Arthritis Care & Research

Aims and Scope

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

Volume 73, No. 7, July 2021

Arthritis Care & Research

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

Cover design: Sandra Pulmano

This journal is printed on acid-free paper.

Arthritis Care & Research

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

VOLUME 73 • July 2021 • NO. 7

Special Articles

Pediatrics

Rheumatoid Arthritis

COVID-19

Osteoarthritis

Systemic Lupus Erythematosus

Ehlers-Danlos Syndrome

Gout

Letters

Cover image: The image on the cover (from Buldt et al, page 1023) shows Dorsiplantar radiographic measurement techniques for bone length, width and angles.

ARP Announcements

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

Download the New ACR Publications Mobile App

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

ARP Membership

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

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

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

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 919–920
DOI 10.1002/acr.24714

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

Arthritis Care & Research: Celebrations and Opportunities

Kelli D. Allen,[1] S. Sam Lim,[2] and Todd A. Schwartz[3]

We are deeply honored to begin our service as editors of *Arthritis Care & Research* (AC&R), the official journal of the Association of Rheumatology Professionals (ARP) and an essential tool in the mission of the American College of Rheumatology (ACR) "to empower rheumatology professionals to excel in their specialty." We each have a long history of involvement with the ACR/ARP, and these experiences have significantly influenced our professional lives. We are passionate about furthering the mission of the ACR/ARP, and we are particularly enthusiastic to contribute to the continued growth, reach, and impact of AC&R.

We would like to begin this editorial by celebrating AC&R and those whose work has allowed it to thrive. AC&R has been led by a series of outstanding editorial teams, most recently Drs. Marian Hannan and Leslie Crofford, along with their excellent team of Associate Editors (1), Managing Editor Maggie Parry, and Assistant Editor Margaret Graton. The knowledge, skills, and efforts of each of these individuals have truly been essential pieces in the success of the journal, and we give our sincere thanks. We extend special gratitude to Dr. Hannan, who for 10 years has led AC&R with excellence and devotion to the ACR/ARP. We thank her for being so generous and supportive during the transition.

As we embark on our AC&R editorial term, we are privileged to lead a journal that is incredibly strong and poised for future growth and innovation. AC&R has been on a growth trajectory for years, based on a range of metrics, including the number of subscribers and submissions, international representation, Impact Factor, and other measures of journal influence (2,3). However, the greatest strength of AC&R is the importance of its underlying mission: to promote excellence in the practice of rheumatology, through an interdisciplinary lens. This mission sets it apart from other journals in the field, guides article selection, creates themed issues and initiatives, and leads to real-world impacts. Pursuit of this mission will continue to drive the activities and success of AC&R.

As an editorial team, we are committed to the core activities that promote a growing and impactful journal. These include attracting and retaining authors and readers, seeking to publish the highest-quality studies within the journal's scope, providing quality manuscript reviews, and promoting visibility of key study findings, particularly in the context of a changing media landscape. In addition, there are 3 specific areas of focus for our incoming editorial team that we believe are timely and closely aligned with the mission of the ACR/ARP. First, we will explore additional avenues for connecting with front-line rheumatology care and providers. This is already a strength of AC&R, and our multidisciplinary editorial team will consider ways in which we can further enhance the journal's impact in the evidence-to-practice space, through manuscripts and other products.

Second, we will seek to publish and promote the use of emerging, innovative research methods in clinical rheumatology

[1]Kelli D. Allen, PhD: University of North Carolina, Chapel Hill, and Department of Veterans Affairs Health Care System, Durham, North Carolina; [2]S. Sam Lim, MD, MPH: Emory University, Atlanta, Georgia; [3]Todd A. Schwartz, PhD: University of North Carolina, Chapel Hill.

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

Address correspondence to Kelli D. Allen, PhD, Thurston Arthritis Research Center, 3300 Thurston Bldg., CB# 7280, Chapel Hill, NC 27599. Email: kdallen@email.unc.edu.

Submitted for publication May 18, 2021; accepted in revised form May 20, 2021.

research. AC&R has an opportunity to showcase innovative statistical and study design methodologies particularly applied to issues of clinical importance to readers. Our editorial team will pursue opportunities to highlight these methods, thereby contributing to both the research and educational missions of the ACR/ARP.

Third, our team has a strong interest in health disparities and equity, areas in which AC&R already has a strong track record. Given the importance of these topics globally, we will seek to publish a variety of publications that add meaningfully to this critical issue facing rheumatology and other health professions. AC&R is positioned to disseminate new findings and perspectives regarding health disparities and equity from an international and multidisciplinary perspective, and our team will seek to maximize the journal's impact and reach in this area.

In closing, we are grateful for the opportunity to work alongside the incoming group of outstanding Associate Editors: Joshua Baker, MD, MSCE, Nancy Baker, ScD, MPH, OT, Cheryl C. M. Barnabe, MD, MSc, Bonnie Bermas, MD, Lorinda Chung, MD, MS, Maria I. Danila, MD, MPH, Robert F. DeVellis, PhD, Afton L. Hassett, PsyD, Puja P. Khanna, MD, MPH, Kanta Kumar, PhD, Crystal MacKay, PhD, MHSc, BScPT, Natalie McCormick, PhD, Kaleb Michaud, PhD, Eli Miloslavsky, MD, Michael H. Weisman, MD, Pamela F. Weiss, MD, MSCE, and Daniel K. White, PT, ScD, MSc. We would like to note that Drs. Bermas, DeVellis, Michaud, and White are continuing their service from the prior editorial team, providing continuity and relevant experience. These Associate Editors represent a variety of clinical and methodologic specialties, which is critical for achieving high quality manuscript reviews, as well as for overall leadership of the journal. Finally, we thank AC&R readers, authors, and reviewers for their continued to support. We look forward to working with this community to ensure the ongoing success of AC&R.

REFERENCES

1. Hannan MT, Crofford LJ. Arthritis Care & Research: a look back and a view forward [editorial]. Arthritis Care Res (Hoboken) 2021;73:765–6.
2. Hannan MT. A quarter-century of excellence and still growing [editorial]. Arthritis Care Res (Hoboken) 2013;65:1–3.
3. Hannan MT. Arthritis Care & Research: continued success and evolution [editorial]. Arthritis Care Res (Hoboken) 2011;63:925–6.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 921–923
DOI 10.1002/acr.24536

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

Evolution in the Understanding of Pediatric-Onset Axial Spondyloarthritis

Daniel J. Lovell and Hermine I. Brunner

Over the last decade, there has been substantial evolution of the concept of axial spondyloarthritis (SpA) with the recognition of nonradiographic SpA in adult patients. Axial SpA is a chronic inflammatory rheumatic disease that primarily affects the axial skeleton, but patients with the illness also frequently demonstrate peripheral arthritis and enthesitis (1). In the US, the estimated prevalence of axial SpA is 0.9–1.4% (2,3). In the current concept of this disease, axial SpA in adults includes both those with radiographic evidence of structural damage of the sacroiliac (SI) joints or lumbar spine (ankylosing spondylitis [AS]) and those without radiographic damage of the SI joints or spine (nonradiographic SpA). The Assessment of Spondyloarthritis international Society (ASAS) developed classification criteria for nonradiographic SpA, for use in adults, that require either magnetic resonance imaging (MRI) evidence of SI inflammation or the presence of HLA–B27 plus clinical features (4). Recently, treatment guidelines for both AS and nonradiographic SpA in adults have been developed. In 2015, a collaborative between the American College of Rheumatology, the Spondylitis Association of America, and the Spondyloarthritis Research and Treatment Network published treatment guidelines for both the AS and nonradiographic axial SpA groups. It is recommended, for both groups, to use nonsteroidal antiinflammatory drugs (NSAIDs) as well as anti–tumor necrosis factor (anti-TNF) inhibitors for active arthritis that fails to respond to NSAIDs (5).

Juvenile idiopathic arthritis (JIA) is the most common rheumatic disease in children. In the currently used diagnostic criteria from the International League of Associations for Rheumatology (ILAR), JIA is divided into 7 distinct categories (6). Enthesitis-related arthritis (ERA) is the category used to describe the pediatric-onset equivalent of SpA in adults. As per the ILAR criteria, a classification of ERA requires the presence of arthritis plus enthesitis. Another classification of ERA requires arthritis or enthesitis and ≥2 of the following: SI joint tenderness or inflammatory back pain, HLA–B27 positivity, first-degree relative with disease associated with HLA–B27, acute anterior uveitis, and arthritis in a male individual older than 6 years (6). Thus, although the criteria for ERA do not require axial arthritis to always be present, many of the same clinical characteristics are considered in order to classify a child as having ERA as what is considered for the classification of nonradiographic SpA, as proposed by the ASAS (4).

In a recent publication, ERA was present in approximately 17% of all patients with JIA in 2 large JIA registries, including one from the Cincinnati Children's Hospital Medical Center Rheumatology Division and the other from the Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry (7). It is known that axial arthritis can occur in another JIA subset (i.e., juvenile psoriatic arthritis [PsA]), but the clinical phenotype of this JIA category is very heterogeneous (8). The US Food and Drug Administration (FDA) granted pharmaceutical companies that are studying new treatments in adult SpA automatic full waivers from doing studies in children for new medications for "axial spondyloarthropathies, including ankylosing spondylitis" (up until July 2020); these waivers were likely granted due to the differences in the nomenclature but also because of differences in the classification criteria for children with ERA and adults with axial SpA. Recent evidence supports the commonality of adult SpA and ERA with respect to genetics, pathogenesis, and clinical manifestations (9–13).

There are currently no drugs that have been approved for the ERA category of JIA. Recent research has indicated a large unmet medical need in the treatment of JIA, with 52–65% of all JIA patients (including those with ERA and juvenile PsA) having been treated with ≥1 biologic disease-modifying antirheumatic drug (DMARD) and 15–19% having been treated with a biologic that was not approved by the FDA (7). In those with ERA or juvenile

Daniel J. Lovell, MD, MPH, Hermine I. Brunner, MD, MSc, MBA: Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio.

Dr. Lovell has received consulting fees, speaking fees, and/or honoraria from AstraZeneca, Wyeth, Amgen, Abbott, Pfizer, Hoffmann-La Roche, Novartis, UCB, Takeda, Janssen, GlaxoSmithKline, Boehringer Ingelheim, Celgene, Roche, Bristol Myers Squibb, AbbVie, Forest Research, and NIN NIAMS (less than $10,000 each). Dr. Brunner has received consulting fees, speaking fees, and/or honoraria from Novartis, Roche, GlaxoSmithKline, Ablynx, AbbVie, Astra Zeneca-Medimmune, Biogen, Boehringer, Bristol Myers Squibb, Celgene, Eli Lilly and Company, EMD Serono, Genzyme, GlaxoSmithKline, F. Hoffmann-La Roche, Merck, Novartis, R-Pharm, and Sanofi (less than $10,000 each). No other potential conflicts of interest relevant to this article were reported.

Address correspondence to Hermine I. Brunner, MD, MSc, MBA, Division of Rheumatology, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, OH 45208. Email: Hermine.brunner@cchmc.org.

Submitted for publication November 5, 2020; accepted in revised form December 8, 2020.

PsA, 72–79% of the children had been treated with a biologic DMARD, although no biologic DMARD has ever been approved by the FDA for these JIA categories. Even more worrisome is that as many as 31–55% of the children demonstrated chronically uncontrolled arthritis, despite the use of these unapproved biologic DMARDs (7).

In this issue of *Arthritis Care & Research*, Rumsey et al report on the characteristics of JIA patients with pediatric-onset axial SpA using data from the CARRA registry (14). At the time of the analysis, this registry included longitudinal data from >5,600 children with JIA, including 522 children with ERA and 380 with juvenile PsA. Select characteristics of children for whom sacroiliitis was ever reported were compared between the first registry visit with clinically active sacroiliitis (which came first in 72% of cases) and the first registry visit without clinically active sacroiliitis. Presence of sacroiliitis was identified by the treating provider or could be based on either clinical or imaging data (MRI or computed tomography scan). Sacroiliitis was identified in 28% of the total group of children (n = 902) with either ERA or juvenile PsA (40% with ERA and 12% with juvenile PsA). The diagnosis of sacroiliitis was based on clinical characteristics in 38% and imaging in 62%. The characteristics of those with sacroiliitis highlighted several important findings. More than half were female, HLA–B27 was present in 36%, and 81% had been treated with ≥1 biologic DMARD. These children with sacroiliitis had significantly greater disease burden with higher physician assessment of disease activity, higher parent/patient global assessment of wellbeing, and higher disease activity as measured by the Juvenile Idiopathic Arthritis Disease Activity Score (15) compared to the children with ERA or juvenile PsA without sacroiliitis.

The study by Rumsey and colleagues has the following limitations: 1) the lack of specific criteria to determine if the ILAR criteria for ERA and juvenile PsA categories were accurately applied; 2) the absence of well-defined criteria for the presence/absence of sacroiliitis; 3) the lack of information on radiologic findings compatible with SI joint involvement that were used to diagnose radiologic as opposed to nonradiologic SpA; and 4) the large amount of missing data. However, despite these limitations, Rumsey et al demonstrated that there is a significant proportion of the overall JIA population with nonradiologic SpA, with phenotypes that are similar to those of nonradiographic SpA in adults. The authors reported that the group of children with nonradiologic SpA have a significant disease burden, and the vast majority of them were being treated with unapproved biologic DMARDs. Evaluation of disease manifestations (4,16) and MRI assessment and scoring (17,18) has been validated for use in children and adolescents with SpA and are similar to those used in adults with nonradiographic SpA.

In conclusion, the recent decision to remove children with juvenile SpA from the automatic waiver list is supported by the data from this study by Rumsey et al. This study showed that there is a significant proportion of the JIA population with active sacroiliitis with high disease burden despite very frequently (>80% of the population) being treated with unstudied and unapproved biologic DMARDs. Other recent publications demonstrate a significant number of these children continue with chronically uncontrolled arthritis despite use of these unapproved biologic DMARDs. It is now time for the pharmaceutical industry to perform FDA-monitored clinical trials of children and adolescents with SpA. This will allow for the scientific assessment of proper dosing, efficacy, and safety of the increasing number of new medications that are being licensed by the FDA for the treatment of SpA, such as anti-TNF inhibitors, anti–interleukin-17 [IL-17], and anti–IL-23 biologics, and perhaps JAK agents to address this unmet medical need in these patients with juvenile SpA.

AUTHOR CONTRIBUTIONS

Both authors drafted the article, revised it critically for important intellectual content, approved the final version to be published, and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Mease PJ, Heijde DV, Karki C, Palmer JB, Liu M, Pandurengan R, et al. Characterization of patients with ankylosing spondylitis and nonradiographic axial spondyloarthritis in the US-based Corrona Registry. Arthritis Care Res (Hoboken) 2018;70:1661–70.
2. Reveille JD. Epidemiology of spondyloarthritis in North America. Am J Med Sci 2011;341:284–6.
3. Reveille JD, Weisman MH. The epidemiology of back pain, axial spondyloarthritis and HLA-B27 in the United States. Am J Med Sci 2013;345:431–6.
4. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777–83.
5. Ward MM, Deodhar A, Akl EA, Lui A, Ermann J, Gensler LS, et al. American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and Treatment Network 2015 recommendations for the treatment of ankylosing spondylitis and nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2016;68:282–98.
6. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
7. Brunner HI, Schanberg LE, Kimura Y, Dennos A, Co DO, Colbert RA, et al. New medications are needed for children with juvenile idiopathic arthritis. Arthritis Rheumatol 2020;0:1–7.
8. Ravelli A, Consolaro A, Schiappapietra B, Martini A. The conundrum of juvenile psoriatic arthritis. Clin Exp Rheumatol 2015;33 Suppl 93:S40–3.
9. Hinks A, Martin P, Flynn E, Eyre S, Packham J, Childhood Arthritis Prospective Study C, et al. Subtype specific genetic associations for juvenile idiopathic arthritis: ERAP1 with the enthesitis related arthritis subtype and IL23R with juvenile psoriatic arthritis. Arthritis Res Ther 2011:R12.
10. Mielants H, Veys EM, Cuvelier C, De Vos M, Goemaere S, Maertens M, et al. Gut inflammation in children with late onset pauciarticular juvenile chronic arthritis and evolution to adult spondyloarthropathy: a prospective study. J Rheumatol 1993;20:1567–72.
11. Colbert RA, Navid F, Gill T. The role of HLA-B*27 in spondyloarthritis. Best Pract Res Clin Rheumatol 2017;31:797–815.

12. Rizzo A, Ferrante A, Guggino G, Ciccia F. Gut inflammation in spondyloarthritis. Best Pract Res Clin Rheumatol 2017;31:863–76.
13. Weiss PF, Colbert RA. Juvenile spondyloarthritis: a distinct form of juvenile arthritis. Pediatr Clin North Am 2018;65:675–90.
14. Rumsey DG, Lougee A, Matsouaka R, Collier DH, Schanberg LE, Schenfeld J, et al. Juvenile spondyloarthritis in the Childhood Arthritis and Rheumatology Research Alliance Registry: high biologic use, low prevalence of HLA–B27, and equal sex representation in sacroiliitis. Arthritis Care Res (Hoboken) 2021;73:940–6.
15. Consolaro A, Ruperto N, Bazso A, Pistorio A, Magni-Manzoni S, Filocamo G, et al. Paediatric Rheumatology International Trials Organisation. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. Arthritis Rheum 2009;61:658–66.
16. Goirand M, Breton S, Chevallier F, Duong NP, Uettwiller F, Melki I, et al. Clinical features of children with enthesitis-related juvenile idiopathic arthritis / juvenile spondyloarthritis followed in a French tertiary care pediatric rheumatology centre. Pediatr Rheumatol Online J 2018;16:21.
17. Weiss PF, Maksymowych WP, Xiao R, Biko DM, Francavilla ML, Lambert RG, et al. Spondyloarthritis Research Consortium of Canada sacroiliac joint inflammation and structural scores: change score reliability and recalibration utility in children. Arthritis Res Ther 2020;22:58.
18. Weiss PF, Maksymowych WP, Lambert RG, Jaremko JL, Biko DM, Paschke J, et al. Feasibility and reliability of the Spondyloarthritis Research Consortium of Canada sacroiliac joint structural score in children. J Rheumatol 2018;45:1411–7.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 924–939
DOI 10.1002/acr.24596

2021 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis

Liana Fraenkel,[1] Joan M. Bathon,[2] Bryant R. England,[3] E. William St.Clair,[4] Thurayya Arayssi,[5] Kristine Carandang,[6] Kevin D. Deane,[7] Mark Genovese,[8] Kent Kwas Huston,[9] Gail Kerr,[10] Joel Kremer,[11] Mary C. Nakamura,[12] Linda A. Russell,[13] Jasvinder A. Singh,[14] Benjamin J. Smith,[15] Jeffrey A. Sparks,[16] Shilpa Venkatachalam,[17] Michael E. Weinblatt,[16] Mounir Al-Gibbawi,[18] Joshua F. Baker,[19] Kamil E. Barbour,[20] Jennifer L. Barton,[21] Laura Cappelli,[22] Fatimah Chamseddine,[18] Michael George,[23] Sindhu R. Johnson,[24] Lara Kahale,[18] Basil S. Karam,[18] Assem M. Khamis,[18] Iris Navarro-Millán,[25] Reza Mirza,[26] Pascale Schwab,[21] Namrata Singh,[27] Marat Turgunbaev,[28] Amy S. Turner,[28] Sally Yaacoub,[18] and Elie A. Akl[18]

Guidelines and recommendations developed and/or endorsed by the American College of Rheumatology (ACR) are intended to provide general guidance for commonly encountered clinical scenarios. The recommendations do not dictate the care for an individual patient. The ACR considers adherence to the recommendations described in this guideline to be voluntary, with the ultimate determination regarding their application to be made by the clinicians in light of each patient's individual circumstances. Guidelines and recommendations are intended to promote beneficial or desirable outcomes but cannot guarantee any specific outcome. Guidelines and recommendations developed and endorsed by the ACR are subject to periodic revision as warranted by the evolution of medical knowledge, technology, and practice. ACR recommendations are not intended to dictate payment or insurance decisions, or drug formularies or other third-party analyses. Third parties that cite ACR guidelines should state that these recommendations are not meant for this purpose. These recommendations cannot adequately convey all uncertainties and nuances of patient care.

The American College of Rheumatology is an independent, professional, medical and scientific society that does not guarantee, warrant, or endorse any commercial product or service.

Objective. To develop updated guidelines for the pharmacologic management of rheumatoid arthritis.

Methods. We developed clinically relevant population, intervention, comparator, and outcomes (PICO) questions. After conducting a systematic literature review, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach was used to rate the certainty of evidence. A voting panel comprising clinicians and patients achieved consensus on the direction (for or against) and strength (strong or conditional) of recommendations.

Results. The guideline addresses treatment with disease-modifying antirheumatic drugs (DMARDs), including conventional synthetic DMARDs, biologic DMARDs, and targeted synthetic DMARDs, use of glucocorticoids, and use of DMARDs in certain high-risk populations (i.e., those with liver disease, heart failure, lymphoproliferative disorders, previous serious infections, and nontuberculous mycobacterial lung disease). The guideline includes 44 recommendations (7 strong and 37 conditional).

Conclusion. This clinical practice guideline is intended to serve as a tool to support clinician and patient decision-making. Recommendations are not prescriptive, and individual treatment decisions should be made through a shared decision-making process based on patients' values, goals, preferences, and comorbidities.

The findings and conclusions herein are those of the authors and do not represent the official position of the Centers for Disease Control and Prevention. This study did not involve human subjects, and therefore, approval from Human Studies Committees was not required.

This article is published simultaneously in *Arthritis & Rheumatology*.

Supported by the American College of Rheumatology.

[1]Liana Fraenkel, MD, MPH: Berkshire Medical Center, Pittsfield, Massachusetts, and Yale University School of Medicine, New Haven, Connecticut; [2]Joan M. Bathon, MD: Columbia University Irving Medical Center, New York Presbyterian Hospital, New York, New York; [3]Bryant R. England, MD, PhD: University of Nebraska Medical Center and VA Nebraska–Western Iowa Health Care System, Omaha, Nebraska; [4]E. William St.Clair, MD: Duke

INTRODUCTION

To support high-quality clinical care, the American College of Rheumatology (ACR) regularly updates clinical practice guidelines for the management of rheumatoid arthritis (RA), with the most recent update reported in 2015 (1). The current recommendations address treatment with the following: 1) conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), biologic DMARDs (bDMARDs), and targeted synthetic DMARDs (tsDMARDs); 2) glucocorticoids; and 3) use of these medications in certain high-risk populations. The use of vaccines and nonpharmacologic treatment approaches (although initially part of this project) will be covered in future ACR treatment guideline publications. For recommendations regarding pretreatment screening and routine laboratory monitoring, we refer readers to the 2008, 2012, and 2015 guidelines (1–3), with newly approved therapies following the screening process recommended for other medications in the same class. Recommendations for the perioperative management of patients undergoing elective orthopedic surgery are addressed in the 2017 guideline for perioperative management (4). For recommendations regarding reproductive health, we refer readers to the 2020 ACR Guideline for the Management of Reproductive Health in Rheumatic and Musculoskeletal Diseases (5).

In keeping with the Grading of Recommendations Assessment, Development and Evaluation [GRADE] methodology), the ACR panel developed recommendations for commonly encountered clinical scenarios (6–8). Both ***strong*** and ***conditional*** recommendations required achieving a 70% level of agreement by the voting panel. Each recommendation is qualified as being strong or conditional. In this context, strong recommendations are those for which the panel is highly confident that the recommended option favorably balances the expected benefits and risks for the majority of patients in clinical practice. In contrast, conditional recommendations are those for which the panel is less confident that the potential benefits outweigh the risks. A recommendation can be conditional either because of low or very low certainty in the evidence supporting one option over another, or because of an expectation of substantial variations in patient preferences for the options under consideration.

METHODS

This guideline follows the ACR guideline development process and ACR policy guiding the management of conflicts of interest and disclosures (https://www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines) (6,8), which includes GRADE methodology (6,8), and abides by the AGREE Reporting Checklist to ensure the completeness and transparency of reporting in practice guidelines (9). Supplementary Appendix 1, available on the *Arthritis Care & Research* website at http://onlin

University Medical Center, Durham, North Carolina; [5]Thurayya Arayssi, MD: Weill Cornell Medicine–Qatar, Doha, Qatar; [6]Kristine Carandang, PhD, OTR/L: University of California, San Diego; [7]Kevin D. Deane, MD, PhD: University of Colorado, Aurora; [8]Mark Genovese, MD: Stanford University Medical Center, Palo Alto, California; [9]Kent Kwas Huston, MD: The Center for Rheumatic Disease/Allergy and Immunology, Kansas City, Missouri; [10]Gail Kerr, MD, FRCP (Edin): Veterans Affairs Medical Center, Georgetown and Howard University, Washington, DC; [11]Joel Kremer, MD: Albany Medical College and The Center for Rheumatology, Albany, New York; [12]Mary C. Nakamura, MD: University of California, San Francisco; [13]Linda A. Russell, MD: Hospital for Special Surgery, New York, New York; [14]Jasvinder A. Singh, MD, MPH: University of Alabama at Birmingham and Birmingham Veterans Affairs Medical Center, Birmingham, Alabama; [15]Benjamin J. Smith, DMSc, PA-C: Florida State University College of Medicine School of Physician Assistant Practice, Tallahassee; [16]Jeffrey A. Sparks, MD, MMSc, Michael E. Weinblatt, MD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [17]Shilpa Venkatachalam, PhD: Global Healthy Living Foundation, Upper Nyack, New York; [18]Mounir Al-Gibbawi, MD, Fatimah Chamseddine, MD, Lara Kahale, RN, MSc, PhD, Basil S. Karam, MD, Assem M. Khamis, MD, MPH, Sally Yaacoub, RPh, MPH, Elie A. Akl, MD, MPH, PhD: American University of Beirut, Beirut, Lebanon; [19]Joshua F. Baker, MD, MSCE: Corporal Michael J. Crescenz VA Medical Center and the University of Pennsylvania, Philadelphia, Pennsylvania; [20]Kamil E. Barbour, PhD, MPH, MS: Centers for Disease Control and Prevention, Atlanta, Georgia; [21]Jennifer L. Barton, MD, MCR, Pascale Schwab, MD: Oregon Health & Science University and VA Portland Health Care System, Portland, Oregon; [22]Laura Cappelli, MD, MHS, MS: Johns Hopkins Medicine, Baltimore, Maryland; [23]Michael George, MD, MSCE: University of Pennsylvania, Philadelphia; [24]Sindhu R. Johnson, MD, PhD, FRCPC: Toronto Western Hospital, Mount Sinai Hospital, Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada; [25]Iris Navarro-Millán, MD, MSPH: Weill Cornell Medicine, New York, New York; [26]Reza Mirza, MD: University of Toronto, Toronto, Ontario, Canada; [27]Namrata Singh, MD, MSCI, FACP: University of Washington, Seattle; [28]Marat Turgunbaev, MD, MPH, Amy S. Turner: American College of Rheumatology, Atlanta, Georgia.

Dr. Bathon has received consulting fees, speaking fees, and/or honoraria from Gilead Biosciences (less than $10,000). Dr. St.Clair has received consulting fees, speaking fees, and/or honoraria from Bristol Myers Squibb (less than $10,000) and royalties from UpToDate. Dr. Deane has received consulting fees, speaking fees, and/or honoraria from Bristol Myers Squibb, Inova Diagnostics, Janssen, and Microdrop (less than $10,000 each). Dr. Genovese has received consulting fees, speaking fees, and/or honoraria from AbbVie, Amgen, BeiGene, Eli Lilly and Company, GlaxoSmithKline, Pfizer, Sanofi, R-Pharm, and SetPoint Medical (less than $10,000 each) and owns stock or stock options in Gilead Sciences. Dr. Huston has received consulting fees and/or honoraria from Trio Health Advisory (less than $10,000). Dr. Kerr has received consulting fees, speaking fees, and/or honoraria from Jansen, Sanofi, and Genzyme (less than $10,000 each). Dr. Kremer has received consulting fees, speaking fees, and/or honoraria from AbbVie, Bristol Myers Squibb, Gilead, Eli Lilly and Company, and Pfizer (less than $10,000 each) and owns stock or stock options in Corrona. Dr. Singh has received consulting fees, speaking fees, and/or honoraria from MediSYS, Fidia Pharma, UBM, Trio Health, Medscape, WebMD, Clinical Care Options, Clearview Healthcare Partners, Putnam Associates, Focus Forward, Navigant Consulting, Spherix, Practice Point Communications (less than $10,000 each), and Simply Speaking (more than $10,000) and owns stock or stock options in Vaxart Pharmaceuticals. Dr. Weinblatt has received consulting fees, speaking fees, and/or honoraria from AbbVie, Amgen, Canfite, Crescendo, GlaxoSmithKline, Horizon, Johnson & Johnson, Genentech/Roche, Scipher, SetPoint, Tremeau (less than $10,000 each), Eli Lilly and Company, Bristol Myers Squibb, Gilead, and Corrona (more than $10,000 each), royalties from Elsevier, and owns stock or stock options in Canfite, Scipher, Inmedix, and Vorso. Dr. Baker has received consulting fees, speaking fees, and/or honoraria from Bristol Myers Squibb, Gilead, Burns-White, and Simply Speaking Rheumatology (less than $10,000 each). Dr. Cappelli has received consulting fees, speaking fees, and/or honoraria from AbbVie (less than $10,000) and research support from Bristol Myers Squibb. Dr. George has received honoraria from Paradigm Medical Communications (less than $10,000) and research support from Bristol Myers Squibb. Dr. Johnson has received consulting fees and/or honoraria from Boehringer Ingelheim and Ikaria (less than $10,000 each) and research support from GlaxoSmithKline, Corbus, Roche, Merck, Boehringer Ingelheim, and Bayer. No other disclosures relevant to this article were reported.

Address correspondence to Liana Fraenkel, MD, MPH, 725 North Street, Pittsfield, MA 01201. Email: liana.fraenkel@yale.edu.

Submitted for publication September 8, 2020; accepted in revised form March 15, 2021.

elibrary.wiley.com/doi/10.1002/acr.24596/abstract), includes a detailed description of the methods. Briefly, the core leadership team drafted clinical population, intervention, comparator, and outcomes (PICO) questions. The literature review team performed systematic literature reviews for the PICO questions, selected and evaluated individual studies and graded the quality of the body of evidence available for each outcome, and produced the evidence report that summarizes these assessments (see Supplementary Appendix 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract). The core team defined the critical study outcome as disease activity for most PICO questions. Because the ACR has, in a separate project, endorsed several disease activity measures for use in clinical practice, this guideline does not define levels of disease activity or the instruments a clinician should use to measure it (10). For PICO questions related to tapering, the critical outcomes were disease flare and subsequent return to the treatment target. Physical function, radiographic progression, quality of life, other patient-reported outcome measures, and adverse events were defined as important outcomes. Additional clinical outcomes were defined for PICO questions pertaining to select high-risk conditions (see Supplementary Appendix 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract). When available, cost-effectiveness studies were included with the evidence reports. Cost estimates (average wholesale prices) were retrieved from Lexicomp (see Supplementary Appendix 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract). The panel considered these estimates from a societal perspective, i.e., based on the list price, and not the copay.

An in-person panel of 10 patients with RA, moderated by the project's principal investigator, reviewed the evidence report (along with a summary and interpretation by the moderator) and provided patient perspectives for consideration by the voting panel. The voting panel (13 clinicians and 2 patients) reviewed the evidence reports and patient perspectives and voted on recommendation statements. Rosters of the core leadership, literature review team, and panel members are listed in Supplementary Appendix 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract.

