. These efforts will help reduce the substantial burden of fragility fractures in RA and improve patient quality of life.

As Taiwanese is my native language, I used Microsoft Copilot from Office 365 for language editing and proofreading to help minimize linguistic errors. I reviewed and approved the final content, taking full responsibility for its publication.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25484.

Jui-Man Chang, PMHNP
Lien-Chung Wei, MD, MPH
wlc078@berkeley.edu
Taoyuan Psychiatric Center
Taoyuan City, Taiwan

1. Jones CA, Guy P, Xie H, et al. Incidence of and risk of mortality after hip fractures in rheumatoid arthritis relative to the general population. Arthritis Care Res (Hoboken) 2025;77(5):604–613.
2. Pawar A, Desai RJ, He M, et al. Comparative risk of nonvertebral fractures among patients with rheumatoid arthritis treated with biologic or targeted synthetic disease-modifying antirheumatic drugs. ACR Open Rheumatol 2021;3(8):531–539.
3. Brenton-Rule A, Dalbeth N, Bassett S, et al. The incidence and risk factors for falls in adults with rheumatoid arthritis: a systematic review. Semin Arthritis Rheum 2015;44(4):389–398.
4. Lin YC, Li YH, Chang CH, et al. Rheumatoid arthritis patients with hip fracture: a nationwide study. Osteoporosis Int 2015;26(2):811–817.
5. Kwon HY, Kim HH, Sung YK, et al. Incidence and mortality of osteoporotic fracture in rheumatoid arthritis in South Korea using nationwide claims data. J Bone Metab 2019;26(2):97–104.
6. Kanis JA, Oden A, McCloskey EV, et al. A systematic review of hip fracture incidence and probability of fracture worldwide. Osteoporosis Int 2012;23(9):2239–2256.
7. Saag KG, McDermott MT, Adachi J, et al. The effect of discontinuing denosumab in patients with rheumatoid arthritis treated with glucocorticoids. Arthritis Rheumatol 2022;74(4):604–611.