Several guiding principles, definitions, and assumptions were established a priori (Table 1). Because poor prognostic factors (11) have had less impact than other factors on prior RA treatment recommendations, they were not explicitly considered in formulating the PICO questions. However, poor prognostic factors were considered as possible influential factors in physicians' and patients' decision-making when developing recommendations. In contrast to the 2015 guideline (1), recommendations were not provided for subgroups defined by early versus late RA disease duration. This change was made because current disease activity, prior therapies used, and the presence of comorbidities were felt to be more relevant than disease duration for most treatment decisions. However, early diagnosis and treatment in RA is associated with improved outcomes and is thus an important overarching principle in its management (12). Recommendations are intended for the general RA patient population and assume that patients do not have contraindications to the options under consideration.

Table 1. Guiding principles*

RA requires early evaluation, diagnosis, and management.
Treatment decisions should follow a shared decision-making process.
Treatment decisions should be reevaluated within a minimum of 3 months based on efficacy and tolerability of the DMARD(s) chosen.
Disease activity levels refer to those calculated using RA disease activity measures endorsed by the ACR (10).
Recommendations are intended for the general RA patient population and assume that patients do not have contraindications to the options under consideration.
Recommendations are limited to DMARDs approved by the US FDA for treatment of RA. csDMARDs: hydroxychloroquine, sulfasalazine, methotrexate, leflunomide bDMARDs: TNF inhibitors (etanercept, adalimumab, infliximab, golimumab, certolizumab pegol), T cell costimulatory inhibitor (abatacept), IL-6 receptor inhibitors (tocilizumab, sarilumab), anti-CD20 antibody (rituximab)† tsDMARDs: JAK inhibitors (tofacitinib, baricitinib, upadacitinib)
Triple therapy refers to hydroxychloroquine, sulfasalazine, and either methotrexate or leflunomide.
Serious infection refers to an infection requiring intravenous antibiotics or hospitalization.
Biosimilars are considered equivalent to FDA-approved originator bDMARDs.
Recommendations referring to bDMARDs exclude rituximab unless patients have had an inadequate response to TNF inhibitors (in order to be consistent with FDA approval) or have a history of lymphoproliferative disorder for which rituximab is an approved therapy.
Treat-to-target refers to a systematic approach involving frequent monitoring of disease activity using validated instruments and modification of treatment to minimize disease activity with the goal of reaching a predefined target (low disease activity or remission).
Target refers to low disease activity or remission.
Recommendations specify that patients be at target (low disease activity or remission) for at least 6 months prior to tapering.
Dose reduction refers to lowering the dose or increasing the dosing interval of a DMARD. Gradual discontinuation of a DMARD is defined as gradually lowering the dose of a DMARD and subsequently stopping it.

* RA = rheumatoid arthritis; DMARDs = disease-modifying antirheumatic drugs; ACR = American College of Rheumatology; FDA = Food and Drug Administration; csDMARDs = conventional DMARDs; bDMARDs = biologic DMARDs; TNF = tumor necrosis factor; IL-6 = interleukin-6; tsDMARDs = targeted synthetic DMARDs.
† Anakinra was not included due to infrequent use for patients with RA.

RESULTS/RECOMMENDATIONS

The recommendations are based on a set of 81 PICO questions. The literature review initially identified 22,971 manuscripts (for the full set of PICO questions covering both pharmacologic and nonpharmacologic treatment). After excluding 18,333 titles

and abstracts, 4,038 full-text articles were screened, of which 1,392 were excluded and 2,646 were considered for the evidence report. After full-text screening, 133 manuscripts were mapped to ≥1 PICO questions addressing pharmacologic treatment (see Supplementary Appendix 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract). The literature review did not identify any evidence for 41% (n = 33) of the PICO questions.

Recommendations for DMARD-naive patients with moderate-to-high disease activity (Table 2)

DMARD monotherapy

Methotrexate is strongly recommended over hydroxychloroquine or sulfasalazine for DMARD-naive patients with moderate-to-high disease activity

This recommendation is strongly in favor of methotrexate despite very low-certainty evidence for hydroxychloroquine and low-certainty evidence for sulfasalazine based on the amount of data supporting the disease-modifying properties of methotrexate monotherapy compared to hydroxychloroquine or sulfasalazine and concerns over the long-term tolerability of sulfasalazine (13,14).

Methotrexate is conditionally recommended over leflunomide for DMARD-naive patients with moderate-to-high disease activity

Despite low-certainty evidence of comparable efficacy, methotrexate is preferred over leflunomide because of the evidence supporting its value as an anchor DMARD in combination regimens. Additional advantages of methotrexate include its greater dosing flexibility and lower cost.

Methotrexate monotherapy is strongly recommended over bDMARD or tsDMARD monotherapy for DMARD-naive patients with moderate-to-high disease activity

There is low-certainty evidence suggesting superiority of tocilizumab monotherapy (15) over methotrexate monotherapy and moderate-certainty evidence suggesting greater efficacy

Table 2. Disease-modifying antirheumatic drugs (DMARDs) initiation*

Recommendations	Certainty of evidence	Based on the evidence report(s) of the following PICO(s)†	Evidence table(s), in Supp. App. 2
Initiation of treatment in DMARD-naive patients with moderate-to-high disease activity			
Methotrexate monotherapy is **strongly** recommended over:			
Hydroxychloroquine or sulfasalazine	Very low/low‡	PICO 2a.C1/C2	p. 14–5
bDMARD or tsDMARD monotherapy	Very low/moderate	PICO 5a.C1–4/C5§	p. 61–78
Combination of methotrexate plus a non–TNF inhibitor bDMARD or tsDMARD¶	Low/very low	PICO 6a.C2–4/C5§	p. 109, 117–28
Methotrexate monotherapy is **conditionally** recommended over:			
Leflunomide	Low	PICO 2a.C3	p. 18
Dual or triple csDMARD therapy¶	Moderate	PICO 4a.C1–C2	p. 46–9
Combination of methotrexate plus a TNF inhibitor¶	Low	PICO 6a.C1	p. 110
Initiation of a csDMARD without short-term (<3 months) glucocorticoids is **conditionally** recommended over initiation of a csDMARD with short-term glucocorticoids.	Very low	PICO 7a	p. 167
Initiation of a csDMARD without longer-term (≥3 months) glucocorticoids is **strongly** recommended over initiation of a csDMARD with longer-term glucocorticoids.	Moderate	PICO 8a	p. 170
Initiation of treatment in DMARD-naive patients with low disease activity			
Hydroxychloroquine is **conditionally** recommended over other csDMARDs.	Very low	PICO 1a.C1–4	p. 1–6
Sulfasalazine is **conditionally** recommended over methotrexate.	Very low	PICO 1a.C2	p. 2
Methotrexate is **conditionally** recommended over leflunomide.	Very low	PICO 1a.C3	p. 5
Initiation of treatment in csDMARD-treated, but methotrexate-naive, patients with moderate-to-high disease activity#			
Methotrexate monotherapy is **conditionally** recommended over the combination of methotrexate plus a bDMARD or tsDMARD.**	Moderate/very low	PICO 6b.C1–4/C5§	p. 136–56

* PICO = population, intervention, comparator, and outcomes; Supp. App. 2 = Supplementary Appendix 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract; bDMARD = biologic DMARD; tsDMARD = targeted synthetic DMARD; TNF = tumor necrosis factor; csDMARD = conventional synthetic DMARD.
† The closest matching PICO questions to each recommendation are provided.
‡ The first certainty of evidence applies to the first listed option; the second certainty of evidence applies to the second listed option.
§ The original PICO included individual DMARDs as comparators. The recommendation considers bDMARDs as a group.
¶ The direction of the beneficial effect is in favor of the nonpreferred option.
# Other recommendations for this patient population are the same as those for DMARD-naive patients.
** The direction of the beneficial effect is in favor of the nonpreferred option. The certainty of evidence is high for the combination of methotrexate plus a TNF inhibitor and moderate for other bDMARDs.

of JAK inhibitor monotherapy over methotrexate monotherapy. The study by van Vollenhoven et al (16) was not considered by the voting panel as it was published after the evidence report was updated. However, methotrexate monotherapy is preferred because of its established efficacy and safety as a first-line DMARD and low cost. Moreover, tocilizumab and JAK inhibitors are not approved by the US Food and Drug Administration (FDA) for use in csDMARD-naive patients. Safety concerns released in early 2021 associated with JAK inhibitors (17,18) further support the recommendation of methotrexate monotherapy over tsDMARDs as initial DMARD therapy at this time.

Methotrexate monotherapy is conditionally recommended over dual or triple csDMARD therapy for DMARD-naive patients with moderate-to-high disease activity

The recommendation favors methotrexate monotherapy because the higher burden of combination therapy (e.g., multiple medications, higher cost) outweighs the moderate-quality evidence suggesting greater improvements in disease activity associated with combination csDMARDs (19). The recommendation is conditional because some patients may choose csDMARD combination therapy for an increased probability of obtaining a better response despite the added burden of taking multiple medications.

Methotrexate monotherapy is conditionally recommended over methotrexate plus a tumor necrosis factor (TNF) inhibitor for DMARD-naive patients with moderate-to-high disease activity

Despite low-certainty evidence supporting greater improvement in disease activity with methotrexate plus a TNF inhibitor, methotrexate monotherapy is preferred over the combination because many patients will reach their goal on methotrexate monotherapy and because of the additional risks of toxicity and higher costs associated with TNF inhibitors. The recommendation is conditional because some patients, especially those with poor prognostic factors, may prioritize more rapid onset of action and greater chance of improvement associated with combination therapy (20–22) over the additional risks and costs associated with initial use of methotrexate in combination with a TNF inhibitor.

Methotrexate monotherapy is strongly recommended over methotrexate plus a non–TNF inhibitor bDMARD or tsDMARD for DMARD-naive patients with moderate-to-high disease activity

There is very low-certainty evidence supporting the superiority of methotrexate plus a non–TNF inhibitor bDMARD or tsDMARD over methotrexate monotherapy in DMARD-naive patients; thus, methotrexate monotherapy is strongly preferred given the lack of proven benefit and additional risks and costs associated with the addition of a non–TNF inhibitor bDMARD or tsDMARD in this patient population.

Glucocorticoids

Initiation of a csDMARD without short-term (<3 months) glucocorticoids is conditionally recommended over initiation of a csDMARD with short-term glucocorticoids for DMARD-naive patients with moderate-to-high disease activity

While the voting panel agreed that glucocorticoids should not be systematically prescribed, the recommendation is conditional because all members acknowledged that short-term glucocorticoids are frequently necessary to alleviate symptoms prior to the onset of action of DMARDs. Treatment with glucocorticoids should be limited to the lowest effective dose for the shortest duration possible. The toxicity associated with glucocorticoids was judged to outweigh potential benefits.

Initiation of a csDMARD without longer-term (≥3 months) glucocorticoids is strongly recommended over initiation of a csDMARD with longer-term glucocorticoids for DMARD-naive patients with moderate-to-high disease activity

Although some patients may require longer-term glucocorticoids, this strong recommendation *against* longer-term glucocorticoid therapy is made because of its significant toxicity.

Recommendations for DMARD-naive patients with low disease activity (Table 2)

Hydroxychloroquine is conditionally recommended over other csDMARDs, sulfasalazine is conditionally recommended over methotrexate, and methotrexate is conditionally recommended over leflunomide for DMARD-naive patients with low disease activity

Hydroxychloroquine is conditionally recommended over other csDMARDs because it is better tolerated and has a more favorable risk profile in patients with RA. Sulfasalazine is recommended over methotrexate because it is less immunosuppressive, and the patient panel felt that many patients with low disease activity would prefer to avoid the side effects associated with methotrexate. The recommendations are conditional because methotrexate may be the preferred initial therapy in patients at the higher end of the low disease activity range and in those with poor prognostic factors (11). Methotrexate is recommended over leflunomide because of its greater dosing flexibility and lower cost.

Recommendation for patients who have been treated with csDMARDs, excluding methotrexate, and who have moderate-to-high disease activity (Table 2)

Recommendations are the same as for DMARD-naive patients except for this population. The strength of the following recommendation is conditional for all bDMARDs and tsDMARDs.

Methotrexate monotherapy is conditionally recommended over the combination of methotrexate plus a bDMARD or tsDMARD

The recommendation is conditional because the voting panel thought that some patients who have already had persistent disease activity despite use of ≥1 csDMARD will prefer combination treatment for a more rapid response.

Recommendations for administration of methotrexate (Table 3)

Oral methotrexate is conditionally recommended over subcutaneous methotrexate for patients initiating methotrexate

Oral administration is preferred, despite moderate evidence suggesting superior efficacy of subcutaneous injections, due to the ease of oral administration and similar bioavailability at typical starting doses (23).

Initiation/titration of methotrexate to a weekly dose of at least 15 mg within 4 to 6 weeks is conditionally recommended over initiation/titration to a weekly dose of <15 mg

The recommendation is conditional because there are few studies comparing different dosing strategies and wide variation in physician and patient preferences regarding the tradeoff between the increased efficacy and risks of toxicity associated with higher starting doses. This recommendation refers only to the initial prescribing of methotrexate and is not meant to limit further dose escalation, which often provides additional efficacy (24).

A split dose of oral methotrexate over 24 hours or weekly subcutaneous injections, and/or an increased dose of folic/folinic acid, is conditionally recommended over switching to alternative DMARD(s) for patients not tolerating oral weekly methotrexate

Despite the very low certainty of evidence supporting these strategies for alleviating side effects related to methotrexate, split dosing, changing to the subcutaneous route of administration, and increased doses of folic/folinic acid are the preferred initial strategies over switching to another DMARD because of the efficacy, long-term safety, and low costs associated with methotrexate. The recommendation is conditional because patient preferences play an important role in the decision whether to continue methotrexate or switch to other DMARDs.

Switching to subcutaneous methotrexate is conditionally recommended over the addition of/switching to alternative DMARD(s) for patients taking oral methotrexate who are not at target

This recommendation is consistent with the voting panel's overarching principle of maximizing use of methotrexate prior to switching/adding DMARDs. However, there are no data comparing outcomes in patients who switch to subcutaneous methotrexate versus another treatment strategy that includes other DMARDs. The recommendation is conditional because patient preferences and the magnitude of previous response to methotrexate play an important role in this decision.

Table 3. Methotrexate administration*

Recommendations	Certainty of evidence	Based on the evidence report(s) of the following PICO(s)	Evidence table(s), in Supp. App. 2
Oral methotrexate is **conditionally** recommended over subcutaneous methotrexate for patients initiating methotrexate.	Moderate	PICO 9	p. 181
Initiation/titration of methotrexate to a weekly dose of at least 15 mg within 4 to 6 weeks is **conditionally** recommended over initiation/titration to a weekly dose of <15 mg.†	Moderate/ very low‡	PICO 10.C1–C3	p. 184–5
A split dose of oral methotrexate over 24 hours or subcutaneous injections, and/or an increased dose of folic/folinic acid, is **conditionally** recommended over switching to alternative DMARD(s) for patients not tolerating oral weekly methotrexate.	Very low	PICO 16 and PICO 15	p. 206–10
Switching to subcutaneous methotrexate is **conditionally** recommended over the addition of/switching to alternative DMARD(s) for patients taking oral methotrexate who are not at target.	Very low	PICO 18	p. 235

* PICO = population, intervention, comparator, and outcomes; Supp. App. 2 = Supplementary Appendix 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract; DMARD = disease-modifying antirheumatic drug.
† This recommendation refers only to the initial prescribing of methotrexate and is not meant to limit further dose escalation, which often provides additional efficacy.
‡ The first certainty of evidence applies to the first listed option; the second certainty of evidence applies to the second option.

Recommendations for treatment modification in patients treated with DMARDs who are not at target (Table 4)

Treat-to-target

A treat-to-target approach is strongly recommended over usual care for patients who have not been previously treated with bDMARDs or tsDMARDs

This recommendation applies to dose optimization of methotrexate and to the subsequent addition of DMARDs when required. The recommendation is strong despite low-certainty evidence because of the recognized importance of systematic monitoring and adjustment of treatment to minimize inflammation to prevent joint damage, as well as other long-term sequelae including cardiovascular disease and osteoporosis.

A treat-to-target approach is conditionally recommended over usual care for patients who have had an inadequate response to bDMARDs or tsDMARDs

The recommendation is conditional because of the uncertain incremental benefits of treat-to-target over usual care in this patient population. In this context, usual care refers to commonly employed practice patterns, i.e., adjustment of treatment based on shared decision-making, albeit typically without systematic monitoring of disease activity using validated measures to reach a predefined target. Moreover, 1) the number of remaining available treatment options, 2) the impact of noninflammatory causes of pain, comorbidities, and/or damage on the accuracy of validated disease activity assessments, and 3) the patient's threshold for changing medications may have a more significant influence on the decision to follow a treat-to-target approach in this population compared to patients who are bDMARD- and tsDMARD-naive.

A minimal initial treatment goal of low disease activity is conditionally recommended over a goal of remission

An initial target of low disease activity is preferred because remission by established criteria may not be achievable for many patients (25). In addition, the patient panel emphasized that failure to reach a specified target may be disheartening and stressful for some patients. They emphasized that it would be preferable to *initially* aim for low disease activity and *subsequently* consider a goal of remission. However, treatment goals should be systematically reassessed over time and individualized to each patient to ensure that remission is targeted when possible. The recommendation is conditional because remission is a reasonable initial goal for patients with early disease and minimal exposure to bDMARDs and tsDMARDs, and patient preferences play a significant role in this decision.

Modification of DMARD(s)

Addition of a bDMARD or tsDMARD is conditionally recommended over triple therapy (i.e., addition of sulfasalazine and hydroxychloroquine) for patients taking maximally tolerated doses of methotrexate who are not at target

The panel vigorously debated whether to recommend addition of a bDMARD or tsDMARD versus sulfasalazine and

Table 4. Treatment modification*

Recommendations	Certainty of evidence	Based on the evidence report(s) of the following PICO(s)	Evidence table(s), in Supp. App. 2
A TTT approach is **strongly** recommended over usual care for patients who have not been previously treated with bDMARDs or tsDMARDs.	Low	PICO 12.a	p. 191
A TTT approach is **conditionally** recommended over usual care for patients who have had an inadequate response to bDMARDs or tsDMARDs.	Very low	PICO 12.b	p. 199
A minimal initial treatment goal of low disease activity is **conditionally** recommended over a goal of remission.	Low	PICO 13	p. 201
Addition of a bDMARD or tsDMARD is **conditionally** recommended over triple therapy for patients taking maximally tolerated doses of methotrexate who are not at target.	Very low	PICO 19.C2–C6†	p. 240–1
Switching to a bDMARD or tsDMARD of a different class is **conditionally** recommended over switching to a bDMARD or tsDMARD belonging to the same class for patients taking a bDMARD or tsDMARD who are not at target.	Very low	PICO 24–27†	p. 293–338
Addition of/switching to DMARDs is **conditionally** recommended over continuation of glucocorticoids for patients taking glucocorticoids to remain at target.	Very low	PICO 23	p. 292
Addition of/switching to DMARDs (with or without IA glucocorticoids) is **conditionally** recommended over the use of IA glucocorticoids alone for patients taking DMARDs who are not at target.	Very low	PICO 28.C1–C2	p. 339–40

* PICO = population, intervention, comparator, and outcomes; Supp. App. 2 = Supplementary Appendix 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract; TTT = treat-to-target; bDMARDs = biologic disease-modifying antirheumatic drugs; tsDMARDs = targeted synthetic DMARDs; IA = intraarticular.
† The original PICO included individual DMARDs as comparators. The recommendation considers bDMARDs as a group.

hydroxychloroquine (triple therapy) for patients with an inadequate response to methotrexate monotherapy in view of very low-certainty evidence favoring bDMARDs or tsDMARDs, randomized controlled trials demonstrating equivalent long-term outcomes across both treatment strategies, and significantly less societal cost associated with triple therapy (26–29). Addition of a bDMARD or tsDMARD was ultimately preferred because the patient panel strongly prioritized maximizing improvement as quickly as possible. In addition, both the patient and voting panels valued the greater persistence of methotrexate plus a bDMARD or tsDMARD compared to triple therapy (defined in Table 1) (13,30). The recommendations from these studies (13,31) are conditional because triple therapy may be preferred in lower resource settings as well as in patients with specific comorbidities for whom triple therapy may be associated with significantly less risk of adverse events. This choice is highly preference sensitive, and decisions on how best to escalate care should incorporate patients' preferences. There is no current recommendation for a bDMARD versus a tsDMARD when adjusting treatment; however, the voting panel acknowledged that safety data released in early 2021 (17,18) may require a modification of this recommendation when peer-reviewed results are published.

Switching to a bDMARD or tsDMARD of a different class is conditionally recommended over switching to a bDMARD or tsDMARD belonging to the same class for patients taking a bDMARD or tsDMARD who are not at target

The recommendation is based on very low-certainty evidence supporting greater improvement in disease activity and drug survival among patients switching classes. The recommendation is conditional because patient and physician preferences are likely to vary based on prior experiences with specific DMARDs.

Use of glucocorticoids

Addition of/switching to DMARDs is conditionally recommended over continuation of glucocorticoids for patients taking glucocorticoids to remain at target

This recommendation assumes that improved disease control with DMARDs should allow less use of glucocorticoids. The recommendation is conditional because the continued use of glucocorticoids may be required for patients who do not respond to DMARDs even after maximizing methotrexate dosage and switching DMARD classes.

Addition of/switching to DMARDs (with or without intraarticular [IA] glucocorticoids) is conditionally recommended over the use of IA glucocorticoids alone for patients taking DMARDs who are not at target

This recommendation was based on the premise that DMARDs should be adjusted to reduce disease activity, irrespective of treatment with IA glucocorticoids. The recommendation is conditional because patients may choose to defer adding/switching DMARDs if they obtain relief from IA injection(s).

Recommendations for tapering/discontinuing DMARDs (Table 5)

Because of the moderate-to-high risk for flare and the potential for irreversible long-term damage associated with stopping all DMARDs, the following recommendations presume that patients maintain a therapeutic dose of at least 1 DMARD. In addition, the recommendations specify that patients be at target (low disease activity or remission) for at least 6 months

Table 5. Tapering disease-modifying antirheumatic drugs (DMARDs)*

Recommendations	Certainty of evidence	Based on the evidence report(s) of the following PICO(s)	Evidence table(s), in Supp. App. 2
Continuation of all DMARDs at their current dose is **conditionally** recommended over a dose reduction of a DMARD.	Low	PICO 54.a	p. 381
Dose reduction is **conditionally** recommended over gradual discontinuation of a DMARD.	Low	PICO 52.C2 and PICO 53. C2	p. 351–5, p. 372–6
Gradual discontinuation is **conditionally** recommended over abrupt discontinuation of a DMARD.	Low	PICO 52.C1 and PICO 53.C1	p. 351, 372
Gradual discontinuation of sulfasalazine is **conditionally** recommended over gradual discontinuation of hydroxychloroquine for patients taking triple therapy who wish to discontinue a DMARD.	Very low	PICO 58	p. 400
Gradual discontinuation of methotrexate is **conditionally** recommended over gradual discontinuation of the bDMARD or tsDMARD for patients taking methotrexate plus a bDMARD or tsDMARD who wish to discontinue a DMARD.	Very low	PICO 59.C1	p. 401

* PICO = population, intervention, comparator, and outcomes; Supp. App. 2 = Supplementary Appendix 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract; bDMARD = biologic DMARD; tsDMARD = targeted synthetic DMARD.

prior to tapering. Patients in remission for <6 months should not routinely be considered for dose reduction or withdrawal. Although the optimal time at target prior to tapering has not been established, the voting panel considered 6 months to be a reasonable minimal length of time to ensure stable disease control. "Dose reduction" refers to lowering the dose or increasing the dosing interval of a DMARD. "Gradual discontinuation" denotes gradually lowering the dose of a DMARD and subsequently stopping it.

Continuation of all DMARDs at their current dose is conditionally recommended over a dose reduction of a DMARD, dose reduction is conditionally recommended over gradual discontinuation of a DMARD, and gradual discontinuation is conditionally recommended over abrupt discontinuation of a DMARD for patients who are at target for at least 6 months

These recommendations are based on studies demonstrating a higher risk of flare in patients who are 1) lowering the dose of a DMARD versus continuing DMARDs at the same dose, and 2) abruptly versus gradually discontinuing a DMARD (32–36). The recommendations are conditional because patient and physician preferences are expected to vary.

Gradual discontinuation of sulfasalazine is conditionally recommended over gradual discontinuation of hydroxychloroquine for patients taking triple therapy who wish to discontinue a DMARD

Gradually discontinuing sulfasalazine is recommended because of its poorer treatment persistence due to adverse events (14). The recommendation is conditional because patient and physician preferences are expected to vary.

Gradual discontinuation of methotrexate is conditionally recommended over gradual discontinuation of the bDMARD or tsDMARD for patients taking methotrexate plus a bDMARD or tsDMARD who wish to discontinue a DMARD

In the absence of direct evidence, gradually discontinuing methotrexate is preferred because a bDMARD or tsDMARD is typically added following an inadequate response to methotrexate. Thus, the continued use of the bDMARD or tsDMARD is more likely to maintain disease control than the continued use of methotrexate. The recommendation is conditional because gradual discontinuation of the bDMARD or tsDMARD may be favored depending on comorbidities, risk for infection, cost concerns, as well as patient and clinician preferences. The voting panel cautioned that many patients treated with certain monoclonal antibodies may require ongoing treatment with methotrexate to prevent the formation of antidrug antibodies (37).

Recommendations for specific patient populations (Table 6)

Subcutaneous nodules

Methotrexate is conditionally recommended over alternative DMARDs for patients with subcutaneous nodules who have moderate-to-high disease activity

Switching to a non-methotrexate DMARD is conditionally recommended over continuation of methotrexate for patients taking methotrexate with progressive subcutaneous nodules

While accelerated nodulosis has been observed in patients starting methotrexate (38), there are no studies examining comparative strategies for patients with stable or progressive subcutaneous nodules. The preceding 2 recommendations are conditional because patient and clinician preferences are expected to vary. The recommendation to switch is based on the premise that methotrexate is a contributing factor to progressive nodulosis.

Pulmonary disease

Methotrexate is conditionally recommended over alternative DMARDs for the treatment of inflammatory arthritis for patients with clinically diagnosed mild and stable airway or parenchymal lung disease, or incidental disease detected on imaging, who have moderate-to-high disease activity

Studies indicate that preexisting lung disease is a risk factor for methotrexate-related pneumonitis (39,40). However, the overall risk of worsening lung disease attributable to methotrexate is uncertain, and alternative DMARDs have also been associated with lung disease (41–45). The recommendation is in favor of methotrexate because of its important role as an anchor treatment in RA and the lack of alternatives with similar efficacy and/or superior long-term safety profiles. The recommendation is conditional because some clinicians (rheumatologists and pulmonologists) and patients will prefer an alternative option rather than accept any additional risk of lung toxicity. Patients with preexisting lung disease should be informed of their increased risk of methotrexate pneumonitis prior to initiating treatment with methotrexate.

Table 6. Specific patient populations*

Recommendations	Certainty of evidence	Based on the evidence report(s) of the following PICO(s)	Evidence table(s), in Supp. App. 2
Subcutaneous nodules			
Methotrexate is **conditionally** recommended over alternative DMARDs for patients with subcutaneous nodules who have moderate-to-high disease activity.	Very low	PICO 64	p. 427
Switching to a non-methotrexate DMARD is **conditionally** recommended over continuation of methotrexate for patients taking methotrexate with progressive subcutaneous nodules.	Very low	PICO 65	p. 428
Pulmonary disease			
Methotrexate is **conditionally** recommended over alternative DMARDs for the treatment of inflammatory arthritis for patients with clinically diagnosed mild and stable airway or parenchymal lung disease who have moderate-to-high disease activity.	Very low	PICO 67	p. 430
Heart failure			
Addition of a non–TNF inhibitor bDMARD or tsDMARD is **conditionally** recommended over addition of a TNF inhibitor for patients with NYHA class III or IV heart failure and an inadequate response to csDMARDs.	Very low	PICO 70	p. 435
Switching to a non–TNF inhibitor bDMARD or tsDMARD is **conditionally** recommended over continuation of a TNF inhibitor for patients taking a TNF inhibitor who develop heart failure.	Very low	PICO 71	p. 436
Lymphoproliferative disorder			
Rituximab is **conditionally** recommended over other DMARDs for patients who have a previous lymphoproliferative disorder for which rituximab is an approved treatment and who have moderate-to-high disease activity.	Very low	PICO 75 and PICO 76	p. 446–7
Hepatitis B infection			
Prophylactic antiviral therapy is **strongly** recommended over frequent monitoring alone for patients initiating rituximab who are hepatitis B core antibody positive (regardless of hepatitis B surface antigen status).	Very low	PICO 82	p. 459
Prophylactic antiviral therapy is **strongly** recommended over frequent monitoring alone for patients initiating any bDMARD or tsDMARD who are hepatitis B core antibody positive and hepatitis B surface antigen positive.	Very low	PICO 83	p. 464
Frequent monitoring alone is **conditionally** recommended over prophylactic antiviral therapy for patients initiating a bDMARD other than rituximab or a tsDMARD who are hepatitis B core antibody positive and hepatitis B surface antigen negative.	Very low	PICO 84	p. 471
Nonalcoholic fatty liver disease			
Methotrexate is **conditionally** recommended over alternative DMARDs for DMARD-naive patients with nonalcoholic fatty liver disease, normal liver enzymes and liver function tests, and no evidence of advanced liver fibrosis who have moderate-to high disease activity.	Very low	PICO 87	p. 489
Persistent hypogammaglobulinemia without infection			
In the setting of persistent hypogammaglobulinemia without infection, continuation of rituximab therapy for patients at target is **conditionally** recommended over switching to a different bDMARD or tsDMARD.	Very low	PICO 66	p. 429
Previous serious infection			
Addition of csDMARDs is **conditionally** recommended over addition of a bDMARD or tsDMARD for patients with a serious infection within the previous 12 months who have moderate-to-high disease activity despite csDMARD monotherapy.	Very low	PICO 88	p. 490
Addition of/switching to DMARDs is **conditionally** recommended over initiation/dose escalation of glucocorticoids for patients with a serious infection within the previous 12 months who have moderate-to-high disease activity.	Very low	PICO 90 and PICO 91	p. 496–7
Nontuberculous mycobacterial lung disease			
Use of the lowest possible dose of glucocorticoids (discontinuation if possible) is **conditionally** recommended over continuation of glucocorticoids for patients with nontuberculous mycobacterial lung disease.	Very low	No relevant PICO	
Addition of csDMARDs is **conditionally** recommended over addition of a bDMARD or tsDMARD for patients with nontuberculous mycobacterial lung disease who have moderate-to-high disease activity despite csDMARD monotherapy.	Very low	PICO 92	p. 498
Abatacept is **conditionally** recommended over other bDMARDs and tsDMARDs for patients with nontuberculous mycobacterial lung disease who have moderate-to-high disease activity despite csDMARDs.	Very low	PICO 93	p. 499

* PICO = population, intervention, comparator, and outcomes; Supp. App. 2 = Supplementary Appendix 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24596/abstract; DMARDs = disease-modifying antirheumatic drugs; TNF = tumor necrosis factor; bDMARD = biologic DMARD; tsDMARD = targeted synthetic DMARD; NYHA = New York Heart Association; csDMARDs = conventional synthetic DMARDs.

Table 7. Key clinical questions requiring further research*

Methotrexate administration
At what dose and route of administration should methotrexate be started?
Does switching to non-methotrexate DMARDs improve tolerability over increasing the dose of folic acid, or using folinic acid or using split dose or subcutaneous dosing, for RA patients with side effects when taking methotrexate?
TTT
What is the efficacy of TTT in different patient populations (early versus late, bDMARD- or tsDMARD-exposed, elderly-onset, comorbidities)?
What is the optimal target and method of assessment of disease activity for TTT in different populations?
Comparative effectiveness/safety
What is the comparative effectiveness/safety between bDMARDs and tsDMARDs?
What is the comparative effectiveness/safety between adding bDMARDs or tsDMARDs to methotrexate and switching to bDMARD or tsDMARD monotherapy?
What is the comparative effectiveness/safety between TTT by maximizing use of methotrexate (i.e., escalating dose via subcutaneous route) and adding/switching to bDMARD or tsDMARD monotherapy?
When, which, and how should DMARDs be tapered/discontinued?
Do clinical or biologic markers predict a differential response to DMARDs?
Comorbidities
What is the effectiveness/safety of alternative treatment strategies in RA patients with clinical lung disease or NAFLD?
Which DMARDs can be initiated or continued after receiving checkpoint inhibitor therapy?
Which DMARDs should be used in patients with solid malignancies, including skin cancer?
Is there a time frame before which DMARDs can be started/resumed in patients with concomitant solid malignancies?

* DMARDs = biologic disease-modifying antirheumatic drugs; RA = rheumatoid arthritis; TTT = treat-to-target; bDMARD = biologic DMARD; tsDMARD = targeted synthetic DMARD; NAFLD = nonalcoholic fatty liver disease.

Heart failure

Addition of a non–TNF inhibitor bDMARD or tsDMARD is conditionally recommended over addition of a TNF inhibitor for patients with New York Heart Association (NYHA) class III or IV heart failure and an inadequate response to csDMARDs

Switching to a non–TNF inhibitor bDMARD or tsDMARD is conditionally recommended over continuation of a TNF inhibitor for patients taking a TNF inhibitor who develop heart failure

These recommendations are based on the risk of worsening heart failure observed in randomized clinical trials of TNF inhibitors in patients with NYHA class III or IV heart failure without RA (46,47). Both recommendations are conditional because of the very low-certainty evidence supporting these PICO questions.

Lymphoproliferative disorder

Rituximab is conditionally recommended over other DMARDs for patients who have a previous lymphoproliferative disorder for which rituximab is an approved treatment and who have moderate-to-high disease activity

Rituximab is preferred over other DMARDs, regardless of previous DMARD experience, because it would not be expected to increase the risk of recurrence or worsening of these lymphoproliferative disorders. The recommendation is conditional because of the very low-certainty evidence supporting this PICO question.

Hepatitis B infection

Prophylactic antiviral therapy is strongly recommended over frequent monitoring of viral load and liver enzymes alone for patients initiating rituximab who are hepatitis B core antibody positive (regardless of hepatitis B surface antigen status)

Prophylactic antiviral therapy is strongly recommended over frequent monitoring alone for patients initiating any bDMARD or tsDMARD who are hepatitis B core antibody positive and hepatitis B surface antigen positive

Frequent monitoring alone of viral load and liver enzymes is conditionally recommended over prophylactic antiviral therapy for patients initiating a bDMARD other than rituximab or a tsDMARD who are hepatitis B core antibody positive and hepatitis B surface antigen negative

These recommendations were made based on the risk of hepatitis B reactivation due to core antibody and surface antigen status and the specific DMARD being initiated and are consistent with the updated American Association for the Study of Liver Diseases guidance (48). Patients at risk for hepatitis B reactivation should be comanaged with a hepatologist. The third recommendation is conditional because it is less certain whether the benefit of prophylactic antiviral therapy outweighs the risks and cost of this treatment in the specified patient population.

Nonalcoholic fatty liver disease (NAFLD)

Methotrexate is conditionally recommended over alternative DMARDs for DMARD-naive patients with NAFLD, normal liver enzymes and liver function tests, and no evidence of advanced liver fibrosis who have moderate-to-high disease activity

Given the concerns about the risk of hepatotoxicity associated with methotrexate therapy in patients with NAFLD, use of methotrexate should be restricted to patients with normal liver enzymes and liver function tests and without evidence of liver disease or liver fibrosis (Stage 3 or 4). Noninvasive testing to diagnose and stage liver fibrosis as well as consultation with a gastroenterologist or hepatologist should be considered in patients prior to initiating methotrexate (49). In addition, more frequent monitoring should be performed in this patient population (every 4 to 8 weeks). The recommendation is conditional because patients' and clinicians' risk tolerance varies.

Persistent hypogammaglobulinemia without infection

In the setting of persistent hypogammaglobulinemia without infection, continuation of rituximab therapy for patients at target is conditionally recommended over switching to a different bDMARD or tsDMARD

Continuing rituximab in patients who are at target is preferred because of the uncertain clinical significance of hypogammaglobulinemia in patients without infection. Although an increased risk of infection has been described in RA patients with hypogammaglobulinemia, it is not known if a switch in DMARDs in patients who are at target is more effective in lowering infection risk while maintaining disease control than continuation of rituximab. The recommendation is conditional because physician and patient risk tolerance is likely to vary depending on the degree of hypogammaglobulinemia and patient-specific risk factors for infection.

Previous serious infection

Addition of csDMARDs is conditionally recommended over addition of a bDMARD or tsDMARD for patients with a serious infection within the previous 12 months who have moderate-to-high disease activity despite csDMARD monotherapy

This conditional recommendation is made based on observational data suggesting a lower risk of infection associated with combination csDMARDs (dual or triple therapy) compared to bDMARDs or tsDMARDs (50). Some clinicians may prefer csDMARDs even if the serious infection occurred >12 months prior to considering a change.

Addition of/switching to DMARDs is conditionally recommended over initiation/dose escalation of glucocorticoids for patients with a serious infection within the previous 12 months who have moderate-to-high disease activity

This conditional recommendation is made based on observational studies suggesting a strong association between dose and duration of glucocorticoids with the risk of serious infection (51–53).

Nontuberculous mycobacterial (NTM) lung disease

Given the variability of NTM lung disease severity and response to treatment, patients should be closely comanaged with an infectious disease or pulmonary specialist.

Use of the lowest possible dose of glucocorticoids (discontinuation if possible) is conditionally recommended over continuation of glucocorticoids without dose modification for patients with NTM lung disease

This recommendation is based on studies suggesting an increased risk of NTM lung disease in patients receiving either inhaled or oral glucocorticoids (54,55).

Addition of csDMARDs is conditionally recommended over addition of a bDMARD or tsDMARD for patients with NTM lung disease who have moderate-to-high disease activity despite csDMARD monotherapy

This recommendation is based on the lower expected risk of NTM lung disease associated with csDMARDs compared to bDMARDs and tsDMARDs (56).

Abatacept is conditionally recommended over other bDMARDs and tsDMARDs for patients with NTM lung disease who have moderate-to-high disease activity despite csDMARDs

Abatacept is conditionally recommended over other bDMARDs and tsDMARDs based on population data extrapolated from studies on tuberculosis (57). There is considerable uncertainty regarding the risk of mycobacterial infections associated with non–TNF inhibitor bDMARDs and tsDMARDs; however, TNF inhibitors are associated with increased rates of mycobacterial infections and should be avoided (58).

The preceding 3 recommendations are conditional because of the very low-certainty evidence supporting the analysis of the differences in treatment outcomes posed by these PICO questions.

DISCUSSION

The ACR guidelines were developed to provide clinicians with recommendations for decisions frequently faced in clinical practice. Several new topics are included in this update, including recommendations for administration of methotrexate, use of methotrexate in patients with subcutaneous nodules, pulmonary disease, and NAFLD, use of rituximab in patients with hypogammaglobulinemia, and treatment of RA in patients with NTM lung disease. Areas covered in the 2015 guidelines that are not covered in this update include recommendations for patients with hepatitis C and solid malignancies. The panel did not vote on specific recommendations for patients with hepatitis C because curative antiviral therapy is now widely available. The panel did deliberate over PICO questions related to use of DMARDs in patients with solid malignancies. However, given the changing landscape of personalized treatments for many solid malignancies, the voting panel felt that a generalized recommendation was not possible.

On February 4, 2021, the FDA released a Drug Safety Alert noting a possible increased risk of major cardiovascular events and malignancies (excluding non-melanoma skin cancer) in patients with RA (over the age of 50 years with at least 1 risk factor for cardiovascular disease) participating in a randomized controlled trial designed to compare the safety of tofacitinib to adalimumab (18). Recommendations will be reviewed once peer-reviewed results are published. Rapidly evolving comparative effectiveness and safety signals associated with JAK inhibitors highlight the need to engage in a shared decision-making process when adjusting DMARDs (16,59). In addition, although previous recommendations cautioned against the use of TNF inhibitors in patients with skin cancer (1), the results of more recently published studies examining specific DMARD-related risks of non-melanoma skin cancer and melanoma do not support making a definite recommendation for or against specific DMARDs (60,61).

The panel also considered PICO questions related to current use of checkpoint inhibitor therapy, but the variability in current practice patterns and differences in treatment for specific cancer types precluded the development of specific recommendations for patients who are candidates for, or are currently receiving checkpoint inhibitor therapy. We anticipate that additional recommendations for patients with systemic rheumatic diseases and solid malignancies will be developed as further data become available. There were vigorous discussions pertaining to recommendations for specific DMARDs in patients with moderate-to-high disease activity despite csDMARDs and with a history of serious infection. However, the evidence was insufficient to support a recommendation. Future studies (using large registries and network meta-analyses) are needed to support specific recommendations for this patient population.

The recommendation statements in this update are not directly comparable to the ACR 2015 guidelines (1) because they do not retain the early versus established RA subgroups. Nevertheless, there are some notable differences. First, the 2015 guidelines recommend csDMARD monotherapy, preferably with methotrexate, for patients with both low and moderate/high disease activity, whereas this update recommends an initial trial of hydroxychloroquine or sulfasalazine for those with low disease activity. Second, the 2015 guidelines recommended DMARD tapering for patients who are in remission. In this update, tapering recommendations are made for patients who are in low disease activity or remission in the face of a paucity of data about when and how best to taper. The panel recommended that careful tapering might be considered if the patient wishes to cut back on their use of DMARDs. However, patients should be closely evaluated during any taper, and if a flare occurs, the prior regimen should be reinstituted promptly. Last, this update includes several recommendations *against* the use of glucocorticoid therapy. These recommendations were made in recognition of the frequent difficulty tapering glucocorticoids leading to undesirable prolonged use and the increasing evidence of the negative impact of glucocorticoids on long-term patient outcomes, including risk for infection, osteoporosis, and cardiovascular disease, in RA and other rheumatic diseases (62–65).

While consensus was easily reached on the majority of statements, 2 issues required prolonged discussion and debate. The decision on whether patients with an inadequate response to methotrexate should escalate to a bDMARD, tsDMARD, or triple therapy engendered much discussion with contrasting points of view. In the end, a recommendation was made in favor of a bDMARD or tsDMARD because of the more rapid onset of benefit and concerns related to the poor tolerability and durability of triple therapy in real-world practice (13,14). In particular, the patient panel highlighted the importance of a rapid onset of benefit after already having had an inadequate response to methotrexate. The conditional recommendation to initiate methotrexate therapy for patients with preexisting mild, stable lung disease was also rigorously debated. While minimizing the risk of toxicity is paramount, the voting panel favored a conditional recommendation to initiate methotrexate therapy in this clinical setting because of the vital role of this DMARD in the overall treatment of RA and lack of other comparable therapies without pulmonary risks.

Members of the voting panel agreed with the patient panel on the direction and strength of all but 2 recommendations. Patients were in favor of initial treatment with combination csDMARDs over methotrexate monotherapy because they placed greater value on the incremental benefits associated with combination therapy compared to clinicians. This preference was also stated in the 2015 guidelines (66). Patients also strongly preferred

discontinuing over a dose reduction of a DMARD whenever possible, whereas most clinicians on the voting panel preferred dose reduction. This discordance reflects patient preference to minimize use of medications once they reach target versus physician preference to minimize flare. However, both the patient and voting panel stressed the variability in patient preferences for tapering. These differences reinforce the importance of using a shared decision-making approach in RA.

When clinically relevant, recommendations specify the level of disease activity in the patient population (Table 1). However, evidence tables include pooled data from studies that often use different measures of disease activity; thus, specific definitions of low versus moderate-to-high disease activity are not provided for specific recommendations. Despite the large body of literature related to pharmacologic treatments for RA, the review team did not identify high-certainty evidence for many of the questions addressed. This discrepancy is due to the differences between clinically important PICO questions and the specific objectives of clinical trials. For example, few studies have examined how to best dose and administer methotrexate, the most effective and safe use of DMARDs in high-risk populations, and the risk–benefit tradeoffs associated with glucocorticoid use. Moreover, many trials could not be matched to specific PICO questions because of differences between the trials and the PICO questions' specified study populations and treatment comparisons. Thus, many recommendations are based largely on very low-certainty or low-certainty evidence. Incorporating medical evidence and expert input and consensus into clinical guidelines is core to the GRADE process and strengthens recommendations, particularly when there is limited evidence. Important gaps in knowledge are described in Table 7.

In summary, this update includes recommendations related to initiation and adjustment of DMARD therapy in patients with RA. It also emphasizes the importance of minimizing use of glucocorticoids. It is expected that additional data may modify the direction and/or strength of specific recommendations. The ACR will update the recommendations and answer these and other questions as new data are published.

ACKNOWLEDGMENTS

We thank Stanley B. Cohen, MD, for his contributions during the scoping phase of the project. We thank Cassandra Calabrese, DO, Leonard Calabrese, DO, Joseph Lim, MD, Danielle Antin-Ozerkis, MD, Marcus R. Pereira, MD, MPH, Jason E. Stout, MD, MHS, and Kevin Winthrop, MD, MPH, for their expert input as the voting panel considered recommendations related to comorbidities. We thank Nancy Baker, ScD, MPH, OTR/l, and Daniel White, PT, ScD, MSc, for their contributions during the scoping phase of the project, as well as their input during early deliberations by the voting panel. We thank the patients who (along with authors Kristine Carandang, PhD, OTR/L, and Shilpa Venkatachalam, PhD) participated in the patient panel meeting: Susan Campbell, Rachelle Crow-Hercher, MEd, Doug Davis, Sidney Harper, Faith Powell, Zenethia Roberts, Mary Turner, and C. Whitney White, PharmD. We thank the ACR staff, including Regina Parker for assistance in organizing the face-to-face meeting and coordinating the administrative aspects of the project, Robin Lane for assistance in manuscript preparation, and Jocelyn Givens for assistance with recruitment for the patient panel. We thank Janet Waters and Janet Joyce for their assistance in developing the literature search strategy, as well as performing the initial literature search and update searches.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Fraenkel, Bathon, England, St.Clair, Deane, Genovese, Kerr, Kremer, Sparks, Venkatachalam, Weinblatt, George, Johnson, Turner, Yaacoub, Akl.

Acquisition of data. Fraenkel, Bathon, England, St.Clair, Carandang, Deane, Genovese, Kerr, Kremer, J. Singh, Sparks, Al-Gibbawi, Baker, Barton, Cappelli, George, Johnson, Kahale, Karam, Khamis, Navarro-Millán, Mirza, Schwab, N. Singh, Turgunbaev, Turner, Yaacoub, Akl.

Analysis and interpretation of data. Fraenkel, Bathon, England, St.Clair, Arayssi, Deane, Genovese, Huston, Kerr, Kremer, Nakamura, Russell, J. Singh, Smith, Sparks, Venkatachalam, Weinblatt, Al-Gibbawi, Barbour, Barton, Chamseddine, Johnson, Kahale, Karam, Khamis, Navarro-Millán, Mirza, Schwab, N. Singh, Turgunbaev, Yaacoub, Akl.

ADDITIONAL DISCLOSURES

Author Genovese was employed by Stanford University Medical Center during development of this guideline but at the time of publication will also be employed by Gilead Sciences. Gilead Sciences had no financial or other interest in this project, had no input in the design, content, data collection, or analysis, and had no role in the writing or approval of this article.

REFERENCES

1. Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheumatol 2016;68:1–26.
2. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR, et al. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. Arthritis Rheum 2008;59:762–84.
3. Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, Kremer JM, et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64:625–39.
4. Goodman SM, Springer B, Guyatt G, Abdel MP, Dasa V, George M, et al. 2017 American College of Rheumatology/American Association of Hip and Knee Surgeons guideline for the perioperative management of antirheumatic medication in patients with rheumatic diseases undergoing elective total hip or total knee arthroplasty. Arthritis Rheumatol 2017;69:1538–51.
5. Sammaritano LR, Bermas BL, Chakravarty EE, Chambers C, Clowse ME, Lockshin MD, et al. 2020 American College of Rheumatology guideline for the management of reproductive health in rheumatic and musculoskeletal diseases. Arthritis Care Res (Hoboken) 2020; 72:461–88.
6. Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going from evidence to recommendation—determinants of a recommendation's direction and strength. J Clin Epidemiol 2013;66:726–35.

7. Alexander PE, Gionfriddo MR, Li SA, Bero L, Stoltzfus RJ, Neumann I, et al. A number of factors explain why WHO guideline developers make strong recommendations inconsistent with GRADE guidance. J Clin Epidemiol 2016;70:111–22.
8. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924–6.
9. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ 2010;182:E839–42.
10. England BR, Tiong BK, Bergman MJ, Curtis JR, Kazi S, Mikuls TR, et al. 2019 Update of the American College of Rheumatology recommended rheumatoid arthritis disease activity measures. Arthritis Care Res (Hoboken) 2019;71:1540–55.
11. Albrecht K, Zink A. Poor prognostic factors guiding treatment decisions in rheumatoid arthritis patients: a review of data from randomized clinical trials and cohort studies. Arthritis Res Ther 2017; 19:68.
12. Burgers LE, Raza K, van der Helm-van Mil AH. Window of opportunity in rheumatoid arthritis—definitions and supporting evidence: from old to new perspectives. RMD Open 2019;5:e000870.
13. Curtis JR, Palmer JL, Reed GW, Greenberg J, Pappas DA, Harrold LR, et al. Real-world outcomes associated with triple therapy versus tumor necrosis factor inhibitor/methotrexate therapy. Arthritis Care Res (Hoboken) doi: http://onlinelibrary.wiley.com/doi/10.1002/acr.24253/abstract. E-pub ahead of print.
14. Erhardt DP, Cannon GW, Teng CC, Mikuls TR, Curtis JR, Sauer BC. Low persistence rates in patients with rheumatoid arthritis treated with triple therapy and adverse drug events associated with sulfasalazine. Arthritis Care Res (Hoboken) 2019;71:1326–35.
15. Burmester GR, Rigby WF, van Vollenhoven RF, Kay J, Rubbert-Roth A, Kelman A, et al. Tocilizumab in early progressive rheumatoid arthritis: FUNCTION, a randomised controlled trial. Ann Rheum Dis 2016;75:1081–91.
16. Van Vollenhoven R, Takeuchi T, Pangan AL, Friedman A, Mohamed ME, Chen S, et al. Efficacy and safety of upadacitinib monotherapy in methotrexate-naive patients with moderately-to-severely active rheumatoid arthritis (SELECT-EARLY): a multicenter, multi-country, randomized, double-blind, active comparator–controlled trial. Arthritis Rheumatol 2020;72:1607–20.
17. Pfizer. Pfizer shares co-primary endpoint results from post-marketing required safety study of Xeljanz (Tofacitinib) in subjects with rheumatoid arthritis (RA). URL: https://investors.pfizer.com/investor-news/press-release-details/2021/Pfizer-Shares-Co-Primary-Endpoint-Results-from-Post-Marketing-Required-Safety-Study-of-XELJANZ-tofacitinib-in-Subjects-with-Rheumatoid-Arthritis-RA/default.aspx.
18. US Food and Drug Administration. Xeljanz, Xeljanz XR (tofacitinib): drug safety communication—initial safety trial results find increased risk of serious heart-related problems and cancer with arthritis and ulcerative colitis medicine. URL: https://www.fda.gov/safety/medical-product-safety-information/xeljanz-xeljanz-xr-tofacitinib-drug-safety-communication-initial-safety-trial-results-find-increased?utm_medium=email&utm_source=govdelivery.
19. Katchamart W, Trudeau J, Phumethum V, Bombardier C. Efficacy and toxicity of methotrexate (MTX) monotherapy versus MTX combination therapy with non-biological disease-modifying antirheumatic drugs in rheumatoid arthritis: a systematic review and meta-analysis. Ann Rheum Dis 2009;68:1105–12.
20. Emery P, Bingham CO, Burmester GR, Bykerk VP, Furst DE, Mariette X, et al. Certolizumab pegol in combination with dose-optimised methotrexate in DMARD-naïve patients with early, active rheumatoid arthritis with poor prognostic factors: 1-year results from C-EARLY, a randomised, double-blind, placebo-controlled phase III study. Ann Rheum Dis 2017;76:96–104.
21. Detert J, Bastian H, Listing J, Weiss A, Wassenberg S, Liebhaber A, et al. Induction therapy with adalimumab plus methotrexate for 24 weeks followed by methotrexate monotherapy up to week 48 versus methotrexate therapy alone for DMARD-naive patients with early rheumatoid arthritis: HIT HARD, an investigator-initiated study. Ann Rheum Dis 2013;72:844–50.
22. Nam JL, Villeneuve E, Hensor EM, Wakefield RJ, Conaghan PG, Green MJ, et al. A randomised controlled trial of etanercept and methotrexate to induce remission in early inflammatory arthritis: the EMPIRE trial. Ann Rheum Dis 2014;73:1027–36.
23. Schiff MH, Jaffe JS, Freundlich B. Head-to-head, randomised, crossover study of oral versus subcutaneous methotrexate in patients with rheumatoid arthritis: drug-exposure limitations of oral methotrexate at doses ≥15 mg may be overcome with subcutaneous administration. Ann Rheum Dis 2014;73:1549–51.
24. Visser K, van der Heijde D. Optimal dosage and route of administration of methotrexate in rheumatoid arthritis: a systematic review of the literature. Ann Rheum Dis 2009;68:1094–9.
25. Scott IC, Ibrahim F, Panayi G, Cope AP, Garrood T, Vincent A, et al. The frequency of remission and low disease activity in patients with rheumatoid arthritis, and their ability to identify people with low disability and normal quality of life. Semin Arthritis Rheum 2019; 49:20–6.
26. Bansback N, Phibbs CS, Sun H, O'Dell JR, Brophy M, Keystone EC, et al. Triple therapy versus biologic therapy for active rheumatoid arthritis: a cost-effectiveness analysis. Ann Intern Med 2017; 167:8–16.
27. Moreland LW, O'Dell JR, Paulus HE, Curtis JR, Bathon JM, St.Clair EW, et al. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the Treatment of Early Aggressive Rheumatoid Arthritis Trial. Arthritis Rheum 2012;64:2824–35.
28. O'Dell JR, Mikuls TR, Taylor TH, Ahluwalia V, Brophy M, Warren SR, et al. Therapies for active rheumatoid arthritis after methotrexate failure. N Engl J Med 2013;369:307–18.
29. Van Vollenhoven RF, Geborek P, Forslind K, Albertsson K, Ernestam S, Petersson IF, et al. Conventional combination treatment versus biological treatment in methotrexate-refractory early rheumatoid arthritis: 2 year follow-up of the randomised, non-blinded, parallel-group Swefot trial. Lancet 2012;379:1712–20.
30. Sauer BC, Teng CC, Tang D, Leng J, Curtis JR, Mikuls TR, et al. Persistence with conventional triple therapy versus a tumor necrosis factor inhibitor and methotrexate in US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2017;69:313–22.
31. Bergstra SA, Winchow LL, Murphy E, Chopra A, Salomon-Escoto K, Fonseca JE, et al. How to treat patients with rheumatoid arthritis when methotrexate has failed? The use of a multiple propensity score to adjust for confounding by indication in observational studies. Ann Rheum Dis 2019;78:25–30.
32. Pavelka K, Akkoç N, Al-Maini M, Zerbini CA, Karateev DE, Nasonov EL, et al. Maintenance of remission with combination etanercept-DMARD therapy versus DMARDs alone in active rheumatoid arthritis: results of an international treat-to-target study conducted in regions with limited biologic access. Rheumatol Int 2017;37:1469–79.
33. Ghiti Moghadam M, Vonkeman HE, ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, et al. Stopping tumor necrosis factor inhibitor treatment in patients with established rheumatoid arthritis in remission or with stable low disease activity: a pragmatic multicenter, open-label randomized controlled trial. Arthritis Rheumatol 2016;68:1810–7.

34. Van Herwaarden N, van der Maas A, Minten MJ, van den Hoogen FH, Kievit W, van Vollenhoven RF, et al. Disease activity guided dose reduction and withdrawal of adalimumab or etanercept compared with usual care in rheumatoid arthritis: open label, randomised controlled, non-inferiority trial. BMJ 2015;350:h1389.
35. Van Vollenhoven RF, Østergaard M, Leirisalo-Repo M, Uhlig T, Jansson M, Larsson E, et al. Full dose, reduced dose or discontinuation of etanercept in rheumatoid arthritis. Ann Rheum Dis 2016; 75:52–8.
36. Weinblatt ME, Bingham CO III, Burmester GR, Bykerk VP, Furst DE, Mariette X, et al. A phase III study evaluating continuation, tapering, and withdrawal of certolizumab pegol after one year of therapy in patients with early rheumatoid arthritis. Arthritis Rheumatol 2017;69:1937–48.
37. Strand V, Balsa A, Al-Saleh J, Barile-Fabris L, Horiuchi T, Takeuchi T, et al. Immunogenicity of biologics in chronic inflammatory diseases: a systematic review. BioDrugs 2017;31:299–316.
38. Patatanian E, Thompson DF. A review of methotrexate-induced accelerated nodulosis. Pharmacotherapy 2002;22:1157–62.
39. Alarcon GS, Kremer JM, Macaluso M, Weinblatt ME, Cannon GW, Palmer WR, et al. Risk factors for methotrexate-induced lung injury in patients with rheumatoid arthritis: a multicenter, case-control study. Ann Intern Med 1997;127:356–64.
40. Bartram SA. Experience with methotrexate-associated pneumonitis in northeastern England: comment on the article by Kremer et al [letter]. Arthritis Rheum 1998;41:1327–8.
41. Kawashiri SY, Kawakami A, Sakamoto N, Ishimatsu Y, Eguchi K. A fatal case of acute exacerbation of interstitial lung disease in a patient with rheumatoid arthritis during treatment with tocilizumab. Rheumatol Int 2012;32:4023–6.
42. Ostor AJ, Crisp AJ, Somerville MF, Scott DG. Fatal exacerbation of rheumatoid arthritis associated fibrosing alveolitis in patients given infliximab. BMJ 2004;329:1266.
43. Schoe A, van der Laan-Baalbergen NE, Huizinga TW, Breedveld FC, van Laar JM. Pulmonary fibrosis in a patient with rheumatoid arthritis treated with adalimumab. Arthritis Rheum 2006;55:157–9.
44. Taki H, Kawagishi Y, Shinoda K, Hounoki H, Ogawa R, Sugiyama E, et al. Interstitial pneumonitis associated with infliximab therapy without methotrexate treatment. Rheumatol Int 2009;30:275–6.
45. Roubille C, Haraoui B. Interstitial lung diseases induced or exacerbated by DMARDS and biologic agents in rheumatoid arthritis: a systematic literature review. Semin Arthritis Rheum 2014; 43:613–26.
46. Chung ES, Packer M, Lo KH, Fasanmade AA, Willerson JT. Randomized, double-blind, placebo-controlled, pilot trial of infliximab, a chimeric monoclonal antibody to tumor necrosis factor-alpha, in patients with moderate-to-severe heart failure: results of the anti-TNF Therapy Against Congestive Heart Failure (ATTACH) trial. Circulation 2003;107:3133–40.
47. Mann DL, McMurray JJ, Packer M, Swedberg K, Borer JS, Colucci WS, et al. Targeted anticytokine therapy in patients with chronic heart failure: results of the Randomized Etanercept Worldwide Evaluation (RENEWAL). Circulation 2004;109:1594–602.
48. Terrault NA, Lok ASF, McMahon BJ, Chang KM, Hwang JP, Jonas MM, et al. Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. Hepatology 2018;67:1560–99.
49. American Association for the Study of Liver Diseases. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. Clin Liver Dis (Hoboken) 2018;11:81.
50. Ozen G, Pedro S, England BR, Mehta B, Wolfe F, Michaud K. Risk of serious infection in patients with rheumatoid arthritis treated with biologic versus nonbiologic disease-modifying antirheumatic drugs. ACR Open Rheumatol 2019;1:424–32.
51. Lacaille D, Guh DP, Abrahamowicz M, Anis AH, Esdaile JM. Use of nonbiologic disease-modifying antirheumatic drugs and risk of infection in patients with rheumatoid arthritis. Arthritis Rheum 2008; 59:1074–81.
52. Dixon WG, Abrahamowicz M, Beauchamp ME, Ray DW, Bernatsky S, Suissa S, et al. Immediate and delayed impact of oral glucocorticoid therapy on risk of serious infection in older patients with rheumatoid arthritis: a nested case–control analysis. Ann Rheum Dis 2012;71:1128–33.
53. George MD, Baker JF, Winthrop K, Alemao E, Chen L, Connolly S, et al. Risk of biologics and glucocorticoids in patients with rheumatoid arthritis undergoing arthroplasty: a cohort study. Ann Intern Med 2019;170:825–36.
54. Liu VX, Winthrop KL, Lu Y, Sharifi H, Nasiri HU, Ruoss SJ. Association between inhaled corticosteroid use and pulmonary nontuberculous mycobacterial infection. Ann Am Thorac Soc 2018;15:1169–76.
55. Liao TL, Lin CF, Chen YM, Liu HJ, Chen DY. Risk factors and outcomes of nontuberculous mycobacterial disease among rheumatoid arthritis patients: a case-control study in a TB endemic area. Sci Rep 2016;6:29443.
56. Brode SK, Jamieson FB, Ng R, Campitelli MA, Kwong JC, Paterson JM, et al. Increased risk of mycobacterial infections associated with anti-rheumatic medications. Thorax 2015;70:677–82.
57. Cantini F, Niccoli L, Goletti D. Tuberculosis risk in patients treated with non-anti-tumor necrosis factor-α (TNF-α) targeted biologics and recently licensed TNF-α inhibitors: data from clinical trials and national registries. J Rheumatol Suppl 2014;91:56–64.
58. Winthrop KL, Baxter R, Liu L, Varley CD, Curtis JR, Baddley JW, et al. Mycobacterial diseases and antitumour necrosis factor therapy in USA. Ann Rheum Dis 2013;72:37–42.
59. Rubbert-Roth A, Enejosa J, Pangan AL, Haraoui B, Rischmueller M, Khan N, et al. Trial of upadacitinib or abatacept in rheumatoid arthritis. N Engl J Med 2020;383:1511–21.
60. Buchbinder R, Barber M, Heuzenroeder L, Wluka AE, Giles G, Hall S, et al. Incidence of melanoma and other malignancies among rheumatoid arthritis patients treated with methotrexate. Arthritis Rheum 2008;59:794–9.
61. Mercer LK, Askling J, Raaschou P, Dixon WG, Dreyer L, Hetland ML, et al. Risk of invasive melanoma in patients with rheumatoid arthritis treated with biologics: results from a collaborative project of 11 European biologic registers. Ann Rheum Dis 2017;76:386–91.
62. Del Rincón I, Battafarano DF, Restrepo JF, Erikson JM, Escalante A. Glucocorticoid dose thresholds associated with all-cause and cardiovascular mortality in rheumatoid arthritis. Arthritis Rheumatol 2014;66:264–72.
63. Bijlsma JW, Buttgereit F. Adverse events of glucocorticoids during treatment of rheumatoid arthritis: lessons from cohort and registry studies. Rheumatology (Oxford) 2016;55 Suppl 2:ii3–5.
64. Curtis JR, Westfall AO, Allison J, Bijlsma JW, Freeman A, George V, et al. Population-based assessment of adverse events associated with long-term glucocorticoid use. Arthritis Rheum 2006;55: 420–6.
65. Huscher D, Thiele K, Gromnica-Ihle E, Hein G, Demary W, Dreher R, et al. Dose-related patterns of glucocorticoid-induced side effects. Ann Rheum Dis 2009;68:1119–24.
66. Fraenkel L, Miller AS, Clayton K, Crow-Hercher R, Hazel S, Johnson B, et al. When patients write the guidelines: patient panel recommendations for the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:26–35.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 940–946
DOI 10.1002/acr.24537

BRIEF REPORT

Juvenile Spondyloarthritis in the Childhood Arthritis and Rheumatology Research Alliance Registry: High Biologic Use, Low Prevalence of HLA–B27, and Equal Sex Representation in Sacroiliitis

Dax G. Rumsey,[1] Aimee Lougee,[2] Roland Matsouaka,[2] David H. Collier,[3] Laura E. Schanberg,[2] Jennifer Schenfeld,[3] Natalie J. Shiff,[4] Matthew L. Stoll,[5] Scott Stryker,[3] Pamela F. Weiss,[6] and Timothy Beukelman,[5] for the Childhood Arthritis and Rheumatology Research Alliance Registry Investigators

Objective. To describe characteristics of children with enthesitis-related arthritis (ERA) and juvenile psoriatic arthritis (PsA) who were enrolled in the Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry.

Methods. All children with ERA and those with juvenile PsA were identified. Demographic characteristics, clinical characteristics, and treatments were described. The children with sacroiliitis and those without sacroiliitis were compared. In the children with sacroiliitis, the first visit with clinically active sacroiliitis (which came first in 72% of cases) was compared to the first visit without clinically active sacroiliitis.

Results. A total of 902 children with ERA or juvenile PsA were identified. Children with ERA were older at diagnosis (ages 10.8 years versus 8.2 years; $P < 0.01$) and were more likely to be male (56% versus 38%; $P < 0.01$). Polyarticular involvement was reported in 57% of children with ERA and in 72% of those with juvenile PsA. Of the children tested, HLA–B27 was positive in 38% of those in the ERA group and in 12% of those in the juvenile PsA group. At least 1 biologic was taken by 72% of those with ERA and 64% of those with juvenile PsA.

Sacroiliitis (diagnosed clinically and/or by imaging) was reported in 28% of the children (40% of those with ERA and 12% of those with juvenile PsA). Of these, 54% of the children were female, 36% were HLA–B27 positive, and 81% took at least 1 biologic. In children with sacroiliitis, scores according to the physician global assessment of disease activity, parent/patient global assessment of well-being, and clinical Juvenile Arthritis Disease Activity Score 10 were all significantly worse at the first visit with clinically active sacroiliitis versus the first visit without active sacroiliitis.

Conclusion. In this registry, there are more than 900 children with ERA or juvenile PsA. There was high biologic use in this population, especially in those with sacroiliitis. Further, there was equal sex representation in those children with sacroiliitis.

INTRODUCTION

Juvenile spondyloarthritis (SpA) is an umbrella term for a group of related conditions in childhood characterized by arthritis, enthesitis, increased risk of axial disease, and association with HLA–B27 positivity (1).

The most widely used classification system for juvenile arthritis at present, the International League of Associations for Rheumatology (ILAR) system, is not designed to properly capture these conditions (2–4). Under ILAR classification, there are 7 categories of juvenile idiopathic arthritis (JIA). The 2 categories that encompass most of the children with spondyloarthritis are enthesitis-related

The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the agencies or organizations with which the authors are affiliated. The Childhood Arthritis and Rheumatology Research Alliance (CARRA) is a North American research organization incorporated in the US as a 501c3 nonprofit. It receives funding from multiple sources, primarily the Arthritis Foundation, as well as research grants from the NIH and the Patient-Centered Outcomes Research Institute, as well as advocacy organizations and industry. The CARRA Registry is sponsored by CARRA and receives funds from Novartis and Roche to support the pharmacosurveillance program. The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the agencies or organizations with which the authors are affiliated.

Supported by Amgen.

Drs. Collier and Stryker and Ms Schenfeld have ownership or partnership and stock options or bond holdings in Amgen. Dr. Schanberg has received research grants and/or consulting fees, speaking fees, or honoraria from Sobi, UCB, Aurania, Bristol Myers Squibb, and Data Safety Monitoring Board Sanofi

SIGNIFICANCE & INNOVATIONS

- This study describes many children with enthesitis-related arthritis and juvenile psoriatic arthritis, followed from June 2015 to June 2018 in the Childhood Arthritis and Rheumatology Research Alliance registry.
- There was high biologic use among these children, with the majority (69%) having been treated with at least 1 biologic. The rate was particularly high in children with sacroiliitis (81%).
- There was lower than expected prevalence of HLA–B27 positivity in this study population, including in those with sacroiliitis (36%).
- There was equal sex representation in children with sacroiliitis, highlighting the need for increased awareness of axial disease in female children, just as the awareness of spondyloarthritis in women has increased.

arthritis (ERA) and juvenile psoriatic arthritis (PsA). An uncertain number of children with ERA or juvenile PsA–like disease are classified in the undifferentiated category, but that category is heterogenous and not limited to those with this type of arthritis (5).

It has been well described that the 2 distinct groups within the juvenile PsA category are 1) younger children who meet the juvenile PsA criteria but resemble children with antinuclear antibody–positive oligoarthritis and 2) older children who meet the juvenile PsA criteria but who have a presentation more consistent with an adult with spondyloarthritis (e.g., more male individuals, more enthesitis) (6).

Children with ERA are more likely to have a clinical picture with predominantly peripheral arthritis, typically described as an oligoarthritis involving the lower extremities with high risk of axial disease, relative to the other categories of JIA (1). The prevalence of HLA–B27 is lower in children with ERA than in adults with spondyloarthritis. Classically, more male individuals have been found to have ERA than female individuals (1).

In the present study, we describe the characteristics of children with a physician-assigned diagnosis of ERA or juvenile PsA who were enrolled in the Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry (see Appendix A for CARRA registry site principal investigators, subinvestigators, and research coordinators) from June 2015 to June 2018. This study represents the largest cohort of children with ERA and juvenile PsA described to date. The extent to which these conditions have been described has been limited (6,7). Children with ERA are reported to have higher pain intensity and poorer health status in comparison to children with other categories of JIA (8). Thus, improved epidemiologic data is needed to help physicians better recognize and treat these children.

PATIENTS AND METHODS

The general methods of the CARRA registry have been described in detail previously (9). Briefly, it is a registry of children with rheumatic disease, including JIA, that started in 2015. At the inception of the registry, there was selective enrollment of children most likely to be treated with biologics. Observational data are collected retrospectively at enrollment and then prospectively. Data collected includes physician-assigned ILAR category, medications, clinical features, laboratory data, and imaging results. Study visits occur approximately every 6 months, in the context of routine clinical care. Additionally, detailed information is recorded whenever a patient starts a new JIA medication (9). A composite measure of disease activity is collected, including the clinical Juvenile Arthritis Disease Activity Score 10 (JADAS-10), which includes the physician global assessment of disease activity (PGA measured on a 21-point scale from 0 to 10, marked in increments of 0.5), the parent/patient global assessment of well-being (measured on an 11-point scale from 0 to 10), and the active joint count to a maximum of 10 joints (after a standard assessment, typically of 71 joints) (10). Additional measures are captured, including the Childhood Health Assessment Questionnaire and level of pain intensity (11). This analysis includes data from 60 American and 3 Canadian centers that were collected from June 2015 to June 2018.

Patients. All children with the physician-assigned diagnosis of either ERA or juvenile PsA in the registry were included in the analysis. If the JIA category changed over time (e.g., if a patient with oligoarthritis later developed psoriasis and was reassigned to having juvenile PsA), the most recent diagnosis was used. Prior to July 2017, there was preferential enrollment of children with polyarticular involvement and/or receiving biologics. After that date, enrollment was opened to all children with JIA.

Statistical analysis. Statistical analyses were performed using SAS, version 9.4. The characteristics of children with ERA or juvenile PsA, and the combined group (ERA and juvenile PsA) were described using mean ± SDs or medians and interquartile ranges

(less than $10,000 each). Dr. Shiff owns stock or stock options in AbbVie, Novartis, and Mylan NV. Dr. Weiss has received consulting fees from Pfizer and Lilly (less than $10,000 each). Dr. Beukelman has received consulting fees from Novartis and UCB (less than $10,000 each). No other disclosures relevant to this article were reported.

[1]Dax G. Rumsey, MD, MSc: University of Alberta, Edmonton, Alberta, Canada; [2]Aimee Lougee, MS, Roland Matsouaka, PhD, Laura E. Schanberg, MD: Duke University, Durham, North Carolina; [3]David H. Collier, MD, Jennifer Schenfeld, MPH, Scott Stryker, MD, DrPH: Amgen, Thousand Oaks, California; [4]Natalie J. Shiff, MD, MHSc: University of Saskatchewan, Saskatoon, Saskatchewan, Canada; [5]Matthew L. Stoll, MD, PhD, Timothy Beukelman, MD, MSCE: University of Alabama at Birmingham; [6]Pamela F. Weiss, MD, MSCE: University of Pennsylvania, Philadelphia.

Address correspondence to Dax G. Rumsey, MD, MSc, 3-502 ECHA, 11405 87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9. Email: rumsey@ualberta.ca.

Submitted for publication April 23, 2020; accepted in revised form December 8, 2020.

for continuous variables and absolute and relative frequencies (number and %) for categorical variables, as appropriate. *P* values for categorical variables were computed when at least 50% of the cell counts had >0 records. Pearson's chi-square test or Fisher's exact test was performed, depending on the expected cell counts. *P* values for comparisons of mean values were computed by either Wilcoxon's rank sum test of means or *t*-test, depending on whether the normality assumption was met (using the Shapiro-Wilk test).

Characteristics of children with ERA or juvenile PsA in whom sacroiliitis was ever reported were compared to those without current or historically reported sacroiliitis. Sacroiliitis was recorded as present clinically (as judged by the treating rheumatologist) and/or by imaging (magnetic resonance imaging [MRI] or computed tomography [CT] scan). The following instructions were provided about what constitutes clinically active sacroiliitis: "Indicate whether the subject has clinically active sacroiliitis upon current physical examination (at the time of the visit)." Clinical sacroiliitis was defined as the presence of ≥2 of the following: 1) tenderness of the sacroiliac (SI) joint on examination, 2) positive Flexion Abduction External Rotation (Patrick's) test, and 3) inflammatory back pain (the presence of ≥3 of the following conditions: insidious onset, improvement with exercise, no improvement with rest, and pain at night [with improvement upon getting up])." The form asked if there was "imaging evidence of sacroiliitis (synovitis, bone marrow edema) by MRI or CT."

Finally, select characteristics of children for whom sacroiliitis was ever reported were compared between the first observed visit with clinically active sacroiliitis (which came first in 72% of cases) and the first observed visit without clinically active sacroiliitis. We singled out active sacroiliitis to highlight its impact on children and their families.

RESULTS

At the time of data extraction (June 30, 2018), there were 5,641 children with JIA in the registry. A total of 902 of the children (16%) met the criteria for inclusion; 522 children with ERA (9%) and 380 children with juvenile PsA (7%). Of note, 496 patients (55%) were enrolled prior to July 2017 (when registry enrollment was biased toward children with polyarticular involvement and/or biologic treatment) and 406 (45%) were enrolled on or after that date (when the entry criteria opened to all children with JIA).

To assess the potential bias introduced by enrollment date, we compared the children enrolled prior to/on July 2017 to those enrolled after that date. We found that those enrolled before July 2017 had a longer mean disease duration (in days) at the time of last visit than those enrolled after on or after July 2017 (1,398 days versus 1,325 days; $P = 0.047$), as expected. The proportion of children with polyarticular involvement (ever reported) was higher in those enrolled in the first time period than those in the second (72% versus 51%; $P < 0.001$), as expected. However, the proportion of patients who received at least 1 biologic was similar between those who enrolled prior to or on July 2017 compared to those who enrolled after this date (71% versus 66%; $P = 0.083$).

Total group and ERA/juvenile PsA. The combined group (comprising children with ERA and those with juvenile PsA) was predominantly White (81%) and female (52%; 44% of children with ERA and 62% with juvenile PsA) ($P < 0.001$). The mean ± SD age at diagnosis for the combined group was 9.7 ± 4.2 years (10.8 years for ERA and 8.2 years for juvenile PsA) ($P < 0.001$). A majority of children had polyarticular involvement (ever reported); 63% of the combined group, 57% with ERA, and 72% with juvenile PsA ($P < 0.001$) (Table 1). Sacroiliitis was ever reported in 28% of the combined group (40% with ERA and 12% with juvenile PsA) ($P < 0.001$). Of those with sacroiliitis, the diagnosis was clinical in 95 of 252 children (38%), while in the remaining children the diagnosis involved MRI or CT. Enthesitis was ever reported in more than half (54%) of the combined group (78% with ERA and 19% with juvenile PsA). HLA–B27 was reported as present in 222 of 725 (31%) of those tested (38% in the ERA group and 12% in the juvenile PsA group) ($P < 0.001$). The majority of children (69%) had

Table 1. Characteristics of children with enthesitis-related arthritis (ERA) or psoriatic arthritis (PsA) by juvenile idiopathic arthritis (JIA) category at the most recently occurring visit*

Characteristic	Overall (n = 902)	ERA (n = 522)	PsA (n = 380)	*P*†
Age at diagnosis, mean ± SD years	9.7 ± 4.2	10.8 ± 3.4	8.2 ± 4.6	<0.001
Female sex	467 (52)	231 (44)	236 (62)	<0.001
Reported White race‡	731 (81)	417 (80)	314 (83)	0.30
Polyarticular involvement, no./total number (%)§	563/888 (63)	293/514 (57)	270/374 (72)	<0.001
Sacroiliitis, no./total number§	252 (28)	208 (40)	44 (12)	<0.001
Enthesitis, no./total number (%)§	476/876 (54)	407 (78)	69/354 (19)	<0.001
HLA–B27 present, no./total number (%)	222/725 (31)	197 (38)	25/203 (12)	<0.001
≥1 biologic taken	622 (69)	377 (72)	245 (64)	0.01
≥1 TNF inhibitor taken	609 (68)	372 (71)	237 (62)	0.005

* Values are the number (%) unless indicated otherwise. TNF = tumor necrosis factor.
† All *P* values are for comparisons between ERA and PsA.
‡ Approximately 8% responded with either "multiple races," "other," or "prefer not to answer."
§ Ever reported.

been treated with at least 1 biologic by the time of analysis (72% in the ERA group and 64% in the juvenile PsA group) ($P = 0.01$) (Table 1). The vast majority of children had been treated with at least 1 tumor necrosis factor (TNF) inhibitor (71% with ERA and 62% with juvenile PsA) (Table 1).

Children with versus without sacroiliitis. When comparing children with sacroiliitis to those without sacroiliitis (ever reported), several important results were observed. First, male sex was not associated with the presence of sacroiliitis. In fact, 54% of the children with sacroiliitis versus 50% without sacroiliitis (ever reported) were female ($P = 0.29$). (Nota Bene Imaging was only conducted in those in whom sacroiliitis was clinically suspected). Children with sacroiliitis (ever reported) were older at JIA diagnosis (mean ± SD age 11.1 ± 3.7 years versus 9.3 ± 4.2 years) ($P < 0.001$). Among those with reported test results, HLA–B27 was present in 86 of 239 (36%) of the children with sacroiliitis and 132 of 464 (28%) of those without sacroiliitis ($P = 0.04$). Most children with sacroiliitis (81%) had been receiving treatment with ≥1 biologic, compared to 65% of those without sacroiliitis ($P < 0.001$) (Table 2). The vast majority of children had been receiving treatment with ≥1 TNF inhibitor (79% of those with sacroiliitis and 64% of those without sacroiliitis) ($P < 0.001$) (Table 2).

Children with sacroiliitis: first visit with active sacroiliitis versus first visit without. When considering only children with ERA or juvenile PsA in whom sacroiliitis had ever been reported, there was a significantly higher mean ± SD PGA score at the first visit with active sacroiliitis (2.9 ± 2.2) versus their first visit without active sacroiliitis (1.8 ± 1.9) ($P < 0.001$). The mean ± SD parent/patient global assessment of disease activity score for these children at their first visit with active sacroiliitis (3.5 ± 2.5) was higher than at their first visit without sacroiliitis (2.9 ± 2.5) ($P = 0.023$). Further, the median active peripheral (non-SI) joint count was not statistically different at the visit with sacroiliitis (median 0.5 [interquartile range (IQR) 0.0–3.5]) versus the visit without sacroiliitis (median 0.0 [IQR 0.0–2.0]) ($P = 0.096$). The mean ± SD clinical JADAS-10 score was significantly higher at the first visit with active sacroiliitis (9.3 ± 6.2) versus the first visit without sacroiliitis (5.8 ± 5.0) ($P < 0.001$). There was no significant difference in any of the other measures (Table 3). To assess the potential confounder of increased treatment intensity, the proportion of patients taking biologic therapy at each of these visits was compared and found to be higher in those without active SI disease (Table 3). Further, the duration of biologic treatment at each visit was assessed and found to not differ (Table 3).

Table 2. Characteristics of children with enthesitis-related arthritis (ERA) or psoriatic arthritis (PsA) by sacroiliitis category (ever reported by clinical examination or imaging)*

Characteristic	Sacroiliitis ever reported†	Sacroiliitis never reported†	*P*
Age at diagnosis, mean ± SD years	11.1 ± 3.7	9.3 ± 4.2	<0.001
Female sex	137 (54)	312 (50)	0.29
Reported White race†	192 (76)	517 (84)	0.012
HLA–B27 present, no./total no. (%)†	86/239 (36)	132/464 (28)	0.04
≥1 biologic taken	203 (81)	403 (65)	<0.001
≥1 TNF inhibitor taken	198 (79)	396 (64)	<0.001

* Values are the number (%) unless indicated otherwise. For sacroiliitis ever reported, n = 252 (28%); for sacroiliitis never reported, n = 619 (69%). TNF = tumor necrosis factor.
† Sacroiliitis status was unknown for 31 patients (3%).

DISCUSSION

In this study, the characteristics of more than 900 children with the physician-assigned diagnosis of ERA or juvenile PsA in the CARRA registry were described. As expected, there were clear differences between children with ERA versus those with juvenile PsA. Children with ERA were older at diagnosis, more likely to be male, more likely to have had enthesitis (ever reported), and more likely to be HLA–B27 positive than those with juvenile PsA.

A majority of children in this cohort with either ERA (57%) or juvenile PsA (72%) had polyarticular involvement (ever reported), unlike traditional descriptions of children with these diagnoses (1). Biologic use was frequent in both groups, especially in children with sacroiliitis, which is reflective of recent trends in JIA treatment (12). It is not surprising that biologic use is particularly high in those with sacroiliitis, as conventional disease-modifying antirheumatic drugs are generally ineffective for axial disease (13).

The high prevalence of polyarticular involvement and high biologic use were likely influenced by registry entry criteria prior to July 2017, which favored enrollment of children with these characteristics. Since then, the registry has opened up to all children with JIA. As mentioned, 55% of our patients were enrolled prior to July 2017 and the remainder were enrolled on or after that date.

Interestingly, however, the proportion of children with polyarticular involvement in this cohort is similar to that reported in the large Canadian cohort of children with JIA (Research in Arthritis in Canadian, Emphasizing Outcomes (ReACCh-Out)]. In that registry, which was open to all children with JIA, 57% of children with enthesitis (regardless of JIA category) had polyarticular arthritis (ever reported) (7).

The prevalence of HLA–B27 in the current study was lower than prior estimates, including in those with sacroiliitis (36%). This may be partially due to the mix of children with juvenile PsA and ERA (as opposed to just ERA) and because some cases of sacroiliitis were diagnosed clinically (i.e., not confirmed with imaging).

Perhaps the most surprising finding of the current study was the approximately equal sex representation in children with

Table 3. Clinical characteristics at first visit with clinically active sacroiliitis versus at first visit without clinically active sacroiliitis among children with ever reported sacroiliitis*

Characteristic	First visit with clinically active sacroiliitis	No./total no. of children	First visit with no clinically active sacroiliitis	No./total no. of children	*P*
Taking biologics at first visit, no. (%)	91 (56)	–	108 (66)	–	0.010†
Duration of biologic treatment, median (IQR) days‡	268.0 (99.0–601.0)§	–	221.0 (127.0–565.0)¶	–	0.94
PGA	2.9 ± 2.2	154/163	1.8 ± 1.9	150/163	<0.001
Parent/patient global assessment	3.5 ± 2.5	132/163	2.9 ± 2.5		0.023
Active peripheral joint count, median (IQR)	0.5 (0.0–3.5)	160/163	0.0 (0.02.0)	162/163	0.096
Clinical JADAS-10#	9.3 ± 6.2	126/163	5.8 ± 5.0	115/163	<0.001
Physical function mobility	31.3 ± 8.0	46/163	29.8 ± 4.0	41/163	0.95
C-HAQ	0.5 ± 0.6	138/163	0.4 ± 0.5	116/163	0.09
Pain intensity	4.0 ± 2.5	111/163	3.6 ± 2.5	94/163	0.37
Pain interference	62.3 ± 6.2	88/163	61.5 ± 6.6	77/163	0.42

* Values are the mean ± SD, unless indicated otherwise. C-HAQ = Childhood Health Assessment; IQR = interquartile range; JADAS-10 = Juvenile Arthritis Disease Activity Score 10; PGA = physician global assessment of disease activity (measured on an 11-point scale from 0 to 10).
† The McNemar test was used for this comparison.
‡ Duration from biologic start to visit date (days).
§ N = 89.
¶ N = 106.
# Clinical JADAS-10 is a composite measure of disease activity, which includes the PGA, the parent/patient global assessment of well-being (measured on an 11-point scale from 0 to 10), and the active joint count to a maximum of 10 joints (after a standard assessment, typically of 71 joints).

sacroiliitis (54% female), diagnosed clinically and/or with imaging. It is surprising that either method would diagnose sacroiliitis as frequently in female children as in male children. In adults, the diagnosis of sacroiliitis in women is often delayed compared to men due in part to a lower clinical index of suspicion (14). Interestingly, many of the children identified with sacroiliitis may have a condition akin to the adult diagnosis of nonradiographic axial spondyloarthritis, which has an almost equal prevalence in men and women (14).

In children with sacroiliitis and either ERA or juvenile PsA, the initial visit with active sacroiliitis was compared to the first visit without active sacroiliitis. At the active visit, the PGA, parent/patient global assessment of disease activity, and the clinical JADAS-10 (which includes the 2 preceding items) were all significantly worse (15), confirming the clinical impression that active sacroiliitis significantly impacts children and their families. A potential confounder is that more patients were taking biologics at the inactive visit, which may have had positive effects beyond treating the SI arthritis.

One study limitation is the imperfect classification of ERA or juvenile PsA in the CARRA registry. First, the ILAR classification system does not adequately describe these forms of arthritis (2–4). Thus, patients in the "unclassified" category of JIA may be more accurately categorized as ERA or juvenile PsA under an improved classification system. Second, we relied on physician-assigned diagnoses and do not know how accurately the physicians applied the ILAR criteria. Relying on clinical diagnoses may have resulted in misclassification of patients but is more accurately reflective of clinical care.

Another limitation is how sacroiliitis was captured in the registry during the study period. Sacroiliitis was recorded as present based on the physician's clinical impression and/or imaging evidence of active sacroiliac arthritis. There is now an improved CARRA registry data collection form in place.

A third limitation is the changing criteria for entry into the CARRA registry over the study period. Half of the study patients entered when the enrollment criteria were stricter, biasing the patient population toward those with polyarticular disease and/or use of biologics, while the other half entered when the enrollment criteria were changed to include all children with JIA.

In the CARRA registry, there are currently more than 900 children with physician-diagnosed ERA or juvenile PsA. Children with ERA are phenotypically different from those with juvenile PsA, although biologic use (especially TNF inhibitors) is high in both groups, particularly in those with sacroiliitis. Further, there was equal sex representation in children with sacroiliitis, suggesting that a higher level of suspicion in girls for sacroiliac disease may improve health outcomes.

Clearly describing and defining ERA and juvenile PsA populations is the first step to understanding the pathophysiology and determining optimal treatment of these diseases. Hopefully, future classification criteria will more fully capture children with ERA and juvenile PsA. An improved case report form for sacroiliitis and other refinements to the registry will help make this an even more valuable resource for studying these children going forward.

ACKNOWLEDGMENTS

This work could not have been accomplished without the aid of the following organizations: The NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases and the Arthritis Foundation. The authors thank all participants and hospital sites that recruited patients for the CARRA registry and the CARRA registry site principal investigators, subinvestigators, and research coordinators.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Rumsey, Beukelman.

Acquisition of data. Schanberg, Weiss, Beukelman.

Analysis and interpretation of data. Rumsey, Lougee, Matsouaka, Collier, Schanberg, Schenfeld, Shiff, Stoll, Stryker.

ROLE OF THE STUDY SPONSOR

Amgen helped facilitate the study design and reviewed and approved the manuscript prior to submission. The authors independently collected the data, interpreted the results, and had the final decision to submit the manuscript for publication.

REFERENCES

1. Weiss PF, Colbert RA. Juvenile spondyloarthritis: a distinct form of juvenile arthritis. Pediatr Clin North Am 2018;65:675–90.
2. Guillaume-Czitrom S, Sibilia J, Nordal E. Growing up with chronic arthritis: the confusing matter of classification. RMD Open 2017;3:e000417.
3. Hofer M, Southwood TR. Classification of childhood arthritis. Best Pract Res Clin Rheum 2002;16:379–96.
4. Martini A. It is time to rethink juvenile idiopathic arthritis classification and nomenclature. Ann Rheum Dis 2012;71:1437–9.
5. Chan MO, Petty RE, Guzman J. A family history of psoriasis in a first-degree relative in children with JIA: to include or exclude? J Rheumatol 2016;43:944–7.
6. Zisman D, Gladman DD, Stoll ML, Strand V, Lavi I, Hsu JJ, et al. The juvenile psoriatic arthritis cohort in the CARRA registry: clinical characteristics, classification, and outcomes. J Rheumatol 2017; 44:342–51.
7. Rumsey DG, Guzman J, Rosenberg AM, Huber AM, Scuccimarri R, Shiff NJ, et al. Characteristics and course of enthesitis in a juvenile idiopathic arthritis inception cohort. Arthritis Care Res (Hoboken) 2018;70:303–8.
8. Weiss PF, Beukelman T, Schanberg LE, Kimura Y, Colbert RA, CARRA Registry Investigators. Enthesitis-related arthritis is associated with higher pain intensity and poorer health status in comparison with other categories of juvenile idiopathic arthritis: the Childhood Arthritis and Rheumatology Research Alliance Registry. J Rheumatol 2012; 39:2341–51.
9. Beukelman T, Kimura Y, Ilowite NT, Mieszkalski K, Natter MD, Burrell G, et al. The new Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry: design, rationale, and characteristics of patients enrolled in the first 12 months. Pediatr Rheumatol Online J 2017;15:30.
10. McErlane F, Beresford MW, Baildam EM, Chieng SE, Davidson JE, Foster HE, et al. Validity of a three-variable juvenile arthritis disease activity score in children with new-onset juvenile idiopathic arthritis. Ann Rheum Dis 2013;72:1983–8.
11. Klepper SE. Measures of pediatric function: the Child Health Assessment Questionnaire (CHAQ), Juvenile Arthritis Functional Assessment Report (JAFAR), Juvenile Arthritis Functional Assessment Scale (JAFAS), Juvenile Arthritis Functional Status Index (JASI), and Pediatric Orthopedic Surgeons of North America (POSNA) Pediatric Musculoskeletal Functional Health Questionnaire. Arthritis Rheum 2003;49:S5–S14.
12. Weiss PF, Xiao R, Brandon TG, Pagnini I, Wright TB, Beukelman T, et al. Comparative effectiveness of tumor necrosis factor agents and disease-modifying antirheumatic therapy in children with enthesitis-related arthritis: the first year after diagnosis. J Rheumatol 2018;45:107–14.
13. Ringold S, Angeles-Han ST, Beukelman T, Lovell D, Cuello CA, Becker ML, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the treatment of juvenile idiopathic arthritis: therapeutic approaches for non-systemic polyarthritis, sacroiliitis, and enthesitis. Arthritis Care Res (Hoboken) 2019;71:717–34.
14. Rusman T, van Vollenhoven RF, van der Horst-Bruinsma IE. Gender differences in axial spondyloarthritis: women are not so lucky. Curr Rheumatol Rep 2018;20:35.
15. Consolaro A, Negro G, Gallo MC, Bracciolini G, Ferrari C, Schiappapietra B, et al. Defining criteria for disease activity states in nonsystemic juvenile idiopathic arthritis based on a three-variable juvenile arthritis disease activity score. Arthritis Care Res (Hoboken) 2014;66:1703–9.

APPENDIX A: CARRA REGISTRY INVESTIGATORS, SUBINVESTIGATORS, AND COORDINATORS

Members of the CARRA registry site principal investigators, subinvestigators and research coordinators are as follows: N. Abel, K. Abulaban, A. Adams, M. Adams, R. Agbayani, J. Aiello, S. Akoghlanian, C. Alejandro, E. Allenspach, R. Alperin, M. Alpizar, G. Amarilyo, W. Ambler, E. Anderson, S. Ardoin, S. Armendariz, E. Baker, I. Balboni, S. Balevic, L. Ballenger, S. Ballinger, N. Balmuri, F. Barbar-Smiley, L. Barillas-Arias, M. Basiaga, K. Baszis, M. Becker, H. Bell-Brunson, E. Beltz, H. Benham, S. Benseler, W. Bernal, T. Bigley, B. Binstadt, C. Black, M. Blakley, J. Bohnsack, J. Boland, A. Boneparth, S. Bowman, C. Bracaglia, E. Brooks, M. Brothers, A. Brown, H. Brunner, M. Buckley, M. Buckley, H. Bukulmez, D. Bullock, B. Cameron, S. Canna, L. Cannon, P. Carper, V. Cartwright, E. Cassidy, L. Cerracchio, E. Chalom, J. Chang, A. Chang-Hoftman, V. Chauhan, P. Chira, T. Chinn, K. Chundru, H. Clairman, D. Co, A. Confair, H. Conlon, R. Connor, A. Cooper, J. Cooper, S. Cooper, C. Correll, R. Corvalan, D. Costanzo, R. Cron, L. Curiel-Duran, T. Curington, M. Curry, A. Dalrymple, A. Davis, C. Davis, C. Davis, T. Davis, F. De Benedetti, D. De Ranieri, J. Dean, F. Dedeoglu, M. DeGuzman, N. Delnay, V. Dempsey, E. DeSantis, T. Dickson, J. Dingle, B. Donaldson, E. Dorsey, S. Dover, J. Dowling, J. Drew, K. Driest, Q. Du, K. Duarte, D. Durkee, E. Duverger, J. Dvergsten, A. Eberhard, M. Eckert, K. Ede, B. Edelheit, C. Edens, C. Edens, Y. Edgerly, M. Elder, B. Ervin, S. Fadrhonc, C. Failing, D. Fair, M. Falcon, L. Favier, S. Federici, B. Feldman, J. Fennell, I. Ferguson, P. Ferguson, B. Ferreira, R. Ferrucho, K. Fields, T. Finkel, M. Fitzgerald, C. Fleming, O. Flynn, L. Fogel, E. Fox, M. Fox, L. Franco, M. Freeman, K. Fritz, S. Froese, R. Fuhlbrigge, J. Fuller, N. George, K. Gerhold, D. Gerstbacher, M. Gilbert, M. Gillispie-Taylor, E. Giverc, C. Godiwala, I. Goh, H. Goheer, D. Goldsmith, E. Gotschlich, A. Gotte, B. Gottlieb, C. Gracia, T. Graham, S. Grevich, T. Griffin, J. Griswold, A. Grom, M. Guevara, P. Guittar, M. Guzman, M. Hager, T. Hahn, O. Halyabar, E. Hammelev, M. Hance, A. Hanson, L. Harel, S. Haro, J. Harris, O. Harry, E. Hartigan, J. Hausmann, A. Hay, K. Hayward, J. Heiart, K. Hekl, L. Henderson, M. Henrickson, A. Hersh, K. Hickey, P. Hill, S. Hillyer, L. Hiraki, M. Hiskey, P. Hobday, C. Hoffart, M. Holland, M. Hollander, S. Hong, M. Horwitz, J. Hsu, A. Huber, J. Huggins, J. Hui-Yuen, C. Hung, J. Huntington, A. Huttenlocher, M. Ibarra, L. Imundo, C. Inman, A. Insalaco, A. Jackson, S. Jackson, K. James, G. Janow, J. Jaquith, S. Jared, N. Johnson, J. Jones, J. Jones, J. Jones, K. Jones, S. Jones, S. Joshi, L. Jung, C. Justice, A. Justiniano, N. Karan, K. Kaufman, A. Kemp, E. Kessler, U. Khalsa, B. Kienzle, S.

Kim, Y. Kimura, D. Kingsbury, M. Kitcharoensakkul, T. Klausmeier, K. Klein, M. Klein-Gitelman, B. Kompelien, A. Kosikowski, L. Kovalick, J. Kracker, S. Kramer, C. Kremer, J. Lai, J. Lam, B. Lang, S. Lapidus, B. Lapin, A. Lasky, D. Latham, E. Lawson, R. Laxer, P. Lee, P. Lee, T. Lee, L. Lentini, M. Lerman, D. Levy, S. Li, S. Lieberman, L. Lim, C. Lin, N. Ling, M. Lingis, M. Lo, D. Lovell, D. Lowman, N. Luca, S. Lvovich, C. Madison, J. Madison, S. Magni Manzoni, B. Malla, J. Maller, M. Malloy, M. Mannion, C. Manos, L. Marques, A. Martyniuk, T. Mason, S. Mathus, L. McAllister, K. McCarthy, K. McConnell, E. McCormick, D. McCurdy, P. McCurdy Stokes, S. McGuire, I. McHale, A. McMonagle, C. McMullen-Jackson, E. Meidan, E. Mellins, E. Mendoza, R. Mercado, A. Merritt, L. Michalowski, P. Miettunen, M. Miller, D. Milojevic, E. Mirizio, E. Misajon, M. Mitchell, R. Modica, S. Mohan, K. Moore, L. Moorthy, S. Morgan, E. Morgan Dewitt, C. Moss, T. Moussa, V. Mruk, A. Murphy, E. Muscal, R. Nadler, B. Nahal, K. Nanda, N. Nasah, L. Nassi, S. Nativ, M. Natter, J. Neely, B. Nelson, L. Newhall, L. Ng, J. Nicholas, R. Nicolai, P. Nigrovic, J. Nocton, B. Nolan, E. Oberle, B. Obispo, B. O'Brien, T. O'Brien, O. Okeke, M. Oliver, J. Olson, K. O'Neil, K. Onel, A. Orandi, M. Orlando, S. Osei-Onomah, R. Oz, E. Pagano, A. Paller, N. Pan, S. Panupattanapong, M. Pardeo, J. Paredes, A. Parsons, J. Patel, K. Pentakota, P. Pepmueller, T. Pfeiffer, K. Phillippi, D. Pires Marafon, K. Phillippi, L. Ponder, R. Pooni, S. Prahalad, S. Pratt, S. Protopapas, B. Puplava, J. Quach, M. Quinlan-Waters, C. Rabinovich, S. Radhakrishna, J. Rafko, J. Raisian, A. Rakestraw, C. Ramirez, E. Ramsay, S. Ramsey, R. Randell, A. Reed, A. Reed, A. Reed, H. Reid, K. Remmel, A. Repp, A. Reyes, A. Richmond, M. Riebschleger, S. Ringold, M. Riordan, M. Riskalla, M. Ritter, R. Rivas-Chacon, A. Robinson, E. Rodela, M. Rodriquez, K. Rojas, T. Ronis, M. Rosenkranz, B. Rosolowski, H. Rothermel, D. Rothman, E. Roth-Wojcicki, K. Rouster – Stevens, T. Rubinstein, N. Ruth, N. Saad, S. Sabbagh, E. Sacco, R. Sadun, C. Sandborg, A. Sanni, L. Santiago, A. Sarkissian, S. Savani, L. Scalzi, S. Scharnhorst, K. Schikler, A. Schlefman, H. Schmeling, K. Schmidt, E. Schmitt, R. Schneider, K. Schollaert-Fitch, G. Schulert, T. Seay, C. Seper, J. Shalen, R. Sheets, A. Shelly, S. Shenoi, K. Shergill, J. Shirley, M. Shishov, C. Shivers, E. Silverman, N. Singer, V. Sivaraman, J. Sletten, A. Smith, C. Smith, J. Smith, J. Smith, E. Smitherman, J. Soep, M. Son, S. Spence, L. Spiegel, J. Spitznagle, R. Sran, H. Srinivasalu, H. Stapp, K. Steigerwald, Y. Sterba Rakovchik, S. Stern, A. Stevens, B. Stevens, R. Stevenson, K. Stewart, C. Stingl, J. Stokes, E. Stringer, S. Sule, J. Sumner, R. Sundel, M. Sutter, R. Syed, G. Syverson, A. Szymanski, S. Taber, R. Tal, A. Tambralli, A. Taneja, T. Tanner, S. Tapani, G. Tarshish, S. Tarvin, L. Tate, A. Taxter, J. Taylor, M. Terry, M. Tesher, A. Thatayatikom, B. Thomas, K. Tiffany, T. Ting, A. Tipp, D. Toib, K. Torok, C. Toruner, H. Tory, M. Toth, S. Tse, V. Tubwell, M. Twilt, S. Uriguen, T. Valcarcel, H. Van Mater, L. Vannoy, C. Varghese, N. Vasquez, K. Vazzana, R. Vehe, K. Veiga, J. Velez, J. Verbsky, G. Vilar, N. Volpe, E. von Scheven, S. Vora, J. Wagner, L. Wagner-Weiner, D. Wahezi, H. Waite, J. Walker, H. Walters, T. Wampler Muskardin, L. Waqar, M. Waterfield, M. Watson, A. Watts, P. Weiser, J. Weiss, E. Wershba, A. White, C. Williams, A. Wise, J. Woo, L. Woolnough, T. Wright, E. Wu, A. Yalcindag, M. Yee, E. Yen, R. Yeung, K. Yomogida, Q. Yu, R. Zapata, A. Zartoshti, A. Zeft, R. Zeft, Y. Zhang, Y. Zhao, A. Zhu, and C. Zic.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 947–954
DOI 10.1002/acr.24204

American College of Rheumatology
Empowering Rheumatology Professionals

Common Functional Ability Score for Young People With Juvenile Idiopathic Arthritis

Stephanie J. W. Shoop-Worrall,[1] Martijn A. H. Oude Voshaar,[2] Janet E. McDonagh,[3] Mart A. F. J. Van de Laar,[2] Nico Wulffraat,[4] Wendy Thomson,[3] Kimme L. Hyrich,[3] and Suzanne M. M. Verstappen[3]

Objective. As young people enter adulthood, the interchangeable use of child and adult outcome measures may inaccurately capture changes over time. This study aimed to use item response theory (IRT) to model a continuous score for functional ability that can be used no matter which questionnaire is completed.

Methods. Adolescents (ages 11–17 years) in the UK Childhood Arthritis Prospective Study (CAPS) self-completed an adolescent Childhood Health Assessment Questionnaire (CHAQ) and a Health Assessment Questionnaire (HAQ). Their parents answered the proxy-completed CHAQ. Those children with at least 2 simultaneously completed questionnaires at initial presentation or 1 year were included. Psychometric properties of item responses within each questionnaire were tested using Mokken analyses to assess the applicability of IRT modeling. A previously developed IRT model from the Pharmachild-NL registry from The Netherlands was validated in CAPS participants. Agreement and correlations between IRT-scaled functional ability scores were tested using intraclass correlations and Wilcoxon's signed rank tests.

Results. In 303 adolescents, the median age at diagnosis was 13 years, and 61% were female. CHAQ scores consistently exceeded HAQ scores. Mokken analyses demonstrated high scalability, monotonicity, and the fact that each questionnaire yielded reliable scores. There was little difference in item response characteristics between adolescents enrolled in CAPS and Pharmachild-NL (maximum item residual 0.08). Significant differences were no longer evident between IRT-scaled HAQ and CHAQ scores.

Conclusion. IRT modeling allows the direct comparison of function scores regardless of different questionnaires being completed by different people over time. IRT modeling facilitates the ongoing assessment of function as adolescents transfer from pediatric clinics to adult services.

INTRODUCTION

Functional ability is an important patient-reported outcome in individuals with juvenile idiopathic arthritis (JIA), both in childhood and later life (1). As a young person with JIA moves through adolescence and into adulthood, their functional ability may be assessed using 1 of 3 versions of the Health Assessment Questionnaire (HAQ), depending on their age and local practice: the proxy-completed Childhood Health Assessment Questionnaire (P-CHAQ) (2), a self-completed adolescent CHAQ (A-CHAQ) with the same items as the P-CHAQ but developmentally appropriate rewording (3), or the self-completed Stanford HAQ, which has fewer items and was originally designed for adults with rheumatoid arthritis (4). The P-CHAQ was adapted from the HAQ and thus assesses similar domains of functional ability, with additional items for tasks more relevant to young people, e.g., writing with a pen/pencil. In addition, a modified HAQ (MHAQ) was developed from the HAQ to reduce the time burden for both patients and health care professionals. The MHAQ includes 1 question from each HAQ domain and can be completed in under 5 minutes by adults with rheumatoid arthritis (5).

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

Supported by the Medical Research Council (UK grant MR/K501311/1) and by the NIHR Manchester Biomedical Research Centre and the Versus Arthritis Centres for Excellence in Epidemiology and Genetics/Genomics (UK grants 20380 and 20542).

[1]Stephanie J. W. Shoop-Worrall, PhD: University of Manchester and Manchester Academic Health Sciences Centre, Manchester, UK; [2]Martijn A. H. Oude Voshaar, PhD, Mart A. F. J. Van de Laar, MD, PhD: University of Twente, Enschede, The Netherlands; [3]Janet E. McDonagh, MD, FRCP, Wendy Thomson, PhD, Kimme L. Hyrich, MD, PhD, FRCPC, Suzanne M. M. Verstappen, PhD: University of Manchester and Manchester Academic Health Sciences Centre, NIHR Manchester Biomedical Research Centre, and Manchester University Hospital NHS Foundation Trust, Manchester, UK; [4]Nico Wulffraat, MD, PhD: University Medical Center Utrecht and European Reference Network–RITA, Utrecht, The Netherlands.

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

Address correspondence to Stephanie J. W. Shoop-Worrall, PhD, 2.908 Stopford Building, The University of Manchester, Manchester, UK, M13 9PT. Email: Stephanie.shoop-worrall@manchester.ac.uk.

Submitted for publication December 13, 2019; accepted in revised form March 31, 2020.

SIGNIFICANCE & INNOVATIONS

- Functional ability is a key outcome for adolescents transitioning into adulthood.
- Using item response theory, a common scale for functional ability has been developed and validated.
- Direct comparison of functional ability through adolescence is now feasible using this common scale.

Directly comparing scores on these 4 similar outcome measures is challenging, since each questionnaire has unique questions, or items. This diversity may lead to differences in scores that are unrelated to actual differences in underlying functional ability (6). In addition, the questionnaires may have been completed by different people over time (e.g., adolescent, parent/care giver). Finally, questionnaires may contain missing values, especially when paper and pencil forms are used. These limitations hinder the continuous assessment of functional ability as adolescents mature and are transferred from pediatric to adult care, with previous research demonstrating that these existing questionnaires produce scores that are similar, but not interchangeable, when completed by/for the same young person at the same time point (6).

To continuously assess functional ability over time, a common function scale is needed. Using a single questionnaire for individuals with JIA of all ages would be difficult, since some functional tasks are age-specific and different people (care giver versus young person) may need to complete the questionnaire over time. One established method used to link scores from different questionnaires to a common scale is item response theory (IRT) (7,8). Within IRT, item and person characteristics are mapped on the same underlying measurement continuum. These characteristics encompass the trait level of the person completing the item (i.e., the level of their functional ability), and the characteristics of the items themselves (e.g., the general difficulty of opening a jar versus a car door). One useful benefit of modeling item responses this way is that the modeling allows the scores to be corrected for these item characteristics (9,10). This way, a single score can be reflective of underlying functional ability, no matter what questionnaires or items have been completed.

The applications of IRT models are increasingly popular in outcome assessments across various medical fields. For example, in the Patient-Reported Outcomes Measurement Information System (PROMIS [11]) project, various item banks have been developed, from which tailored questionnaires with different items and lengths can be derived, with optimal relevance to specific patients (12). In the patient-reported outcome Rosetta Stone (PROsetta Stone) project, IRT was one method used to link legacy measures, those already developed and historically used, with newer PROMIS measures, to allow the retrofitting of existing scores to the newer measures and vice versa (13). In addition, IRT has previously been used to model latent functional ability across multiple questionnaires in adults with rheumatoid arthritis (14). However, to date, its application in JIA, in addition to similar questionnaires that have been sequentially developed, is limited.

Recently, an IRT-based standardized functional ability reporting metric was developed in 16,386 people with inflammatory rheumatic diseases recruited to international registries, including the Pharmachild-NL registry of children and young people with JIA (14). The standardized functional ability scale developed includes 10 commonly used functional ability questionnaires (and their items), including the HAQ, MHAQ, and the P-CHAQ, and can be used to obtain comparable scores from each of the included questionnaires. It could therefore be used in young people with JIA to obtain comparable physical function scores regardless of the particular functional ability questionnaire used.

The aim of the current study was to examine 1) the applicability of this metric in JIA, which could be assessed by examining the assumptions and fit of the IRT model underlying the common metric, in data obtained in a population of adolescents with JIA in the UK; 2) the agreement between IRT-scaled scores obtained using P-CHAQ, A-CHAQ, and HAQ in adolescents with JIA; and 3) the measurement properties of these questionnaires in this population using nonparametric IRT analyses.

SUBJECTS AND METHODS

Development study population. The Pharmachild-NL registry is a web-based register extracting demographic and clinical data from medical records twice yearly for juvenile arthritis in Utrecht, The Netherlands. The cohort has been previously described (14). This cohort included 1,194 prevalent cases of juvenile arthritis who were prescribed methotrexate or biologic therapies and were selected for development of the IRT model. Item responses from the P-CHAQ, HAQ, and MHAQ were extracted from young people contributing these data between 2010 and 2017.

Validation study population. Data were obtained from adolescents enrolled in the Childhood Arthritis Prospective Study (CAPS). CAPS is a longitudinal, UK, multicenter inception cohort following children and young people with inflammatory arthritis with onset before their 16th birthday. Specific inclusion and exclusion criteria for CAPS have been described previously (15). CAPS has been approved by the Northwest Multicentre Research Ethics Committee (REC/02/8/104, IRAS 184042), and written informed consent was provided by proxies for all participants; where possible, patient assent was also obtained.

Between January 2004 and January 2015, adolescents ages 11–17 years who were enrolled in CAPS were asked to self-complete the A-CHAQ and HAQ and for their proxies to complete the P-CHAQ at the same clinic visit. Only those adolescents with data from at least 2 of these 3 questionnaires completed at either initial presentation to pediatric rheumatology (CAPS baseline) or at 1 year following the initial presentation (CAPS 1-year follow-up)

were included in the current analysis. MHAQ scores were calculated using existing HAQ scores where available, with 1 item from each domain included (16).

Additional data collected at baseline from the CAPS cohort included demographic (ethnicity, sex, date of birth, disease onset, and initial presentation) and disease-related variables collected at both baseline and 1 year (disease category, active joint count, limited joint count, erythrocyte sedimentation rate [mm/hour], physician's global assessment of disease [10-cm visual analog scale], and proxy global assessment of well-being [10-cm visual analog scale]).

Statistical analysis. *Calculating CHAQ/HAQ scores in CAPS data.* Item-specific, domain-specific, and overall CHAQ/HAQ scores were calculated using CAPS data at baseline and 1 year. Due to translation discordance between the UK and The Netherlands CHAQ versions, the UK item regarding running errands (Netherlands: run a race) was omitted. IRT models are robust to missing item data and overall scores can be compared using a total of the remaining items (14). To gain an overall score for each questionnaire, the largest possible item scores (0–3) within each domain (8 in total) were summed, for a possible range of 0–24. Dividing by 8 yields a final score ranging from 0 to 3 (increasing scores denote worsening disability). In cases of incomplete data, a final score can be calculated if at least 6 of 8 domains have values, through dividing by the number of domains with available data instead. In this study, the use of aids and devices was not considered when calculating domain-specific scores, in order to assess the effects of item differences on overall scores.

Assessing IRT assumptions in CAPS data. The IRT model that was used for calibrating the items from each questionnaire to a common function ability scale, the generalized partial credit model (17), has 2 assumptions: 1) unidimensionality: that all items from each functional ability questionnaire relate to the common underlying continuous function variable; and 2) monotonicity: that the expected item score functions are monotonically increasing over this latent variable (i.e., the common functional ability scale increases each time an item score increases). Both assumptions were tested by checking the goodness-of-fit of Mokken's model of monotone homogeneity (18). This is a nonparametric IRT model used to verify that patients can be ordered along an underlying latent variable. The model relies on the same assumptions as the generalized partial-credit model. In the Mokken approach, the unidimensionality assumption can be checked using item-level (H_i) and scale-level (H) scalability coefficients. Higher values indicate better scalability. H >0.30 supports unidimensionality and H >0.50 suggests a strong scale (19). The monotonicity assumption was checked using the check.monotonicity function of the Mokken R package. Subsequently, we examined the reliability of the overall scores for each questionnaire using the Molenaar Sijtsma coefficient.

Fitting the IRT model in CAPS data. Differences in item response behavior between adolescents enrolled in Pharmachild-NL (P-CHAQ) and CAPS (P-CHAQ, A-CHAQ, HAQ) were then examined to assess whether the existing item parameters were generalizable. This was completed by testing for differential item functioning (DIF). DIF occurs if adolescents with the same level of functional ability across cohorts have different IRT expected item scores. DIF was examined using Lagrange multiplier statistics and associated effect size statistics (20).

Subsequently, we fitted the previously estimated IRT model in the CAPS data. We tested the fit of the models by calculating the differences between the observed item scores in CAPS and the IRT model predicted scores (i.e., the absolute residuals). Item fit was considered acceptable if an item's score residual was less than ±0.2.

A test characteristic curve and conversion tables were constructed to demonstrate how raw CHAQ, HAQ, and MHAQ scores (as scored in this article with the 19-item HAQ and without the use of aids) can be compared with standardized functional ability scores and/or translated among each other. The conversion tables were constructed according to the expected a posteriori (EAP) approach of Thissen et al for summed scores, using the Lord Wingerky algorithm (21). These stand only where no missing data are evident. To gain more accurate comparisons to latent scores, the converter tool at http://tihealthcare.nl/en/expertise/common-metrics can be used, and an app is currently under development.

Evaluating congruence of IRT scores obtained from different functional ability questionnaires. Finally, the comparability of functional ability scores was assessed between IRT-scaled and raw CHAQ and HAQ scores. Pairwise agreement between EAP IRT scores from the 4 functional ability measures was assessed (22). The EAP score estimation procedure was chosen because of the sizable flooring effect of the CHAQ/HAQ. Pairwise agreements between overall raw scores and between EAP-modeled IRT scores at baseline were assessed using Bland-Altman plots and compared using Wilcoxon's signed rank tests. All analyses were undertaken in Stata software version 14, and R version 3.4.1.

RESULTS

Patient cohort. A total of 303 adolescents in CAPS had completed at least 2 of the 3 full questionnaires at either the baseline (n = 178) or 1 year visit (n = 231). Compared with those adolescents with fewer than 2 questionnaire responses at either time point (n = 77), those included in the study had marginally higher physician global scores (2.5 cm versus 3.1 cm; $P = 0.032$). There were no differences in age, sex, ethnicity, disease duration, International League of Associations for Rheumatology (ILAR) category, pain, or any of the JIA core outcome variables except physician's global scores at baseline between those included

and excluded from the study. Available CHAQ/HAQ scores were equivalent between the 2 groups.

The majority of study participants were female (59%) and of white ethnicity (91%). The median age at initial presentation to pediatric rheumatology was 13 years (interquartile range [IQR] 12–14) with median 7 months symptom duration to that point (IQR 4–17). The most common disease category was oligoarticular JIA (40%). At that time, adolescents had a median of 2 active joints and physician and proxy global scores at ~3 cm on a 10-cm visual analog scale (Table 1).

At baseline, median CHAQ scores were consistent across proxies and adolescents at both baseline (both CHAQ medians 0.6, both IQRs 0.1–1.3) and 1 year (both CHAQ medians 0.3, both IQRs 0.0–0.8). HAQ and MHAQ scores consistently ranked below those of the CHAQ: baseline HAQ 0.5 (IQR 0.0–1.3), 1-year HAQ 0.1 (IQR 0.0–0.8), baseline MHAQ 0.1 (IQR 0.0–0.5), 1-year MHAQ 0.0 (IQR 0.0–0.1) (Table 1).

Table 1. Baseline characteristics of the cohort (n = 303)*

Characteristic	Complete data, %	Value
Female sex, no. (%)	100	180 (59)
White or Caucasian, no. (%)	97	267 (91)
Age at onset, years	97	12 (11–13)
Age at first presentation, years	100	13 (12–14)
Symptom duration at first pediatric rheumatology appointment, months	98	7 (4–17)
ILAR category, no. (%)	100	
Systemic	–	20 (7)
Oligoarticular	–	120 (40)
RF– polyarticular	–	56 (18)
RF+ polyarticular	–	20 (7)
Enthesitis-related	–	30 (10)
Psoriatic	–	30 (10)
Undifferentiated	–	27 (9)
Core outcome variables at baseline		
Active joint count	90	2 (1–6)
Limited joint count	90	1 (1–4)
ESR, mm/hour	70	17 (6–54)
Physician's global assessment, cm	64	3.1 (1.7–5.4)
Proxy global assessment of well-being, cm	77	2.7 (0.7–5.1)
Functional ability at baseline†		
P-CHAQ	87	0.625 (0.125–1.250)
A-CHAQ	89	0.625 (0.125–1.250)
HAQ	87	0.500 (0.000–1.250)
MHAQ	87	0.125 (0.000–0.500)
Functional ability at 1 year†		
P-CHAQ	90	0.250 (0.000–0.750)
A-CHAQ	89	0.250 (0.000–0.750)
HAQ	93	0.125 (0.000–0.750)
MHAQ	93	0.000 (0.000–0.125)

* Values are the median (interquartile range) unless indicated otherwise. A-CHAQ = adolescent Childhood Health Assessment Questionnaire; ESR = erythrocyte sedimentation rate; HAQ = Health Assessment Questionnaire; ILAR = International League of Associations for Rheumatology; MHAQ = modified HAQ; P-CHAQ = proxy CHAQ; RF = rheumatoid factor.
† Of those patients who had ≥2 complete functional ability questionnaires at the time point.

The CAPS cohort was similar in sex, ethnicity, and ILAR distributions to the development population from the Pharmachild-NL registry (65% female, 96% white ethnicity, 48% oligoarthritis). Although Pharmachild-NL included prevalent cases, patient age at CHAQ/HAQ completion was comparable (mean ± SD 13 ± 7 years). Similar to the CAPS cohort, CHAQ scores (median 0.5 [IQR 0.1–1.0]) were higher than HAQ (median 0.4 [IQR 0.0–0.9]) and MHAQ scores (median 0.1 [IQR 0.0–0.5]).

Checking IRT assumptions and the psychometric properties of CHAQ/HAQ scores in CAPS. The IRT model assumptions held for each functional ability measure, suggesting that an IRT approach was applicable to functional ability in JIA using these questionnaires. Strong scalability and unidimensionality were evident for overall P-CHAQ, A-CHAQ, and HAQ scores at both baseline and 1 year (all H >0.5, all SE <0.1). Item-specific associations with the latent functional ability variable varied between items within questionnaires in terms of both scalability coefficients (H_i ranges: P-CHAQ 0.3–0.7, A-CHAQ 0.3–0.7, HAQ 0.4–0.7) and concordance coefficients (coefficient ranges: P-CHAQ 0.4–0.8, A-CHAQ 0.4–0.8, HAQ 0.5–0.8). There were no violations to monotonicity, and the reliability for each questionnaire at each time point was high (all reliability coefficients ≥0.95) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24204/abstract).

Assessing differences in item response behavior between CAPS and Pharmachild-NL and IRT model fit. The DIF analyses are summarized in Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24204/abstract, and suggested no great differences in how adolescents in CAPS and Pharmachild-NL responded to the items. In general, the observed HAQ, P-CHAQ, and A-CHAQ average item scores were similar to the average item scores predicted by a joint IRT calibration of the CAPS and Pharmachild-NL data, with all residuals <0.10, and only 1% of item residuals exceeding ±0.05 (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24204/abstract).

Subsequently, the fit of the item parameters calibrated in Oude Voshaar et al (14) were evaluated in CAPS data. Again, the model-predicted average item scores were generally close to the average item scores observed in the CAPS data, with residuals consistently falling below 0.2 across all questionnaires (see Supplementary Table 2).

Directly comparing latent functional ability across different questionnaires with different completers. Figure 1 shows how the CHAQ and HAQ scores relate to the standardized physical function score metric. In addition,

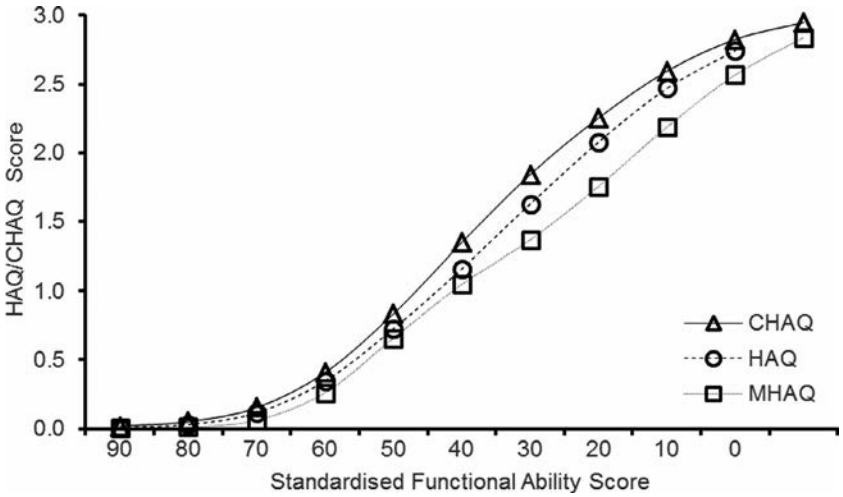

Figure 1. A test characteristic curve demonstrating how latent functional ability can be modeled using either/all of the Childhood Health Assessment Questionnaire (CHAQ), Health Assessment Questionnaire (HAQ), and modified HAQ (MHAQ) scores.

Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24204/abstract, allows the direct comparison of CHAQ, HAQ, and MHAQ scores to this score metric. Increasing values on the standardized function scores indicate better functional ability. The figure and conversion tables can be used to compare CHAQ scores to the standardized physical function scores and to retranslate to HAQ scores if needed. However, this exact relationship only applies where no missing values are evident.

Agreement between scores across modeling techniques. Bland-Altman plots demonstrated greater agreement between IRT-scaled than raw scores, demonstrated by narrower limits of agreement and greater centrality around a mean difference of zero for all pairs of scores (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24204/abstract). The majority of pairings had significant differences between raw scores and nonsignificant differences between IRT-scaled scores. In addition, T values were lower for all IRT-scaled pairings than raw scores, with the exception of the P-CHAQ versus A-CHAQ at baseline (Table 2).

DISCUSSION

Upon reaching adolescence and following transfer from pediatric to adult care, outcomes and adolescents with JIA are measured using self-completed questionnaires rather than via proxy reports. For functional ability, this often means the HAQ is used instead of the P-CHAQ, with the potential intermediate use of the A-CHAQ. Previous work has demonstrated high correlation but only moderate agreement between raw scores using these 3 measures (6,23,24). Therefore, assuming that the scores are interchangeable may result in the false assumption of an improvement in ability where no such change had occurred, based only on the choice of questionnaire. Similarly, longitudinal outcome studies in JIA that capture data across adolescence and young adulthood (25) may also make incorrect conclusions about functional ability over this period if the choice of measure is not considered. The current study demonstrated the applicability of IRT modeling using CHAQ/HAQ item responses. This could be used to understand functional ability in young people with JIA over longer periods of time, retrospectively scale functional ability scores from completed studies to increase standardized comparison, and allow for the interpretation of incomplete functional ability questionnaires.

Models initially developed in an international cohort including children and young people with JIA were validated in a UK multicenter inception cohort. This resulted in greater agreement between overall IRT-scaled scores than between raw scores. The IRT models presented therefore allow the direct comparison of P-CHAQ, A-CHAQ, HAQ, and/or MHAQ scores over time, with an underlying latent variable score and with each other. Further research using any of these measures in JIA should report scaled values alongside raw scores, to allow direct comparison of functional ability between cohorts that may have used different questionnaires.

The psychometric properties of CHAQ/HAQ/MHAQ scores in relation to IRT modeling have rarely been assessed. Previous smaller studies including prevalent cases of JIA have found estimating stable item parameters to be difficult (26,27). In both studies, small sample sizes, in addition to the prevalent flooring effect of the questionnaires, limited the accuracy of generated parametric-IRT (Rasch) parameters. One study resorted to combining the "with much difficulty" and "unable to do" CHAQ categories to force Rasch model fit (26). To overcome these issues, the current

Table 2. Significant differences between pairwise functional ability questionnaires*

	Baseline				One year			
Questionnaire comparison, model	No.	% ceiling†	T‡	*P*‡	No.	% ceiling†	T‡	*P*‡
P-CHAQ vs. A-CHAQ								
Raw data	136	19.3	1.3	0.196	183	41.3	0.6	0.580
IRT: EAP	136	–	1.5	0.138	183	–	0.2	0.843
P-CHAQ vs. HAQ								
Raw data	133	25.7	3.2	0.002	192	45.3	1.3	0.205
IRT: EAP	133	–	1.6	0.109	192	–	−0.2	0.851
P-CHAQ vs. MHAQ								
Raw data	133	23.1	8.7	<0.001	192	43.3	7.1	<0.001
IRT: EAP	133	–	1.9	0.059	192	–	0.8	0.425
A-CHAQ vs. HAQ								
Raw data	136	32.1	3.2	0.002	191	51.2	1.1	0.263
IRT: EAP	136	–	2.1	0.036	191	–	0.0	0.978
A-CHAQ vs. MHAQ								
Raw data	136	46.4	10.1	<0.001	191	61.7	7.1	<0.001
IRT: EAP	136	–	2.6	0.012	191	–	0.8	0.432
HAQ vs. MHAQ								
Raw data	156	24.3	9.9	<0.001	218	42.8	9.1	<0.001
IRT: EAP	156	–	1.0	0.340	218	–	2.0	0.052

* A-CHAQ = adolescent Childhood Health Assessment Questionnaire; EAP = expected a priori; HAQ = Health Assessment Questionnaire; IRT = item response theory; MHAQ = modified HAQ; P-CHAQ = proxy CHAQ.
† Percentage 0 on both scores.
‡ Wilcoxon's signed rank test.

study employed nonparametric IRT models in a population at least twice the sample size than in previous works. These models do not rely on estimated parameters to study the measurement properties of the included scales. Our results therefore provide useful additional information about the psychometric properties of the evaluated questionnaires. We were able to show that all items on the P-CHAQ, A-CHAQ, and HAQ relate to a single underlying functional ability variable and that each instrument yields highly reliable scores.

Once the applicability of IRT modeling to each of the 3 questionnaires had been confirmed, the current study was able to validate existing IRT models developed in young people and adults with JIA in the Pharmachild-NL registry. Previously fitted models successfully summarized the item responses given by adolescents in CAPS. Thus, the results should generalize across other cohorts of patients with JIA, regardless of which questionnaire has been completed. The utility of the models was demonstrated in the increased agreement between pairs of overall scores under these models compared to raw scores, with the former adjusting for item characteristics.

If complete data are available, the conversion table (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24204/abstract) and figure (Figure 1) can be used to access latent functional ability scores. In cases of missing data, or to convert entire data sets at once, the now externally validated models are available at http://tihealthcare.nl/en/expertise/common-metric and can be used to directly access latent functional ability scores for individual patients or cohorts of patients for both clinical and research purposes.

Limitations to the study include the small differences between CHAQ and HAQ items, few of which were entirely unique to each questionnaire. Despite the differences between questionnaire scores being greater than the minimum clinically important differences in functional ability (28,29), this analysis did not demonstrate the full possibilities of IRT modeling. Further applications include its ability to model other functional ability questionnaires with unique items, such as CHAQ compared with the functional ability questions within the Juvenile Arthritis Multidimensional Assessment Report (30). With increasing differences in questionnaires measuring the same disease construct, greater differences between raw scores and IRT-scaled scores would be evident. However, for this study, CHAQ and HAQ scores have been assumed interchangeable, and even with these small changes between questionnaire items, the current study was able to demonstrate 1) greater agreement between IRT-scaled compared with raw scores, 2) scores that are not biased in the presence of incomplete answers compared with raw scores, and 3) the ability to directly compare scores from any of the questionnaires with an underlying construct variable.

In clinical practice, these models facilitate direct comparison of CHAQ scores with HAQ scores upon switching of questionnaires during adolescence. This includes the MHAQ, with lesser burden on adolescents, since only 5 items on the HAQ are required for a total score, taking fewer than 5 minutes to complete (5), with young people previously reporting that the CHAQ was burdensome in length (31). Beyond this advantage, functional ability questionnaires can be tailored to each young person based on personalized relevance from a functional ability item bank such as PROMIS (11). IRT modeling would then allow

for the direct comparison of functional ability over time, even when different items have been completed from these different questionnaires.

Further limitations include the fact that functional ability of the tested cohort was, on average, low to moderate, and thus few very high CHAQ/HAQ scores contributed to the models. The flooring effect of these questionnaires is well known (2), with upper quartile scores extended to only 1.3 of 3.0 even at initial presentation to pediatric rheumatology. While few patients experienced severe limitations in functional ability, this validation cohort represents a generalizable sample of adolescents with newly diagnosed JIA, including those across all ILAR categories. Finally, the current study was able to demonstrate a direct comparison between latent functional ability and a proxy-completed P-CHAQ. However, it is often evident that young people with JIA complete the P-CHAQ themselves, particularly where the A-CHAQ and HAQ are not available. No adolescents in this study self-completed the P-CHAQ. However, the lack of differences in item responses between the proxy-completed P-CHAQ and adolescent-completed A-CHAQ meant that the current study could combine these questionnaires to a single CHAQ score. Thus, the CHAQ model presented should be able to adequately incorporate self-completed P-CHAQ scores. Finally, these data were collected as part of an observational real-world research study. As in any longitudinal observational study, clinical and demographic data are often missing. To allow for adequate validation of the IRT model, we required at least 2 of the CHAQ/HAQ forms to have been completed. Available CHAQ/HAQ scores were equivalent between adolescents included and excluded from the study.

P-CHAQ, A-CHAQ, and HAQ scores can be directly compared to latent functional ability using IRT modeling. This will greatly aid the direct comparison of functional ability across the JIA disease course when adolescents are transferred from pediatric to adult rheumatology services. In additional, scores from different study populations using different functional ability questionnaires can be directly compared, and longer-scale studies can now feasibly compare functional ability even if questionnaires have missing items and/or adolescents switch questionnaires throughout the study.

ACKNOWLEDGMENTS

The authors thank all of the children, young people, and their guardians involved in CAPS and the Pharmachild-NL registry, in addition to all clinical staff and administrators. We also thank the data management team at the University of Manchester, UK.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Shoop-Worrall, Oude Voshaar, McDonagh, Hyrich, Verstappen.

Acquisition of data. Oude Voshaar, McDonagh, Van de Laar, Wulffraat, Thomson, Hyrich, Verstappen.

Analysis and interpretation of data. Shoop-Worrall, Oude Voshaar, Hyrich, Verstappen.

REFERENCES

1. Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. Preliminary definition of improvement in juvenile arthritis. Arthritis Rheum 1997;40:1202–9.
2. Nugent J, Ruperto N, Grainger J, Machado C, Sawhney S, Baildam E, et al. The British version of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ). Clin Exp Rheumatol 2001;19:S163–7.
3. Shaw KL, Southwood TR, McDonagh JE. Growing up and moving on in rheumatology: parents as proxies of adolescents with juvenile idiopathic arthritis. Arthritis Rheum 2006;55:189–98.
4. Kirwan JR, Reeback JS. Stanford Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. Br J Rheumatol 1986;25:206–9.
5. Maska L, Anderson J, Michaud K. Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire disability index (HAQ), modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire II (HAQ-II), Improved Health Assessment Questionnaire (Improved HAQ), and Rheumatoid Arthritis Quality of Life (RAQoL). Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S4–13.
6. Shoop-Worrall SJ, Hyrich KL, Verstappen SM, Sergeant JC, Baildam E, Chieng A, et al. Comparing proxy, adolescent, and adult assessments of functional ability in adolescents with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2020;72:517–24.
7. Reise SP, Waller NG. Item response theory and clinical measurement. Annu Rev Clin Psychol 2009;5:27–48.
8. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. Med Care 2000;38:II28–II42.
9. Van der Linden WJ, Glas CA. Elements of adaptive testing. New York: Springer-Verlag; 2010.
10. Stocking ML, Frederic LM. Developing a common metric in item response theory. Appl Psychol Meas 1983;7:201–10.
11. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol 2010;63:1179–94.
12. Oude Voshaar MA, ten Klooster PM, Glas CA, Vonkeman HE, Taal E, Krishnan E, et al. Calibration of the PROMIS physical function item bank in Dutch patients with rheumatoid arthritis. PLos One 2014;9:e92367.
13. Choi SW, Schalet B, Cook KF, Cella D. Establishing a common metric for depressive symptoms: linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. Psychol Assess 2014;26:513–27.
14. Oude Voshaar MA, Vonkeman HE, Courvoisier D, Finckh A, Gossec L, Leung YY, et al. Towards standardized patient reported physical function outcome reporting: linking ten commonly used questionnaires to a common metric. Qual Life Res 2019;28:187–97.
15. Adib N, Hyrich K, Thornton J, Lunt M, Davidson J, Gardner-Medwin J, et al. Association between duration of symptoms and severity of disease at first presentation to paediatric rheumatology: results from the Childhood Arthritis Prospective Study. Rheumatology (Oxford) 2008;47:991–5.
16. Blalock SJ, Sauter SV, DeVellis RF. The modified Health Assessment Questionnaire difficulty scale: a health status measure revisited. Arthritis Rheumatol 1990;3:182–8.

17. Muraki E. A generalized partial credit model. In: ven der Linden WJ, Hambleton RK, editors. Handbook of modern item response theory. New York: Springer; 1997. p. 153–64.
18. Molenaar IW. Nonparametric models for polytomous responses. In: van der Linden WJ, Hambleton RK, editors. Handbook of modern item response theory. New York: Springer; 1997. p. 369–80.
19. Sijtsma K, van der Ark LA. A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. Br J Math Stat Psychol 2017;70:137–58.
20. Glas CA. Detection of differential item functioning using Lagrange Multiplier Tests. Stat Sin 1998;8:647–67.
21. Thissen D, Pommerich M, Billeaud K, Williams VS. Item response theory for scores on tests including polytomous items with ordered responses. Appl Psychol Meas 1995;19:39–49.
22. Warm TA. Weighted likelihood estimation of ability in item response theory. Psychometrika 1989;54:427–50.
23. Van Pelt PA, Kruize AA, Goren SS, Van Der Net J, Uiterwaal C, Kuis W, et al. Transition of rheumatologic care, from teenager to adult: which health assessment questionnaire can be best used? Clin Exp Rheumatol 2010;28:281–6.
24. Lal SD, McDonagh J, Baildam E, Wedderburn LR, Gardner-Medwin J, Foster HE, et al. Agreement between proxy and adolescent assessment of disability, pain, and well-being in juvenile idiopathic arthritis. J Pediatr 2011;158:307–12.
25. Gore FM, Bloem PJ, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10–24 years: a systematic analysis. Lancet 2011;377:18–24.
26. Pouchot J, Ecosse E, Coste J, Guillemin F. Validity of the Childhood Health Assessment Questionnaire is independent of age in juvenile idiopathic arthritis. Arthritis Rheum 2004;51:519–26.
27. Tennant A, Kearns S, Turner F, Wyatt S, Haigh R, Chamberlain MA. Measuring the function of children with juvenile arthritis. Rheumatology (Oxford) 2001;40:1274–8.
28. Brunner HI, Klein-Gitelman MS, Miller MJ, Barron A, Baldwin N, Trombley M, et al. Minimal clinically important differences of the Childhood Health Assessment Questionnaire. J Rheumatol 2005;32:150–61.
29. Pope JE, Khanna D, Norrie D, Ouimet JM. The minimally important difference for the Health Assessment Questionnaire in rheumatoid arthritis clinical practice is smaller than in randomized controlled trials. J Rheumatol 2009;36:254–9.
30. Filocamo G, Consolaro A, Schiappapietra B, Dalpra S, Lattanzi B, Magni-Manzoni S, et al. A new approach to clinical care of juvenile idiopathic arthritis: the Juvenile Arthritis Multidimensional Assessment Report. J Rheumatol 2011;38:938–53.
31. Parsons S, Thomson W, Cresswell K, Starling B, McDonagh JE. Barbara Ansell National Network for Adolescent Rheumatology (BANNAR). What do young people with rheumatic conditions in the UK think about research involvement? A qualitative study. Pediatr Rheumatol Online J 2018;16:35.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 955–963
DOI 10.1002/acr.24219

American College of Rheumatology
Empowering Rheumatology Professionals

Vertical Drop Jump Performance in Youth With Juvenile Idiopathic Arthritis

Gregor Kuntze,[1] Alberto Nettel-Aguirre,[2] Julia Brooks,[3] Shane Esau,[1] Colleen Nesbitt,[1] Dianne Mosher,[2] Marinka Twilt,[4] Susanne Benseler,[4] Janet L. Ronsky,[1] and Carolyn A. Emery[2]

Objective. Juvenile idiopathic arthritis (JIA) is associated with altered body structure and function outcomes that may expose youth with JIA to a greater risk of secondary joint injury. This study aimed to examine differences in vertical drop jump (VDJ) biomechanics for youth with JIA and healthy youth (control group).

Methods. The present study was a matched pair cohort study. Youth with JIA (n = 30) and their age- and sex-matched control peers participated in this ethics-approved study. Lower-extremity biomechanics information was obtained using a motion analysis system (Motion Analysis) and 2 force plates (AMTI). Biomechanics outcomes included hip, knee, and ankle joint angles, ground reaction forces (GRF), and VDJ phase durations. Other outcomes included disease activity, physical disability, and sports participation. Matched pairs data (JIA–control) were analyzed using a multivariate random coefficient model (version 3.5.0, R Core Team; joint angles, potential confounders) and paired samples *t*-tests with Bonferroni correction (α = 0.0125; GRF, VDJ phase durations).

Results. Youth with JIA had low disease activity, pain, and disability scores. Youth with JIA maintained a more erect posture at the hip (β = –4.0°, $P = 0.004$), knee (β = 7.5°, $P = 0.004$) and ankle (β = –2.6°, $P = 0.001$). GRF and phase durations outcomes did not meet criteria for significant differences. Knee extension increased with participant age (β = –1.0°, $P = 0.002$), while female participants displayed greater hip flexion (β = –6.6°, $P = 0.001$) and less ankle dorsiflexion (β = 2.3°, $P = 0.006$).

Conclusion. This study provides evidence for a stiff knee landing strategy by youth with JIA. These findings inform targets for physical therapy management to mitigate the risks of a secondary joint injury in sports participation.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is the most common chronic childhood rheumatic disease (onset age of 4.8 years, range 0.6–16.4 years), affecting approximately 0.1–4.0 per 1,000 children worldwide (1–3). The primary symptoms of JIA include pain, joint swelling, and stiffness, as well as fever and swollen lymph nodes. Modern pharmacologic management approaches, using early, targeted use of nonsteroidal antiinflammatories, intraarticular glucocorticoid injections, disease-modifying agents, and biologics, have contributed to an approximately 50% probability of remission off medication within 5 years of diagnosis in Canada (4). Increasingly effective pharmacologic management in turn raises important questions regarding physical activity participation for children and youth with JIA in periods of disease remission and the role of exercise therapy for this population (5).

Important considerations regarding physical activity participation include secondary consequences of JIA, which may include body structure and function (6), physical activity, and participation outcomes. A growing body of evidence indicates that youth with JIA may experience reduced physical activity (7,8), impaired postural balance (9), alterations in joint biomechanics of gait and jumping (10–14), decreased physical fitness (15), including bone and muscle structure (16), and strength deficits (17). However, the heterogeneity of patient-specific disease characteristics, differential responses to clinical management, and differences in

Supported by the Cumming School of Medicine and Alberta Health Services (Clinical Research Fund, CRF15-027) and the Alberta Children's Hospital Foundation (Carolyn A. Emery Chair in Pediatric Rehabilitation). Trainee funding was supported by the Vi Riddell Pediatric Rehabilitation Research Program. Additional funding was provided by Alberta Innovates Technology Futures (Janet L. Ronsky iCORE Strategic Chair, Advanced Diagnostics and Devices).

[1]Gregor Kuntze, PhD, Shane Esau, MKin, Colleen Nesbitt, MSc, Janet L. Ronsky, PEng, PhD: University of Calgary, Calgary, Alberta, Canada; [2]Alberto Nettel-Aguirre, PhD, PStat, Dianne Mosher, MD, Carolyn A. Emery, PT, PhD: Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; [3]Julia Brooks, PT: Alberta Children's Hospital, Calgary, Alberta, Canada; [4]Marinka Twilt, MD, PhD, Susanne Benseler, MD, PhD: Cumming School of Medicine, University of Calgary, and Alberta Children's Hospital, Calgary, Alberta, Canada.

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

Address correspondence to Gregor Kuntze, PhD, Faculty of Kinesiology, University of Calgary, 2500 University Drive Northwest, Calgary, Alberta, T2N 1N4, Canada. Email: gkuntze@ucalgary.ca.

Submitted for publication October 6, 2019; accepted in revised form April 7, 2020.

SIGNIFICANCE & INNOVATIONS

- Secondary consequences of juvenile idiopathic arthritis (JIA) include body structure and function deficits that may expose youth with JIA to an increased risk of injury when participating in high-impact movement tasks during activities such as school sports. However, little information is available related to known injury risk factors during sport-specific movement tasks for this population.
- The objective of this study was to identify differences in the performance of a vertical drop jump (VDJ) task between youth with and without JIA, using lower-extremity movement biomechanics. The VDJ was chosen specifically due to existing research evidence on the association of poor movement mechanics and greater risk of injury in youth sports.
- The findings of this study revealed that youth with JIA performed the VDJ task with a stiff knee landing strategy, which is regarded as a risk factor for injury in youth sport. Further evidence on the potential confounding effects of age and sex on biomechanics outcomes highlights the importance of including age and sex in research design to reveal the consequences of JIA on body structure and function outcomes.
- Evidence on the secondary consequences of JIA informs targets for physical therapy to restore normal joint function and prepare youth for a safe return to physical activity in periods of disease remission.

assessment techniques may contribute to disparate research findings. While Houghton et al (9) reported impaired postural balance in youth with JIA using an unstable multi-axis platform, Nesbitt et al (8) did not observe dynamic balance deficits using a triple single-leg hop, while Merker et al (18) reported improved balance for youth with JIA using a uni-axis balance platform. Such differences in research findings illustrate the challenges of assessing body structure and function outcomes and highlight the need for additional research to assess the secondary consequences of JIA in youth using validated and reliable protocols and measurement approaches.

A particular area of concern with respect to physical activity participation in youth with JIA is their ability to accommodate impact loading during sports activities. This concern is especially relevant because current evidence indicates increasing participation in school sports by youth with JIA (19). A stiff landing strategy (20–22), characterized by less total hip and knee joint flexion, as well as valgus knee alignment (23) and medial knee displacement (24), have been suggested as risk factors for anterior cruciate ligament (ACL) injury in female youth athletes. Ford et al (12) demonstrated altered kinematics and kinetics of the hip, knee, and ankle joints during a vertical drop jump (VDJ) task, where youth with JIA appeared to favor a flexed landing posture. While the authors attributed such a landing strategy to a potential adaptation to dissipate the landing forces, the strategy may also reflect a strength deficit of the hip and knee extensors (25) and an inability to adequately control the forces during the landing phase. Given the large heterogeneity of JIA, the increasing efficacy of modern pharmacologic management, and limited research evidence on body structure and function outcomes to date, there is a clear need for further information on the consequences of JIA, and specifically on how JIA affects movement task performance (using quantitative biomechanics outcomes) and sports participation in youth.

The objective of this study was to quantify differences in hip, knee, and ankle joint biomechanics and task performance of a VDJ for youth with JIA compared to their age- and sex-matched healthy peers. This study focused on individuals with JIA who have an involved knee joint and receive modern pharmacologic management as well as targeted physical therapy as needed. Further, this study explored sports participation habits of youth with and without JIA to provide a behavioral context regarding potential biomechanical differences.

SUBJECTS AND METHODS

Participants. This study employed a matched pairs cohort study design, matching for age (within 1.5 years) and sex. Recruitment details for this cohort have been previously reported by Kuntze et al (14). Criteria for matched pairs assignments were based on the closest match in age by month between individuals with JIA (JIA group) and their healthy peers (control group). Youth with JIA were ages 10–20 years and had a current diagnosis of JIA with knee involvement, confirmed by a physician using American College of Rheumatology criteria (26). Inclusion criteria for participants with JIA comprised joint involvement of 1 or both knees without systemic symptoms, no change in medication for 3 weeks prior to testing, and no active ankle joint involvement at the time of testing. Further, joints could be symptomatic or in remission at the time of testing. Age- and sex-matched healthy youth had no history of JIA or other rheumatic diseases. Exclusion criteria for all participants comprised contraindications indicated on the Physical Activity Readiness Questionnaire for Everyone (a 7-item self-administered questionnaire), previous lower-extremity musculoskeletal injury or intraarticular steroid injection within 3 months prior to testing that resulted in time loss (work, school, or sport), diagnosis of any other arthritides, or any current medical problem that prevented participation in the study (e.g., neurologic conditions).

Participants with JIA were recruited sequentially as they presented to their acting physician and physical therapist at the Pediatric Rheumatology clinic at the Alberta Children's Hospital and the Richmond Road Diagnostic and Treatment Centre Rheumatology Clinic in Calgary. Control youth were recruited using the Healthy Infants and Children Clinical Research Program at the Alberta Children's Hospital, participant siblings and friends, and word-of-mouth recruitment. All testing was conducted between July 2016 and January 2018. Ethics approval was granted by the University of Calgary Conjoint Health Research Ethics Board at

the University of Calgary, Canada (ethics ID: REB15-3125) and Alberta Health Services. All participants provided signed informed consent/assent.

Disease activity and sports participation. Details of disease activity assessment and cohort characteristics have been previously reported by Kuntze et al (14). Disease activity was recorded by the same study physician (SB) using the Juvenile Arthritis Disease Activity Score, a sensitive numerical score of disease activity in individuals with JIA (27), which consists of measures of active joint count (10 joints), physician global assessment of disease activity, and evaluation of the child's well-being. Further, the Child Heath Assessment Questionnaire (C-HAQ) was completed by participants. The C-HAQ is a validated and reliable tool of self-assessed physical disability in children and adolescents (28). Items of the C-HAQ are scored on a 4-point ordinal scale (from 0 = without any difficulty, to 3 = unable to do) as well as 2 visual analog scale (VAS) scores of disease-related pain and overall well-being (scale 0–3). Sports participation was assessed using a demographics questionnaire. Section D of the questionnaire focused on sports and physical activity participation in the past year. Participants were provided with a range of 48 sports and physical activity options and were asked to identify which sports they participated in over the past year. An option for other sports or physical activities was provided in case a suitable option was not available.

VDJ assessment. Bilateral joint kinematics and ground reaction forces (GRF) were recorded using a 12-camera motion capture system (Motion Analysis, 240 Hz) and 2 OR6-6 force plates (Advanced Mechanical Technology, 2,400 Hz). Thirty-two reflective spherical markers were attached to the pelvis and lower extremities of the participants (14). After a standing neutral trial, participants performed 5 successful VDJs, stepping off a raised platform (height 33 cm). Successful VDJs consisted of participants landing with both feet at the same time (shoulder width apart) on each of the 2 force plates, immediately performing a maximum effort vertical jump upon landing, and following the jump phase, landing again with each foot on the same force plate. All participants were given a task familiarization period and provided verbal confirmation when they felt comfortable with performing the VDJ. Adherence to the task criteria was visually confirmed during testing by a member of the research team, and data were collected until participants performed 5 successful repetitions. The VDJ has been shown to be predictive of ACL injuries (r^2 = 0.88) and has a high test–retest reliability (intraclass correlation coefficient >0.94) (29,30).

Data processing. Kinematics data were processed using EVaRT (Motion Analysis) and hip, knee, and ankle joint angles were computed using Visual3D (C-Motion) (14). Joint angle time series were normalized to the first support phase of the VDJ (101 data points) by extracting foot contact events for the right and left feet. Foot contacts of the right and left foot were identified from vertical GRF (vGRF) data using custom code implemented in Matlab, version 2016b (MathWorks) (31). The duration of the support phase was determined from the time either leg first contacted the ground until the first time either leg first left the ground. The flight phase was defined as the time from the end of the support phase until either leg first contacted a force plate a second time. Outcomes of the vGRF were computed for the landing and push-off phases ($vGRF_L$ and $vGRF_P$, respectively) of the support phase of the VDJ. The $vGRF_L$ was defined as the maximum vGRF for either force plate (i.e., regardless of leg) during the first 30% of the support phase. Similarly, $vGRF_P$ was defined as the maximum vGRF, regardless of leg during the final 30% of the support phase.

Joint kinematics outcomes included bilateral maximum hip flexion angles, hip adduction/abduction and internal/external rotation angles at 50% of support phase, and maximum knee flexion and ankle dorsiflexion angles. To support the analysis of the participants' maximum jumping ability, mean data were computed using 3 of 5 VDJ repetitions with the longest jump phase durations. Kinematics outcomes were analyzed with respect to the indexed leg (the affected leg of participants with JIA and dominant leg of control participants) and the contralateral leg (the unaffected or less affected leg of participants with JIA and the nondominant leg of control participants). The indexed leg was identified by the study rheumatologist, and in cases of bilateral knee involvement, the

Table 1. Participant characteristics of typically developing (control) youth and youth with juvenile idiopathic arthritis (JIA)*

Characteristic	Control (n = 35)	Matched control (n = 30)	JIA (n = 30)
Age, mean ± SD years	14.8 ± 2.9	15.0 ± 2.7	14.8 ± 2.6
Height, mean ± SD meters	1.61 ± 0.12	1.62 ± 0.12	1.63 ± 0.14
Weight, mean ± SD kg	52.1 ± 13.1	53.4 ± 12.8	55.6 ± 15.0
Female, no. (%)	25 (71.4)	21 (70.0)	21 (70.0)
Disease course			
Oligoarticular, no. (%)	NA	NA	14 (46.7)
Polyarticular, no. (%)	NA	NA	14 (46.7)
Enthesitis-related, no. (%)	NA	NA	2 (6.7)
Time since diagnosis, mean (range) months	NA	NA	80 (0–173)
PGA (0–10), no., mean ± SD	NA	NA	26, 0.5 ± 0.7
PtGA (0–10), no., mean ± SD	NA	NA	20, 1.1 ± 2.0
Active joint count, no., mean ± SD	NA	NA	26, 1.6 ± 4.9
Joints with limited ROM, no., mean ± SD	NA	NA	26, 1.7 ± 4.9
Drug management, no. (%)			
DMARDS	NA	NA	28 (71)
Biologics	NA	NA	28 (36)
Intraarticular steroid injections	NA	NA	28 (32)

* DMARDS = disease-modifying antirheumatic drugs; NA = not applicable; PGA = physician global assessment of disease activity; PtGA = parent global assessment of disease activity; ROM = range of motion.

Table 2. Child Health Assessment Questionnaire outcomes for youth with juvenile idiopathic arthritis (JIA), their typically developing (control) peers, and pair differences*

Outcome	Control (n = 35)	JIA (n = 30)	Pair differences (JIA–control) (n = 30)
Pain (0–3)			
Mean ± SD	0.3 ± 0.6	0.4 ± 0.5	0.2 ± 0.8
Median (Q1, Q3)	0.0 (0.0, 0.20)	0.3 (0.0, 0.7)	0.1 (0.0, 0.6)
Range: min., max.	0.0, 2.0	0.0, 2.3	−2.0, 2.3
Global evaluation (0–3)			
Mean ± SD	0.1 ± 0.2	0.5 ± 0.8	0. 5 ± 0.8
Median (Q1, Q3)	0.0 (0.0, 0.0)	0.2 (0.0, 0.6)	0.2 (0.0, 0.6)
Range: min., max.	0.0, 0.7	0.0, 2.8	−0.5, 2.8
Disability index (0–3)			
Mean ± SD	0.0 ± 0.1	0.2 ± 0.3	0.2 ± 0.3
Median (Q1, Q3)	0.0 (0.0, 0.0)	0.0 (0.0, 0.3)	0.0 (0.0, 0.3)
Range: min., max.	0.0, 0.3	0.0, 0.9	−0.3, 0.9

* max. = maximum; min. = minimum; Q1 = 1st quartile; Q3 = 3rd quartile.

indexed leg was identified by the participant as the leg that they felt was worst affected. The dominant leg of control participants was identified by determining which leg they prefer to kick a ball with.

Statistical analysis. Data analysis was conducted using Matlab and R, version 3.5.0 (R Core Team). The effects of JIA on kinematics outcomes of the indexed leg and contralateral leg (i.e., JIA–control) were investigated using a multivariate (i.e., vector of responses) random coefficient model in R (32) ($\alpha = 0.05$) using the nonlinear mixed-effects package (33). In this approach, joint angle outcomes for each participant were analyzed as a single 5-dimensional DATA vector of correlated outcomes (32). Assumptions for normality of residuals were visually assessed using Q-Q plots and plots of residuals against the fitted values. In the multivariate random coefficient model, JIA and control joint angle data were considered as dependent multivariate samples, with matched pairs considered as random effects. The effects of group (JIA and control) and leg (indexed leg and contralateral leg) as well as the effects of potential confounders (i.e., age, sex, and body mass index [BMI]) were modeled as fixed effects within the multivariate model.

Landing and push-off maximum vGRFs and VDJ support and flight phase durations were analyzed using paired samples *t*-tests with Bonferroni correction for false discovery rate ($\alpha = 0.0125$). Mean matched pair differences and 98.75% confidence intervals (CIs) were computed, where a significant difference was concluded if the CI did not include zero. Differences in participant characteristics were explored using mean ± SDs, interquartile ranges, and minimum and maximum data ranges. Sports participation was assessed with respect to sports with ≥15% participation by youth with and without JIA to provide information on common sports choices.

RESULTS

Cohort characteristics. Sixty-five youth participated in this study (JIA, n = 30; control, n = 35), resulting in 30 matched pairs. The proportion of female pairs (70%) was greater than male (30%), where participants had a mean age of ~15 years, and youth with JIA presented with predominantly oligoarticular and polyarticular JIA subtypes (oligoarticular 46.7%, polyarticular 46.7%, enthesitis-related 6.7%) (Table 1). The median disease duration was 80 months (range 0–173 months). The physician assessment

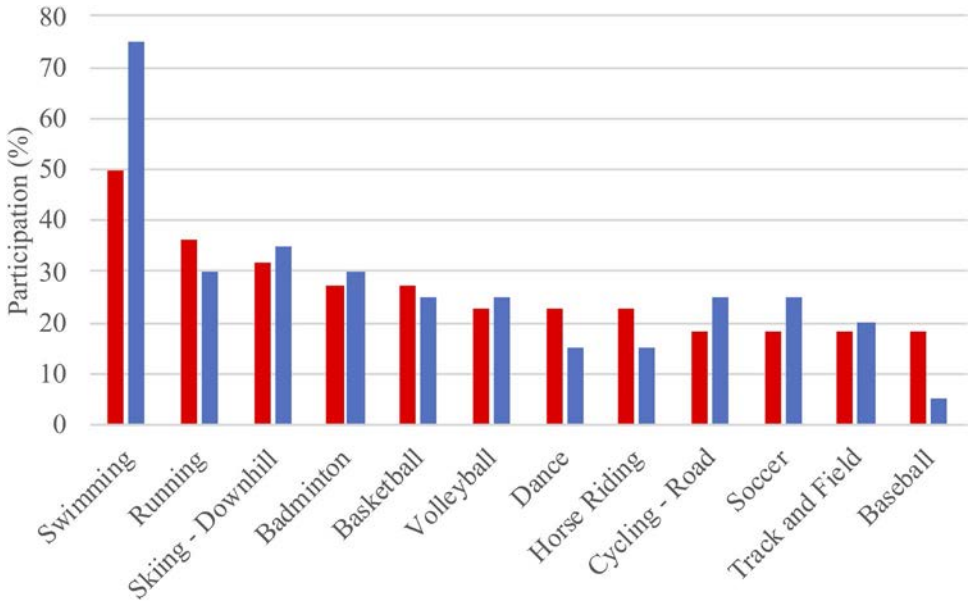

Figure 1. Sports and physical activity participation over the past year by youth with JIA (red bars) and the control group without JIA (blue bars). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24219/abstract.

was conducted for 26 of 30 youth with JIA who displayed low scores for physician global assessment of disease activity (0–10 range) (mean ± SD 0.5 ± 0.7), active joint count (mean ± SD 1.6 ± 4.9 joints), and joints with limited range of motion (mean ± SD 1.7 ± 4.9 joints). Parent assessment of disease activity was completed for 20 of 30 participants (0–10 range; mean ± SD 1.1 ± 2.0) because not all participants were accompanied by their parents. The type of antirheumatic drug used was reported by 28 participants (71% disease-modifying antirheumatic drugs, 36% biologics, 32% intraarticular steroid injections) (Table 1).

All participants completed the C-HAQ pain and global evaluation VAS (0–3 range). Median pair differences indicated slightly elevated pain ratings (median 0.1 [first quartile (Q1)–third quartile (Q3) 0.0–0.6], range 0–3) and reduced global evaluation scores (median 0.2 [Q1–Q3 0.0–0.6], range 0–3) in youth with JIA (Table 2), as well as elevated disability ratings for some individuals with JIA (median 0.0 [Q1–Q3 0.0–0.3], range 0–3) (Table 2). Further, 22 of 30 youth with JIA and 20 of 35 healthy controls completed the sports and physical activity participation questionnaire (Figure 1). Findings for the top 12 sports and physical activities for youth with JIA indicated a greater preference for swimming by control participants. Participation preferences appeared to be similar between groups across other sports.

VDJ biomechanics. Multivariate analysis of hip, knee, and ankle joint kinematics outcomes revealed a significant effect of group (JIA, control) on maximum sagittal and frontal plane hip angles (β = –4.0°, P = 0.004 and β = 2.9°, P = 0.027, respectively), as well as sagittal plane knee (β = 7.5°, P = 0.001) and ankle (β = –2.6°, P = 0.001) angles (Table 3). No significant differences were observed in the transverse plane of the hip (β = –0.2°, P = 0.906). Participants with JIA displayed less hip flexion (JIA indexed leg mean ± SD 88.2 ± 8.8°, contralateral leg 87.2 ± 8.2°; control indexed leg 92.3 ± 7.4°, contralateral leg 91.6 ± 7.6°), and greater hip adduction (JIA indexed leg mean ± SD –8.4 ± 6.5°, contralateral leg –5.2 ± 6.5°; control indexed leg –9.5 ± 7.3°, contralateral leg –10.1 ± 7.7°) than their control peers (Table 4 and Figure 2). Participants with JIA displayed less knee flexion (JIA indexed leg mean ± SD –92.6 ± 8.6°, contralateral leg –92.6 ± 8.0°; control indexed leg –100.6 ± 8.1°, contralateral leg –100.7 ± 7.9°) and less ankle dorsiflexion (JIA indexed leg mean ± SD 32.2 ± 5.2°, contralateral leg 33.9 ± 4.4°; control indexed leg 35.5 ± 4.1°, contralateral leg 35.7 ± 3.4°) than their control peers (Table 3 and Figure 2).

A significant effect of age was observed for maximum sagittal plane knee joint angles (β = –1.0°, P = 0.002), indicating that for every 1 year in increasing age there was a 1.0° increase in the knee flexion angle. Further, a significant effect of sex was observed for maximum sagittal plane ankle and hip joint angles (β = 2.3°, P = 0.006 and β = –6.6°, P = 0.001, respectively). Here, female participants performed the VDJ with less ankle dorsiflexion (mean ± SD female 33.6 ± 4.5°, male 35.9 ± 4.0°) and with greater hip flexion (mean ± SD female 91.8 ± 7.7°, male 85.3 ± 7.5°). No further fixed effects met the criteria for a significant effect. However, effects with significance close to the 0.05 cutoff could be observed for age on frontal plane hip joint angles (β = –0.5°, P = 0.057) and BMI on sagittal knee joint angles (β = 0.5°, P = 0.066). Assessment of the random effect indicates that the variability explained by differences between pairs was low (1.68^{-5}) compared to the variability across participants (4.19), and matched pairs did not have a substantial effect on the outcomes.

Differences in maximum landing and push-off vGRFs between participants with JIA and their control peers did not meet criteria for a significant between-pairs difference ($vGRF_L$ JIA mean pair difference 0.074 body weight [98.75% CI –0.109, 0.257]; $vGRF_P$ 0.011 body weight [98.75% CI –0.079, 0.101]). Despite the CIs of the support phase durations (–0.049 seconds [98.75% CI –0.130, 0.031]) and flight phase durations (–0.024 seconds [98.75% CI –0.0558, 0.007]), including zero, the low value of the upper CI limits and comparatively higher value of the lower CI limits indicate that participants with JIA may have performed the VDJ with approximately 6–7%

Table 3. Multivariate model fixed effects outcomes (JIA–control)*

Outcomes and fixed effect	β	SE	df	*t*-value	*P*
Ankle flexion/extension					
Group	–2.64	0.79	493	–3.36	0.001†
Leg	0.92	0.78	493	1.18	0.240
Age	0.12	0.17	493	0.72	0.471
Sex	2.31	0.84	493	2.74	0.006†
Body mass index	0.14	0.13	493	1.06	0.289
Knee flexion/extension					
Group	7.53	1.48	493	5.10	0.000†
Leg	–0.09	1.46	493	–0.06	0.949
Age	–1.01	0.32	493	–3.15	0.002†
Sex	–1.21	1.58	493	–0.77	0.443
Body mass index	0.46	0.25	493	1.84	0.066
Hip flexion/extension					
Group	–4.04	1.39	493	–2.90	0.004†
Leg	–0.79	1.38	493	–0.58	0.565
Age	0.38	0.30	493	1.25	0.211
Sex	–6.56	1.49	493	–4.39	0.000†
Body mass index	–0.18	0.23	493	–0.79	0.431
Hip adduction/abduction					
Group	2.90	1.31	493	2.21	0.027†
Leg	1.32	1.30	493	1.02	0.308
Age	–0.54	0.28	493	–1.91	0.057
Sex	–0.58	1.40	493	–0.41	0.679
Body mass index	–0.07	0.22	493	–0.31	0.754
Hip internal/external rotation					
Group	–0.18	1.49	493	–0.12	0.906
Leg	0.18	1.48	493	0.12	0.901
Age	–0.27	0.32	493	–0.82	0.412
Sex	–0.97	1.60	493	–0.61	0.545
Body mass index	–0.15	0.25	493	–0.61	0.539

* β = between-group difference; df = degrees of freedom; JIA = juvenile idiopathic arthritis.
† Statistically significant.

Table 4. Joint angle and VDJ phase duration outcomes for youth with juvenile idiopathic arthritis (JIA) and their healthy (control) peers*

Outcome	Control		JIA	
	Indexed	Contralateral	Indexed	Contralateral
Maximum hip flexion, degrees	92.3 ± 7.4	91.6 ± 7.6	88.2 ± 8.8	87.2 ± 8.2
Hip add./abd. (50% support phase), degrees	−9.5 ± 7.3	−10.1 ± 7.7	−8.4 ± 6.5	−5.2 ± 6.5
Hip int./ext. rotation (50% support phase), degrees	−0.1 ± 5.8	3.0 ± 7.5	2.6 ± 9.4	−0.2 ± 8.2
Maximum knee flexion, degrees	−100.6 ± 8.1	−100.7 ± 7.9	−92.6 ± 8.6	−92.6 ± 8.0
Maximum ankle dorsiflexion, degrees	35.5 ± 4.1	35.7 ± 3.4	32.2 ± 5.2	33.9 ± 4.4
Support phase duration, seconds†	0.74 ± 0.14		0.69 ± 0.09	
Flight phase duration, seconds†	0.44 ± 0.05		0.42 ± 0.05	
Maximum vGRF landing, body weight†	1.34 ± 0.24		1.42 ± 0.27	
Maximum vGRF push-off, body weight†	0.98 ± 0.13		0.99 ± 0.10	

* Values are the mean ± SD. add. = adduction; abd. = abduction; ext. = external; int. = internal; VDJ = vertical drop jump; vGRF = vertical ground reaction forces.
† Comparison between controls and JIA group, independent of indexed and contralateral leg.

shorter support and flight phase durations (mean difference). These findings may indicate an overall reduced performance of the 2 phases of the VDJ.

DISCUSSION

The findings of this investigation provide further evidence of the presence of multijoint movement alterations in youth with JIA with knee joint involvement. Despite generally low disease activity, youth with JIA performed the VDJ task with a stiffer landing strategy than their healthy matched peers, indicating functional adaptations when performing high-impact movement tasks.

Youth with JIA had primarily oligoarticular (46.7%) and polyarticular (46.7%) disease and appeared to have an effective disease management based on generally low scores for physician and parent assessments of disease activity, active joint count, joints with limited range of motion, and C-HAQ outcomes (Tables 1 and 2). Disease management involved disease-modifying antirheumatic drugs (71%), biologics (36%), and intraarticular steroid injections (32%) (Table 1). Despite low disease activity, marked

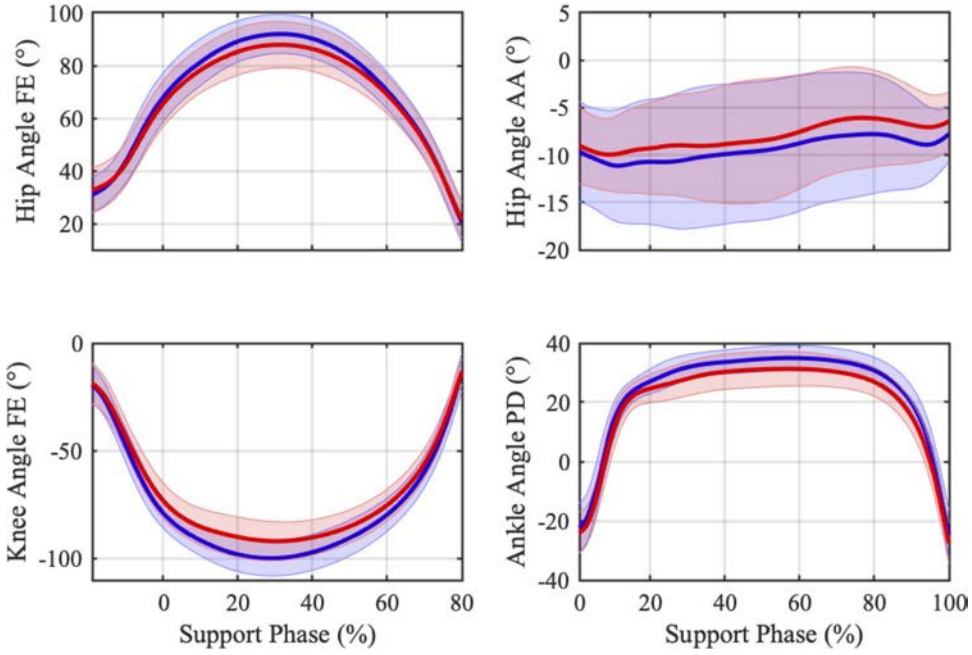

Figure 2. Joint kinematics angle time series of the indexed leg: hip (sagittal and frontal plane), knee (sagittal plane), and ankle (sagittal plane). Data for the support phase of the VDJ are presented for the indexed leg of healthy participants (control) in blue and youth with JIA in red. Positive angles represent hip flexion, hip adduction, knee extension, and ankle dorsiflexion. Shaded bands indicate the SD. AA = adduction/abduction; FE = flexion/extension; PD = plantarflexion/dorsiflexion.

lower-extremity kinematics deviations were observed for the VDJ support phase. Youth with JIA maintained a more erect posture with bilateral reductions in hip and knee flexion, as well as reduced ankle joint dorsiflexion compared to their age- and sex-matched peers (Tables 3 and 4 and Figure 2). The largest differences between individuals with and without JIA were observed at the knee, where youth with JIA had approximately 8° greater knee extension than their healthy peers (Table 4). Further, youth with JIA maintained approximately 4° greater hip extension and approximately 2–3° greater ankle plantar flexion (Table 4).

Interestingly, these findings differed from those of Ford et al (12), who observed greater support phase hip and knee flexion angles for youth with JIA. Participants in these studies appeared to be similar with respect to age, JIA subtypes, and joint involvement. However, differences in participant characteristics (e.g., physical activity and sports participation), clinical management (e.g., physical therapy), and measurement protocol (e.g., VDJ task instructions) may have contributed to differences in joint biomechanics. Differences in task familiarity in particular could have a substantial effect on VDJ task performance. Notably, healthy participants in the study by Ford et al performed the VDJ with comparatively low hip and knee joint flexion angles (12), indicating that task familiarity and task instructions could be potential contributors to the observed differences. Unfortunately, Ford et al did not report on the clinical management for participants with JIA, and no additional information related to participants or VDJ task performance criteria is available.

The findings of reduced sagittal plane joint angles are in line with a stiff landing strategy (22), which has been identified as a possible risk factor for knee ligament injury in female youth athletes (20,21). Specifically, Leppänen et al identified a higher risk of knee ligament injury in female basketball and floorball players who landed with a more extended knee joint and lower hip flexion range of motion. Valgus collapse is frequently stated as a potential risk factor for knee ligament injury (34). The findings of this study appear to support the presence of risk factors associated with a knee valgus alignment. Youth with JIA performed the VDJ with greater hip adduction, with differences of approximately 5° for the contralateral leg (Table 4). While the mechanisms and contributions of sagittal and frontal plane joint kinematics risk factors for knee injury continue to be debated, interventions aimed at increasing hip and knee flexion during landings have proven successful for injury prevention (35,36). Given that youth with JIA appear to experience similar at-risk kinematics, while experiencing similar vGRF loading profiles as their healthy peers (Table 4), injury prevention strategies may have a similar protective effect in youth with JIA.

Findings of a stiff landing strategy and approximately 6–7% longer support phase and shorter flight phase durations (Table 4) may be interpreted as a hesitation or unfamiliarity of youth with JIA to perform high-impact movement tasks. Given that jumping activities in youth with JIA are poorly tolerated when used as part of an exercise program (5), youth with JIA may limit their exposure to jumping tasks. When considering preferences for sports participation, no obvious differences between youth with and without JIA were observed (Figure 1). Unfortunately, more detailed information on the duration of sports participation was not reliably available for this cohort. Therefore, whether JIA affects total sports exposure is currently unknown (including coaching and training), which may affect strength and coordination and contribute to differences in task-specific biomechanics. When viewed from the perspective of generally increasing school sports participation (19), similar preferences for high-impact sports, in the presence of potential biomechanical risk factors, may indicate a greater injury risk in youth with JIA. Unfortunately, the associations between JIA and the risk of injury do not appear to be defined at this stage, limiting opportunities to inform intervention strategies.

Consideration of potential confounders in the statistical analysis is essential in research involving children and youth populations. The findings of this research indicated significant effects of both age and sex across a number of joint angle outcomes. Specifically, older participants appeared to perform the VDJ with greater knee flexion angles (1°/1-year age increase in the knee flexion angle). These findings are in line with those by Gheller et al (37) and may reflect an attempt by older and likely stronger participants to achieve maximum jump height through greater thigh-trunk coupling and resultant increased net impulse. Further, female participants appeared to perform the VDJ with less ankle dorsiflexion and greater hip flexion. While the causes for such sex-specific differences are unknown, they may be related to differences in joint coordination strategies to achieve maximum jump height.

This study focused on a subset of individuals with knee involvement at the time of testing. Therefore, the results of this study reflect a subgroup of the local clinical population only and may not be generalizable across individuals with JIA. Due to the inclusion of participants ages >18 years, the results relating with C-HAQ have to be viewed with caution. Given the limited sample size and generally low disease activity for this cohort, investigating associations between jump landing mechanics and disease activity is currently not possible. Further, soft tissue movement artifact during high-impact movement tasks may influence joint angle outcomes, and frontal and transverse plane joint angles in particular have to be treated with caution.

Youth with JIA exhibited multijoint movement alterations in a VDJ task, despite seemingly effective clinical management. Based on existing evidence for the risk of knee joint injury associated with a stiff landing strategy, these findings inform targets for rehabilitation interventions to mitigate the risks of high impact movements by youth with JIA, such as school sports participation. This evidence informs future research regarding injury risk and the efficacy of physical therapy and exercise interventions to enable a safe return to physical activity for youth with JIA.

ACKNOWLEDGMENTS

The authors thank all patients, volunteers, and their families for their participation in this research study.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Kuntze, Nesbitt, Mosher, Twilt, Benseler, Ronsky, Emery.

Acquisition of data. Kuntze, Brooks, Esau, Nesbitt, Benseler.

Analysis and interpretation of data. Kuntze, Nettel-Aguirre.

REFERENCES

1. Saurenmann RK, Rose JB, Tyrrell P, Feldman BM, Laxer RM, Schneider R, et al. Epidemiology of juvenile idiopathic arthritis in a multiethnic cohort: ethnicity as a risk factor. Arthritis Rheum 2007; 56:1974–84.
2. Laaksonen AL. A prognostic study of juvenile rheumatoid arthritis: analysis of 544 cases. Acta Paediatr Scand 1966;166:49–55.
3. Prieur AM, Le Gall E, Karman F, Edan C, Lasserre O, Goujard J. Epidemiologic survey of juvenile chronic arthritis in France: comparison of data obtained from two different regions. Clin Exp Rheumatol 1987;5:217–23.
4. Guzman J, Oen K, Tucker LB, Huber AM, Shiff N, Boire G, et al. The outcomes of juvenile idiopathic arthritis in children managed with contemporary treatments: results from the ReACCh-Out cohort. Ann Rheum Dis 2015;74:1854–60.
5. Kuntze G, Nesbitt C, Whittaker JL, Nettel-Aguirre A, Toomey C, Esau S, et al. Exercise therapy in juvenile idiopathic arthritis: a systematic review and meta-analysis. Arch Phys Med Rehabil 2018;99:178–93.
6. International classification of functioning, disability and health: ICF. Geneva: World Health Organization; 2001.
7. Bohr AH, Nielsen S, Müller K, Karup Pedersen F, Andersen LB. Reduced physical activity in children and adolescents with juvenile idiopathic arthritis despite satisfactory control of inflammation. Pediatr Rheumatol Online J 2015;13:57.
8. Nesbitt C, Kuntze G, Toomey C, Esau S, Brooks J, Mosher D, et al. Secondary consequences of juvenile idiopathic arthritis in children and adolescents with knee involvement: physical activity, adiposity, fitness and functional performance. Rheumatol Int (under review).
9. Houghton KM, Guzman J. Evaluation of static and dynamic postural balance in children with juvenile idiopathic arthritis. Pediatr Phys Ther 2013;25:150–7.
10. Hartmann M, Kreuzpointner F, Haefner R, Michels H, Schwirtz A, Haas JP. Effects of juvenile idiopathic arthritis on kinematics and kinetics of the lower extremities call for consequences in physical activities recommendations. Int J Pediatr 2010;2010: 835984.
11. Broström E, Hagelberg S, Haglund-Akerlind Y. Effect of joint injections in children with juvenile idiopathic arthritis: evaluation by 3D-gait analysis. Acta Paediatr 2004;93:906–10.
12. Ford KR, Myer GD, Melson PG, Darnell SC, Brunner HI, Hewett TE. Land-jump performance in patients with juvenile idiopathic arthritis (JIA): a comparison to matched controls. Int J Rheumatol 2009;2009:1–5.
13. Frigo C, Bardare M, Corona F, Casnaghi D, Cimaz R, Fovino PL, et al. Gait alteration in patients with juvenile chronic arthritis: a computerized analysis. J Orthop Rheumatol 1988;9:82–90.
14. Kuntze G, Nesbitt C, Nettel-Aguirre A, Esau S, Scholz R, Brooks J, et al. Gait adaptations in youth with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2020;72:917–24.
15. Takken T, Hemel A, van der Net J, Helders PJ. Aerobic fitness in children with juvenile idiopathic arthritis: a systematic review. J Rheumatol 2002;29:2643–7.
16. Burnham JM, Shults J, Dubner SE, Sembhi H, Zemel BS, Leonard MB. Bone density, structure, and strength in juvenile idiopathic arthritis: importance of disease severity and muscle deficits. Arthritis Rheum 2008;58:2518–27.
17. Hansmann S, Benseler SM, Kuemmerle-Deschner JB. Dynamic knee joint function in children with juvenile idiopathic arthritis (JIA). Pediatr Rheumatol Online J 2015;13:8.
18. Merker J, Hartmann M, Kreuzpointner F, Schwirtz A, Haas JP. Excellent balance skills despite active and inactive juvenile idiopathic arthritis: unexpected results of a cross-sectional study. Clin Exp Rheumatol 2017;35:161–8.
19. Milatz F, Klotsche J, Niewerth M, Geisemeyer N, Trauzeddel R, Weißbarth-Riedel E, et al. Participation in school sports among children and adolescents with juvenile idiopathic arthritis in the German National Paediatric Rheumatologic Database, 2000–2015: results from a prospective observational cohort study. Pediatr Rheumatol Online J 2019;17:6.
20. Leppänen M, Pasanen K, Kujala UM, Vasankari T, Kannus P, Äyrämö S, et al. Stiff landings are associated with increased ACL injury risk in young female basketball and floorball players. Am J Sports Med 2017;45:386–93.
21. Leppänen M, Pasanen K, Krosshaug T, Kannus P, Vasankari T, Kujala UM, et al. Sagittal plane hip, knee, and ankle biomechanics and the risk of anterior cruciate ligament injury: a prospective study. Orthop J Sports Med 2017;5:2325967117745487.
22. DeVita P, Skelly WA. Effect of landing stiffness on joint kinetics and energetics in the lower extremity. Med Sci Sports Exerc 1992; 24:108–15.
23. Olsen OE, Myklebust G, Engebretsen L, Bahr R. Injury mechanisms for anterior cruciate ligament injuries in team handball. Am J Sports Med 2004;32:1002–12.
24. Krosshaug T, Steffen K, Kristianslund E, Nilstad A, Mok KM, Myklebust G, et al. The vertical drop jump is a poor screening test for ACL injuries in female elite soccer and handball players. Am J Sports Med 2015;44:874–83.
25. Powers CM. The influence of abnormal hip mechanics on knee injury: a biomechanical perspective. J Orthop Sports Phys Ther 2010;40:42–51.
26. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
27. Consolaro A, Ruperto N, Bazso A, Pistorio A, Magni-Manzoni S, Filocamo G, et al. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. Arthritis Rheum 2009;61:658–66.
28. Andersson Gäre B, Ruperto N, Berg S, Hagelberg S, Jonsson NO, Magnusson B, et al. The Swedish version of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ). Clin Exp Rheumatol 2001;19:S146–50.
29. Hewett TE, Myer GD, Ford KR, Heidt RS, Colosimo AJ, McLean SG, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. Am J Sports Med 2005;33:492–501.
30. Noyes FR, Barber-Westin SD, Fleckenstein C, Walsh C, West J. The drop-jump screening test: difference in lower limb control by gender and effect of neuromuscular training in female athletes. Am J Sports Med 2005;33:197–207.

31. Kuntze G, von Tscharner V, Hutchison C, Ronsky JL. Alterations in lower limb multi-muscle activation patterns during stair climbing in female total knee arthroplasty patients. J Neurophysiol 2015;114:2718–25.
32. UCLA Statistical Consulting. Multivariate random coefficient model. URL: https://stats.idre.ucla.edu/r/faq/multivariate-random-coefficient-model/.
33. Pinheiro J, Bates D, DebRoy S, Sarkar D. Linear and nonlinear mixed effects models. 2019. URL: https://cran.r-project.org/package=nlme.
34. Quatman CE, Hewett TE. The anterior cruciate ligament injury controversy: is "valgus collapse" a sex-specific mechanism? Br J Sports Med 2009;43:328–35.
35. Soligard T, Myklebust G, Steffen K, Holme I, Silvers H, Bizzini M, et al. Comprehensive warm-up programme to prevent injuries in young female footballers: cluster randomised controlled trial. BMJ 2008;337:a2469.
36. Longo UG, Loppini M, Berton A, Marinozzi A, Maffulli N, Denaro V. The FIFA 11+ program is effective in preventing injuries in elite male basketball players: a cluster randomized controlled trial. Am J Sports Med 2012;40:996–1005.
37. Gheller RG, Dal Pupo J, Ache-Dias J, Detanico D, Padulo J, dos Santos SG. Effect of different knee starting angles on intersegmental coordination and performance in vertical jumps. Hum Mov Sci 2015;42:71–80.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 964–974
DOI 10.1002/acr.24183

Potential for Major Therapeutic Changes to Produce Significant Clinical Response Across a Broad Range of Disease Activity: An Observational Study of US Veterans With Rheumatoid Arthritis

Brian C. Sauer,[1] Wei Chen,[1] Jincheng Shen,[1] Neil A. Accortt,[2] David H. Collier,[2] and Grant W. Cannon[1]

Objective. To examine the impact of major therapeutic change (MTC) on clinical response across a broad range of disease activity in US veterans with rheumatoid arthritis (RA).

Methods. This historical cohort analysis evaluated patient visits from the Veterans Affairs RA registry between January 1, 2006 and September 30, 2017. Eligible patient visits were a rheumatology visit with 3 disease activity measures, including the Disease Activity Score in 28 joints, the Clinical Disease Activity Index, and the Routine Assessment of Patient Index Data 3; the follow-up visit for all 3 disease activity measures was 2–6 months later. The full population and a subset of patients with active disease (≥6 tender joints, ≥6 swollen joints) were evaluated. Clinical outcome was based on the American College of Rheumatology criteria for 20% improvement in disease activity (ACR20). The effect of MTC on ACR20 response was presented as crude descriptive statistics and evaluated using standardized regression for population- and disease activity–level conditional effects.

Results. The full population comprised 1,208 patients (6,138 visits) and the active disease subpopulation included 383 patients (1,109 visits). Overall, visits with MTC were associated with increased likelihood of ACR20 response across all disease activity measures for the full population. Risk ratios for overall risk of ACR20 response for visits with MTC versus those without MTC ranged from 1.67 to 2.22 across disease activity measures among the full population and from 1.51 to 1.60 for the subpopulation with active disease.

Conclusion. MTC was associated with clinical improvement, even among patients with longstanding RA who had received multiple prior therapies, which emphasizes the utility of therapy modifications for patients with established and active RA.

INTRODUCTION

The 2015 American College of Rheumatology (ACR) guidelines (1) for the treatment of rheumatoid arthritis (RA) recommend the use of validated disease activity measures to guide disease-modifying antirheumatic drug (DMARD) treatment decisions. Specifically, patients with active RA (i.e., moderate-to-high disease activity) should be treated with DMARDs to achieve and sustain a prespecified target, which is often remission or low disease activity. This approach is referred to as the treat-to-target strategy. The evidence-based guidelines recommend adjusting treatment as needed to achieve treatment targets and maintain control of disease activity after treatment targets are obtained, with the goal to prevent long-term joint destruction. Studies have shown that the treat-to-target strategy lowers disease activity and reduces progressive joint damage when compared to routine care (2,3).

Rheumatologists at collaborating Veterans Affairs Rheumatoid Arthritis (VARA) registry sites collect and document core clinical measurements at each clinic visit that support calculation of multiple validated disease activity measures. The clinical measures include physician assessment of swollen and tender joint counts, provider global assessment, erythrocyte sedimentation rate (ESR), patient-reported pain score, patient global well-being score, and the Multidimensional Health Assessment Questionnaire

Supported by Amgen and, in part, by Specialty Care Center of Innovation, Veterans Health Administration, Department of Veterans Affairs. Drs. Sauer and Cannon's work was supported by Amgen.

Drs. Accortt and Collier have ownership or partnership and stock options or bond holdings in Amgen. No other disclosures relevant to this article were reported.

[1]Brian C. Sauer, PhD, MS, Wei Chen, MStat, Jincheng Shen, PhD, Grant W. Cannon, MD: Salt Lake City VA Medical Center and University of Utah, Salt Lake City; [2]Neil A. Accortt, PhD, David H. Collier, MD: Amgen, Thousand Oaks, California.

Address correspondence to Brian C. Sauer, PhD, MS, Veterans Affairs Salt Lake City Health Care System, University of Utah School of Medicine, Salt Lake City, UT. Email: Brian.Sauer@utah.edu.

Submitted for publication August 1, 2019; accepted in revised form March 3, 2020.

SIGNIFICANCE & INNOVATIONS

- The treat-to-target strategy has been shown to improve clinical outcomes in patients with rheumatoid arthritis.
- The treat-to-target strategy requires implementation of major therapeutic changes (MTC) among patients who have not yet achieved the target of remission or low disease activity.
- We observed a clinical benefit (based on the American College of Rheumatology criteria for 20% improvement in disease activity) among patients with moderate-to-severe disease who received an MTC, even among patients who had longstanding disease and had received multiple prior therapies.
- The Routine Assessment of Patient Index Data 3 measure appeared to overestimate disease activity in the full population and may have limited utility in a treat-to-target framework.

(4). Although providers are collecting and documenting core clinical measures, many providers do not appear to be using an explicit treat-to-target strategy.

Our past work indicates that disease activity measures are 1 component in the complex decision to initiate a major therapeutic change (MTC) among patients with active RA (5,6). We found that more than half of the patients in the VARA registry with active RA did not receive an MTC within 90 days of their index visit (5). Review of the medical notes revealed that providers who were not initiating an MTC in patients with active RA often did not document composite disease activity scores. Additionally, providers, and often patients, believed their disease was under control even though their Disease Activity Score in 28 joints (DAS28) indicated active RA. These results are consistent with qualitative interviews with practicing rheumatologists that suggested clinicians often rely on their understanding of a patient's disease activity as a better assessment of clinical status than a validated disease activity measure (7).

Disease activity measures were selected based on common clinical use and ability to be computed at point-of-care, and included the DAS28, the Clinical Disease Activity Index (CDAI), and the Routine Assessment of Patient Index Data 3 (RAPID3). The DAS28 is a widely used instrument in clinical trials (8) and is the VARA standard, but it requires assessment of the laboratory ESR value in relation to the visit and it may not reflect the patient's status at the point of care if an ESR measurement is not scheduled with the visit. The CDAI incorporates objective provider clinical measures (tender and swollen joint counts) and patient and physician global assessments but does not include acute-phase reactant laboratory measures and can be calculated at the point-of-care to support targeted treatment strategies (9). The RAPID3 is based on 3 patient-reported ACR RA core data set measures that include function, pain, and patient global assessment (10). These composite measures of RA disease activity were strongly correlated in clinical trials and clinical care and produced similar classification of patients into 4 disease activity levels (high, moderate, low, and remission) (10). The RAPID3, however, tends to classify patients into higher disease activity categories (11), especially in populations with prevalent comorbidity, which may affect its utility as a measure to guide treatment decisions and monitor provider adherence to treat-to-target strategies.

The objectives of this study were to describe the relationship between disease activity and MTC using the DAS28, CDAI, and RAPID3, and to describe the impact of MTC on clinical response, measured by the ACR criteria for 20% improvement in disease activity (ACR20 response) (12). For this analysis, we also selected a subpopulation of patients in the VARA registry that represented RA patients with active disease who would qualify for inclusion in a clinical trial based on their number of swollen and tender joints, and who would be expected to have treatment changes in a treat-to-target strategy with remission or low disease activity as the target. We also evaluated patients in the VARA registry across the full spectrum of disease severity.

PATIENTS AND METHODS

Population, data source, and study design. The study population comprised patients who are veterans enrolled in the VARA registry (4,13,14). The VARA registry is a prospective, observational registry involving 11 Veterans Affairs (VA) medical centers. Disease activity measure components are collected and documented during routine patient care using templated notes. The disease activity measures are extracted from medical notes (available in the VA Corporate Data Warehouse [CDW]) (15), using validated text extraction algorithms, or entered manually into the database.

The CDW was the primary data source used to construct patient histories. Key data domains included patient demographic characteristics, pharmacy, laboratory, outpatient diagnoses, and electronic medical notes. VARA enrollment data provided the patient demographic characteristics, disease history, and duration of RA. At a central laboratory for VARA patients, serologic status for rheumatoid factor (RF) and anti–cyclic citrullinated peptide antibodies (ACPA) were assayed on enrollment. When data were available from both the CDW and the VARA database, data from the CDW were compared to data from the VARA enrollment file; VARA data were used to correct CDW data when inconsistencies were identified. An additional chart review was performed to collect remaining missing data on patient demographic characteristics, biomarkers, and disease duration not identified in the CDW or VARA database.

We used a historical cohort design to compare clinical response between patient visits with and without an MTC. The unit of observation was eligible patient visits to a rheumatology clinic during the study period (January 1, 2006 to September 30, 2017). Key features of study design included a baseline measurement period (18 months prior to the eligible visit) to measure covariates

and potential confounders; an exposure period (7 days prior to 30 days after the eligible visit) to assess whether an MTC occurred or not. The 7-day previsit exposure period was selected to capture interventions that may have been taken immediately prior to the visit, (e.g., steroid dose escalation via telephone call or electronic message), and the 30-day postvisit period was selected to capture interventions that started at the visit. An outcome period (2–6 months after the eligible visit) was used to determine whether the patient achieved an ACR20 response. The first visit during the outcome period with complete disease activity measure components was included for analysis (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24183/abstract).

This study was approved by the University of Utah Institutional Review Board, the VA Research Service, and the Scientific and Ethical Advisory Board of the VARA registry for analysis of VARA and VA health care data. All patients provided written consent and authorization for use of health information upon enrollment in the VARA registry.

Visit eligibility and patient exclusion criteria. The study population comprised eligible rheumatology visits that met the following criteria: 1) enrolled in the VARA registry; 2) ≥18 years of age; 3) rheumatology visit with all components of disease activity measures (DAS28, CDAI, RAPID3); 4) ≥18 months of enrollment in the VA health care system prior to eligible visit; and 5) 2 rheumatology visits with documented DAS28 scores during the 18-month baseline period that were ≥60 days apart from each other and ≥60 days before the eligible patient visit (in order to measure disease stability; additional criterion of follow-up visit with documented disease activity measure components between 2 and 6 months after the eligible patient visit) (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24183/abstract). Visits were excluded if they were followed by any surgical procedure within 90 days or any hospitalization within 30 days of the visit. Patients were excluded if they had active cancer, organ transplant, or diagnosis of other autoimmune disorders (e.g., systemic lupus erythematosus).

This study consisted of a full population that met the eligibility criteria, representing the spectrum of disease activity, and a subpopulation of patient visits with features of active disease and who were more likely to be indicated for an MTC based on disease activity scores. The same eligibility criteria that were used to establish the full population were implemented for the active disease subpopulation with 1 additional requirement for ≥6 tender joints and ≥6 swollen joints during the eligible visit.

Measurements. *Exposure: MTC.* An MTC was defined as any of the following occurring within the timeframe of 7 days before and 30 days after an eligible visit: 1) initiation of new DMARD (including switching agents within the same drug class) either as a new agent or after a 90-day gap following the last date of prior therapy; 2) escalation of DMARD dose by ≥25% unless preplanned because of initiation of therapy with methotrexate, infliximab, or abatacept (preplanned defined as documented in the medication schedule or dose increased on the second dispensing event of a new treatment course); 3) initiation of prednisone (either as new agent or after 90-day gap in therapy); 4) increase in monthly average prednisone dose by 25%; and/or 5) injection of ≥2 joints with corticosteroids.

Outcome: clinical improvement measured by ACR20 response criteria. In order to achieve an ACR20 response, patients had to improve by 20% on both tender and swollen joint counts and experience 20% improvement in 3 ACR core disease activity measures (patient assessment of pain; patient global assessment of disease activity; physician global assessment of disease activity; patient assessment of physical function; and acute-phase reactant value) (12). ACR20 response was selected to measure treatment effect because it is a validated measure of clinical response in patients with RA. Using the ACR20 response as the clinical response measure across the 3 disease activity measures provides a measure that had demonstrated sensitivity to detect clinical response to treatment during a time window that is consistent with routine follow-up care (~3 months) (16,17). The first visit with complete set of clinical core measures that occurred within the outcome measurement window (2–6 months after visit) was used to account for variability in visit intervals and reduce the risk of exposure misclassification.

Covariates: potential confounders between MTC and ACR20 response. The goal of covariate adjustment was to remove confounding between MTC and ACR20 response. Potential confounders were selected based on clinician background knowledge of patient and disease characteristics that may influence the decision to initiate an MTC. Potential confounders included demographic characteristics (age, sex, race [White or non-White]), duration of RA, disease activity, Rheumatic Disease Comorbidity Index (RDCI) score (18), disease stability over time before the index visit (5,19), baseline use of DMARDs, and an MTC within 90 days of the eligible visit. The disease stability calculation was based on the European Alliance of Associations for Rheumatology (formerly European League Against Rheumatism) response criteria (20) and compared baseline DAS28 scores to index visit DAS28 scores. DAS28 reductions of ≥0.6 points were categorized as improved and increases of ≥0.6 points were categorized as worsening (5,19).

Disease activity level was thought to be the most important confounder because it establishes the indication for an MTC and is the strongest predictor of MTC in patients with active disease (5). We also believed that the level of disease activity would act as an effect modifier when including all patient visits meeting full population inclusion criteria, since RA patients with persistently high disease activity may not respond to treatment and patients in

Table 1. Demographic and clinical characteristics*

	Full population				Active disease population			
	Without MTC (n = 4,365)	Without MTC, 95% CI	With MTC (n = 1,773)	With MTC, 95% CI	Without MTC (n = 589)	Without MTC, 95% CI	With MTC (n = 520)	With MTC, 95% CI
Age, mean ± SD years	65.9 ± 11.0	65.6–66.2	63.3 ± 10.9	62.8–63.8	64.5 ± 10.8	63.6–65.3	62.5 ± 10.7	61.6–63.4
Male sex	3,974 (91.0)	90.2–91.9	1,625 (91.7)	90.3–92.9	557 (94.6)	92.4–96.3	489 (94.0)	91.6–95.9
White race	3,568 (81.7)	80.6–82.9	1,409 (79.5)	77.5–81.3	529 (89.8)	87.1–92.1	449 (86.3)	83.1–89.2
RA duration, mean ± SD years	15.6 ± 12.3	15.2–16.0	14.2 ± 11.9	13.6–14.7	14.3 ± 11.1	13.4–15.1	12.7 ± 11.0	11.8–13.7
RDCI score, mean ± SD	2.3 ± 1.5	2.2–2.3	2.2 ± 1.5	2.2–2.3	2.3 ± 1.4	2.2–2.4	2.2 ± 1.6	2.1–2.4
RF status								
Positive	3,928 (90.0)	89.1–90.9	1,589 (89.6)	88.1–91.0	529 (89.8)	87.1–92.1	457 (87.9)	84.8–90.6
Negative	433 (9.9)	9–10.8	183 (10.3)	8.9–11.8	60 (10.2)	7.9–12.9	63 (12.1)	9.4–15.2
Missing	4 (0.1)	0–0.2	1 (0.1)	0–0.3	–	–	–	–
ACPA status								
Positive	3,622 (83.0)	81.8–84.1	1,484 (83.7)	81.9–85.4	476 (80.8)	77.4–83.9	422 (81.2)	77.5–84.4
Negative	739 (16.9)	15.8–18.1	288 (16.2)	14.6–18.0	113 (19.2)	16.1–22.6	98 (18.8)	15.6–22.5
Missing	4 (0.1)	0–0.2	1 (0.1)	0–0.3	–	–	–	–
Disease activity measure stability								
Better or no change	3,015 (69.1)	67.7–70.4	1,261 (71.1)	69.0–73.2	488 (82.9)	79.6–85.8	405 (77.9)	74.1–81.4
Worse	1,350 (30.9)	29.6–32.3	512 (28.9)	26.8–31.0	101 (17.1)	14.2–20.4	115 (22.1)	18.6–25.9
Recent dispensing episodes,†								
bDMARD	664 (15.2)	14.2–16.3	194 (10.9)	9.5–12.5	97 (16.5)	13.6–19.7	46 (8.8)	6.5–11.6
csDMARD	1,382 (31.7)	30.3–33.1	387 (21.8)	19.9–23.8	177 (30.1)	26.4–33.9	100 (19.2)	15.9–22.9
Prednisone	396 (9.1)	8.2–10.0	180 (10.2)	8.8–11.7	65 (11.0)	8.6–13.8	44 (8.5)	6.2–11.2
Established dispensing episodes,‡								
bDMARD	1,953 (44.7)	43.3–46.2	671 (37.8)	35.6–40.1	322 (54.7)	50.5–58.7	190 (36.5)	32.4–40.8
csDMARD	4,031 (92.3)	91.5–93.1	1,505 (84.9)	83.1–86.5	545 (92.5)	90.1–94.5	434 (83.5)	80.0–86.6
Prednisone	1,926 (44.1)	42.6–45.6	954 (53.8)	51.5–56.1	312 (53.0)	48.8–57.1	282 (54.2)	49.8–58.6
Baseline MTC before index visit								
8–30 days	98 (2.2)	1.8–2.7	31 (1.7)	1.2–2.5	38 (6.5)	4.6–8.7	11 (2.1)	1.1–3.8
8–60 days	363 (8.3)	7.5–9.2	109 (6.1)	5.1–7.4	101 (17.1)	14.2–20.4	41 (7.9)	5.7–10.5
8–90 days	701 (16.1)	15.0–17.2	201 (11.3)	9.9–12.9	166 (28.2)	24.6–32.0	78 (15.0)	12.0–18.4

***** Values are the number (%) unless indicated otherwise. For full population, n = 1,208 patients (6,138 visits); for active disease population, n = 383 patients (1,109 visits). 95% CI = 95% confidence intervals; ACPA = anti–cyclic citrullinated peptide antibodies; bDMARD = biologic disease-modifying antirheumatic drug; csDMARD = conventional synthetic DMARD; MTC = major therapeutic change; RA = rheumatoid arthritis; RDCI = Rheumatic Disease Comorbidity Index; RF = rheumatoid factor.
† Recent dispensing episodes were first dispensed between 8 to 30 days of eligible visit.
‡ Established dispensing episodes were dispensed between 7 to 370 days of eligible visit.

low disease or remission may not receive additional benefit from treatment modification.

Statistical analysis. Descriptive statistics included the number of observations and percentages for dichotomous variables and number of observations, mean ± SDs, and 95% confidence intervals (95% CIs) for continuous variables. Crude (bivariate) associations between MTC and ACR20 response were represented by risk ratios with 95% CIs.

Standard definitions for disease activity levels (remission, low, moderate, high) for each disease activity measure were used (9,10,21) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24183/abstract). Descriptive analyses of MTC criteria were categorized as changes in oral prednisone (initiating medication, restarting medication after a gap, and/or increase in medication dose), intraarticular prednisone injections, changes in biologic DMARD (bDMARD), and changes in conventional synthetic DMARD (csDMARD).

The effect of MTC on ACR20 response was presented as crude descriptive statistics and further evaluated using direct regression standardization (22–25) for the population and disease activity level conditional effects. Since patient visits were correlated (multiple visits were analyzed per patient), we used the log binomial population average generalized estimating equation model with exchangeable correlation structure (26) as the working model. We computed 95% CIs with 1,000 bootstrap samples (random sampling with replacement) (22). The working models were fit using patient age at visit, sex, race, ACPA status, RF status, disease duration, RDCI score, disease activity measure stability measure (worsening or not), csDMARDs and bDMARDs dispensed in the month prior to visit, and the baseline MTC (MTC during previous 90 days). An interaction term was used to estimate the stratum-specific effect of MTC by level of disease activity. The overall population-level effect averaged across the heterogeneous effects by level of disease activity.

We used quartiles to generate 4 disease activity categories in the active disease subpopulation, allowing us to evaluate stratum-specific effects at a more granular level, because most visits with remission/low disease were removed by subpopulation criteria. The probability of ACR20 response was independent of follow-up month when conditioning on MTC and level of disease activity; therefore, we did not model follow-up interval independently in our analysis. We were not able to fit baseline established medication exposure (dispensed in the previous year) for csDMARDs or bDMARDs in all models (across disease activity measures) because of a lack of variation in these measures; for this reason, they were not included in the final models. Software used for these analyses included Microsoft SQL server and SAS, version 9.4; Enterprise Guide, version 7.1, and Stata 14 were used to prepare data and conduct statistical analysis.

RESULTS

Study visits. Two study populations were evaluated. The full population comprised all patient visits meeting eligibility criteria and included 1,208 patients with 6,138 eligible visits, and the active disease population included 383 patients with 1,109 eligible visits. The full population had 4,365 visits without an MTC and 1,773 visits with an MTC, while the active disease population had 589 visits without an MTC and 520 visits with an MTC.

Patient demographic and clinical characteristics. For the full population and the active disease population, patients were younger during visits with an MTC (mean ages 63.3 years [95% CI 62.8–63.8 years] and 62.5 years [95% CI 61.6–63.4 years], respectively) compared to visits without an MTC (mean ages 65.9 years [95% CI 65.6–66.2 years] and 64.5 years [95% CI 63.6–65.3 years], respectively) (Table 1). Patient visits with an MTC had shorter RA disease duration (mean 14.2 years [95% CI 13.6–14.7 years] and 12.7 years [95% CI 11.8–13.7 years] for the full and active disease populations, respectively) compared to patients visits without an MTC (mean 15.6 years [95% CI 15.2–16.0 years] and 14.3 years [95% CI 13.4–15.1 years]). The percentage of male participants, White race, RF-positive status, ACPA-positive status, and disease stability were similar between visits with or without MTC for both study populations.

For the full and active disease populations, patient visits with MTC had a lower percentage of recent (within the past month) bDMARD-dispensing episodes (10.9% [95% CI 9.5–12.5] and 8.8% [95% CI 6.5–11.6], respectively) compared to patient visits without MTC (15.2% [95% CI 14.2–16.3] and 16.5% [95% CI 13.6–19.7], respectively). Visits with MTC also had a lower percentage of established (within the past year) bDMARD-dispensing episodes (37.8% [95% CI 35.6–40.1] and 36.5% [95% CI 32.4–40.8] for the full and active disease populations, respectively) compared to patient visits without MTC (44.7% [95% CI 43.3–46.2] and 54.7% [95% CI 50.5–58.7] for the full and active disease populations, respectively). Baseline use of csDMARDs was similar to the use of bDMARDs; visits with MTC had lower percentages of active and established csDMARD-dispensing episodes compared to visits without MTC for both study populations (Table 1). Finally, visits with MTC had a lower percentage of MTC during the 90 days prior to the eligible visit (11.3% [95% CI 9.9–12.9] for the full population and 15.0% [95% CI 12.0–18.4] for the active disease population) compared to visits without MTC during the previous 90 days (16.1% [95% CI 15.0–17.2] for the full population and 28.2% [95% CI 24.6–32.0] for the active disease population).

Classification of disease activity levels. In the full population, 65.2%, 63.1%, and 80.9% of visits were categorized as either moderate or high disease by the DAS28, CDAI, and RAPID3, respectively (Figure 1). In the active disease population, 97.1%, 100%, and 93.9% of visits were categorized as moderate or high disease by the DAS28, CDAI, and RAPID3, respectively.

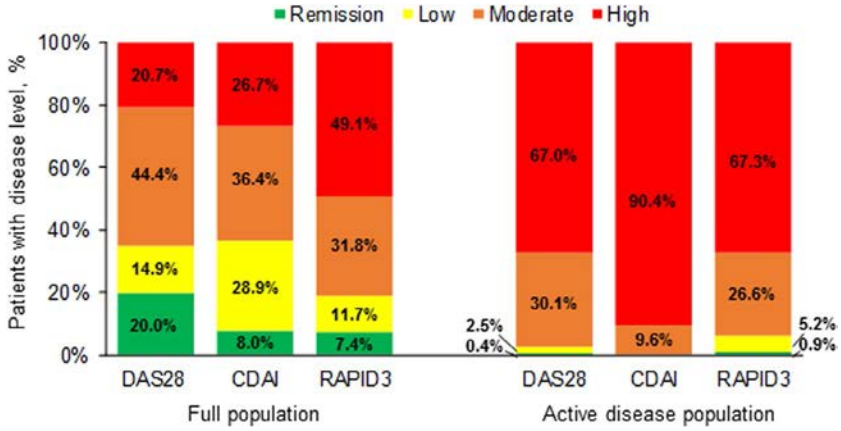

Figure 1. Distribution of disease activity levels by disease activity measure. Bars show the percentage of patients in disease activity levels of remission, low, moderate, or high for the full population and the active disease population. CDAI = Clinical Disease Activity Index; DAS28 = Disease Activity Score in 28 joints; RAPID3 = Routine Assessment of Patient Data 3.

Description of types of MTC. The most common type of MTC across all disease activity measures and levels of disease activity was a change to csDMARD therapy (Table 2). For patient visits with remission or low disease activity, the second most common type of MTC was a change to oral prednisone, and for patients with moderate or high disease activity was a change to bDMARD therapy or to oral prednisone.

Description of MTC and frequency of ACR20 response. In a crude analysis describing the frequency of ACR20 responses by MTC, patient visits with remission or low disease activity were generally not associated with MTC, and as expected these visits with MTC had low rates of ACR20 responses (Table 3). Approximately one-third of visits with high disease activity with MTC were associated with an ACR20 response (range 32.9–36.9%) regardless of disease activity measure. Across all disease activity measures, the prevalence of ACR20 responses was higher in the active disease population than in the full population. Table 4 provides the crude results for the active disease population using quartiles of the disease activity levels. MTC was generally associated with an increased prevalence of ACR20 response across all strata, but not all strata-specific estimates met statistical significance.

Population and stratum-specific model-based effect of MTC on ACR20 response. In the population (overall) analysis, a visit with MTC resulted in statistically significant increase of an ACR20 response across all disease activity measures for both populations (Table 5). The stratum-specific effects varied by disease activity measure and study population. In the full population,

Table 2. Type of MTC by disease activity measure and disease activity category*

	Full population				Active disease population			
DAM category	Change in oral prednisone†	Prednisone injection	Change in bDMARD†	Change in csDMARD†	Change in oral prednisone†	Prednisone injection	Change in bDMARD†	Change in csDMARD†
DAS28								
Remission	59 (4.8)	0	31 (2.5)	124 (10.1)	0	0	0	0
Low	56 (6.1)	3 (0.3)	23 (2.5)	122 (13.4)	2 (7.1)	0	0	6 (21.4)
Moderate	192 (7.0)	20 (0.7)	177 (6.5)	454 (16.6)	23 (6.9)	4 (1.2)	35 (10.4)	77 (23.1)
High	163 (12.8)	16 (1.3)	169 (13.3)	310 (24.4)	116 (15.6)	10 (1.4)	119 (16.0)	193 (26.0)
CDAI								
Remission	23 (4.7)	0	12 (2.4)	48 (9.7)	0	0	0	0
Low	87 (4.9)	4 (0.2)	46 (2.6)	197 (11.1)	0	0	0	0
Moderate	168 (7.5)	16 (0.7)	121 (5.4)	369 (16.5)	3 (2.8)	0	6 (5.7)	17 (16.0)
High	192 (11.7)	19 (1.3)	221 (13.5)	396 (24.2)	138 (13.8)	14 (1.4)	148 (14.8)	259 (25.8)
RAPID3								
Remission	22 (4.9)	1 (0.2)	11 (2.4)	61 (13.5)	0	0	0	4 (40.0)
Low	28 (3.9)	1 (0.1)	24 (3.3)	88 (12.2)	2 (3.5)	0	6 (10.3)	14 (24.1)
Moderate	116 (6.0)	5 (0.3)	83 (4.3)	281 (14.4)	30 (10.2)	0	29 (9.8)	52 (17.6)
High	304 (10.1)	32 (1.1)	282 (9.4)	580 (19.2)	109 (14.6)	14 (1.9)	119 (16.0)	206 (27.6)

* Values are the number (%). For full population, n = 6,138 visits; for active disease population, n = 1,109 visits. bDMARD = biologic disease-modifying antirheumatic drug; CDAI = Clinical Disease Activity Index; csDMARD = conventional synthetic DMARD; DAM = disease activity measure; DAS28 = Disease Activity Score in 28 joints; MTC = major therapeutic change; RAPID3 = Routine Assessment of Patient Index Data 3.
† Changes included initiating medication, restarting medication after a gap, and/or change in medication dose.

Table 3. Crude descriptive analysis of the frequency of MTC and ACR20 response by disease activity measure and category*

	Full population without MTC		Full population with MTC			Active disease population without MTC		Active disease population with MTC		
	Frequency	ACR20 response frequency (%)	Frequency	ACR20 response frequency (%)	RR (95% CI)	Frequency	ACR20 response frequency (%)	Frequency	ACR20 response frequency (%)	RR (95% CI)
Overall	4,365	308 (7.1)	1,773	337 (19.0)	2.7 (2.33–3.11)	589	100 (17.0)	520	174 (33.5)	2.0 (1.59–2.45)
DAS28 category										
Remission	1,019	15 (1.5)	207	4 (1.9)	1.3 (0.44–3.92)	4	1 (25.0)	0	NA	NA
Low	715	20 (2.8)	197	6 (3.1)	1.1 (0.44–2.67)	20	4 (20.0)	8	0	NA
Moderate	1,935	160 (8.3)	792	132 (16.7)	2.0 (1.63–2.50)	204	24 (11.8)	130	33 (25.4)	2.2 (1.34–3.48)
High	696	113 (16.2)	577	195 (33.8)	2.1 (1.70–2.55)	361	71 (19.7)	382	141 (36.9)	1.9 (1.47–2.40)
CDAI category										
Remission	412	0	81	0	NA	0	NA	0	NA	NA
Low	1,454	29 (2.0)	320	8 (2.50)	1.3 (0.58–2.72)	0	NA	0	NA	NA
Moderate	1,599	137 (8.6)	635	112 (17.6)	2.1 (1.63–2.60)	81	9 (11.1)	25	11 (44.0)	4.0 (1.86–8.45)
High	900	142 (15.8)	737	217 (29.4)	1.9 (1.55–2.25)	508	91 (17.9)	495	163 (32.9)	1.8 (1.47–2.30)
RAPID3 category										
Remission	360	9 (2.5)	92	3 (3.3)	1.3 (0.36–4.72)	6	1 (16.7)	4	1 (25.0)	1.5 (0.13–17.7)
Low	583	24 (4.1)	138	12 (8.7)	2.1 (1.08–4.12)	37	2 (5.4)	21	3 (14.3)	2.6 (0.48–14.6)
Moderate	1,494	92 (6.2)	456	67 (14.7)	2.4 (1.77–3.21)	192	32 (16.7)	103	29 (28.2)	1.7 (1.09–2.63)
High	1,928	183 (9.5)	1,087	255 (23.5)	2.5 (2.08–2.94)	354	65 (18.4)	392	141 (36.0)	2.0 (1.52–2.53)

* For the full population, n = 6,138 visits; for the active disease population, n = 1,109 visits. 95% confidence interval = 95% CI; ACR20 = American College of Rheumatology criteria for 20% improvement in disease activity; CDAI = Clinical Disease Activity Index; DAS28 = Disease Activity Score in 28 joints; MTC = major therapeutic change; NA = not applicable; RAPID3 = Routine Assessment of Patient Index Data 3; RR = risk ratio.

Table 4. Crude descriptive analysis of the frequency of MTC and ACR20 response by disease activity quartiles in the active disease population*

	Without MTC (589 visits)		With MTC (520 visits)		
	Frequency	ACR20 response frequency (%)	Frequency	ACR20 response frequency (%)	RR (95% CI)
Overall	589	100 (17.0)	520	174 (33.5)	2.0 (1.59–2.45)
DAS28 quartiles					
2.35–4.87	178	22 (12.4)	100	24 (24.0)	1.9 (1.15–3.28)
4.88–5.60	166	25 (15.1)	111	35 (31.5)	2.1 (1.33–3.30)
5.61–6.29	139	26 (18.7)	138	40 (29.0)	1.5 (1.00–2.39)
6.30–8.76	106	27 (25.5)	171	75 (43.9)	1.7 (1.19–2.49)
CDAI quartiles					
14.7–26.6	182	24 (13.2)	99	31 (31.3)	2.4 (1.48–3.81)
26.7–32.9	148	24 (16.2)	126	34 (27.0)	1.7 (1.04–2.65)
33.0–41.7	135	31 (23.0)	143	53 (37.1)	1.6 (1.11–2.35)
41.9–71.0	124	21 (16.9)	152	56 (36.8)	2.2 (1.40–3.38)
RAPID3 quartiles					
0.33–3.61	183	28 (15.3)	95	21 (22.1)	1.4 (0.87–2.40)
3.62–4.99	153	26 (17.0)	124	47 (37.9)	2.2 (1.47–3.38)
5.00–6.28	143	23 (16.1)	134	39 (29.1)	1.8 (1.14–2.86)
6.31–9.55	110	23 (20.9)	167	67 (40.1)	1.9 (1.28–2.89)

* Values are the number (%) unless indicated otherwise. For the active disease population, n = 1,109 visits. 95% CI = 95% confidence interval; ACR20 = American College of Rheumatology criteria for 20% improvement in disease activity; CDAI = Clinical Disease Activity Index; DAS28 = Disease Activity Score in 28 joints; MTC = major therapeutic change; RAPID3 = Routine Assessment of Patient Index Data 3; RR = risk ratio.

MTC was strongly associated with ACR20 response in the moderate and high categories of disease activity across all disease activity measures. In the active disease population, MTC consistently showed improvements in ACR20 response, but not all strata risk ratios met statistical significance, which is likely due to the smaller sample size for the active disease population.

DISCUSSION

This study was performed to evaluate the effect of MTC on ACR20 response criteria across 3 disease activity measures and corresponding levels of disease activity in a real-world clinical practice environment. The goal was to estimate the potential for clinical improvement when initiating MTC for patients with active RA in a population that is typically seen in clinical practice. In addition, we evaluated a subgroup of patients with active disease to mimic a population of RA patients that would qualify for inclusion in a clinical trial. We chose ACR20 response criteria to standardize the outcome assessment across the 3 disease activity measures, and for this reason we did not specifically look at whether MTC resulted in achievement or sustainment of low disease activity or remission. The findings are intended to shed light on the potential benefit that may be anticipated when escalating treatment in patients with active RA and established disease; this is a population that may not be considered for treatment modification. If the study had been conducted in a population that is typically enrolled in RA clinical trials (i.e., patients with shorter duration of RA and fewer prior treatments for RA), the ACR20 response rate likely would have been higher. Notably, many patients in our analysis still experienced clinical improvements despite longstanding disease and multiple prior therapies.

The present study showed that MTC resulted in consistent ACR20 responses in the population analysis across all disease activity measures. Nevertheless, disease activity level was an effect modifier on the full population and, as expected, MTC only showed clinical benefits in patient visits with active RA. The analysis in the full population should be viewed with caution, as it is sensitive to the underlying distribution of disease activity in our population and the selection process; additionally, nuances in how VARA providers collect core clinical measures likely resulted in a larger fraction of our study population in active RA compared to the true VARA population. For example, we identified 65% of eligible visits as having active RA based on the DAS28 compared to 58% before applying the exclusion criteria. Although VARA providers are encouraged to collect core clinical measures for each RA visit, they have discretion in how they collect and record these measures. VARA providers have reported greater leniency in recording core measures for patients with consistently low disease activity (personal communication); this report is consistent with VARA data. We observed that for half of the visits with patients in remission/low disease activity, there was a a follow-up visit with a recorded DAS28 within 168 days, while for half of the visits with patients in moderate/high disease activity there was a follow-up visits within 116 days. Since the population estimate is sensitive to the distribution of disease activity, the emphasis should be placed on the stratum-specific estimates in the full study population. In the active disease subpopulation, disease activity level did not appear to be a strong effect modifier, since the effect of MTC on ACR20 response

Table 5. Adjusted population (overall) and conditional (severity category) effects of MTC on risk of ACR20 response*

	Visits with MTC, % ACR20 response (95% CI)	Visits without MTC, % ACR20 response (95% CI)	RR (95% CI)
Full population			
DAS28 category			
Overall effect	15 (13–22)	8.5 (7.5–10)	1.77 (1.51–2.43)
Remission	1.1 (0.3–2.6)	0.8 (0.3–1.4)	1.36 (0.25–4.04)
Low	2.4 (0.7–4.8)	2.5 (1.6–3.9)	0.96 (0.23–2,43)
Moderate	16 (13–19)	9 (7–10)	1.81 (1.43–2.25)
High	35 (32–40)	19 (16–23)	1.82 (1.52–2.21)
CDAI category			
Overall effect	17.4 (16.9–18.7)	10.4 (10–11.9)	1.67 (1.43–1.91)
Remission	NA	NA	NA
Low	1.8 (0.6– 3.4)	1.6 (1.0–2.3)	1.12 (0.41–2.50)
Moderate	17 (14–20)	9 (8–11)	1.89 (1.51–2.46)
High	31 (27– 34)	19 (16–23)	1.60 (1.34–1.88)
RAPID3 category			
Overall effect	17 (15–19)	7.6 (6.7–8.5)	2.22 (1.88–2.57)
Remission	2.3 (0.5–5.9)	1.8 (0.7–3.3)	1.31 (0.27–4.33)
Low	8 (4–12)	4 (2–5)	2.13 (1.01–4.22)
Moderate	14 (11–17)	6 (5–8)	2.21 (1.60–2.99)
High	23 (21–26)	10 (9–12)	2.25 (1.91–2.68)
Active disease population			
DAS28 quartile			
Overall effect	34 (30–39)	22 (18–28)	1.52 (1.23–1.92)
2.35–4.87	25 (17–34)	16 (10–22)	1.59 (0.93–2.49)
4.88–5.60	33 (24–43)	20 (14–28)	1.64 (1.04–2.51)
5.61–6.29	30 (23–38)	23 (16–31)	1.30 (0.88–1.98)
6.30–8.76	48 (41–59)	30 (22–41)	1.57 (1.17–2.35)
CDAI quartile			
Overall effect	35 (31–40)	23 (19–31)	1.51 (1.18–1.90)
14.7–26.6	31 (22–42)	15 (10–21)	2.09 (1.28–3.40)
26.7–32.9	27 (20–35)	22 (15–32)	1.23 (0.83–1.95)
33.0–41.7	41 (34–50)	30 (23–40)	1.34 (1.03–1.84)
41.9–71.0	42 (35–52)	26 (18–39)	1.60 (1.14–2.53)
RAPID3 category			
Overall effect	34 (29–38)	21 (18–26)	1.60 (1.28–2.00)
0.33–3.61	22 (14–31)	18 (12–25)	1.23 (0.70–1.96)
3.62–4.99	39 (32–50)	21 (15–28)	1.86 (1.30–2.66)
5.00–6.28	31 (23–39)	20 (14–29)	1.53 (0.90–2.38)
6.31–9.55	43 (36–51)	25 (17–34)	1.69 (1.25–2.66)

* For the full population, n = 1,208 patients (6,138 visits); for the active disease population, n = 383 patients (1,109 visits). 95% CI = 95% confidence interval; ACR20 = American College of Rheumatology criteria for 20% improvement in disease activity; CDAI = Clinical Disease Activity Index; DAS28 = Disease Activity Score in 28 joints; MTC = major therapeutic change; NA = not applicable; RAPID3 = Routine Assessment of Patient Index Data 3; RR = risk ratio.

was fairly consistent across strata of disease activity quartiles; however, some levels did not reach statistical significance. This may be a result of the smaller sample size of this subset.

The RAPID3 has been reported to overestimate disease activity compared to the CDAI and the DAS28 (11). The RAPID3 classified ~15% more patient visits as active RA than the other disease activity measures, indicating these visits for MTC in a treat-to-target framework. Furthermore, when restricting the analysis to the active disease population, the RAPID3 classified ~6% of visits as remission or low disease activity. The effect estimates of MTC on ACR20 response were generally higher when classifying disease activity levels by the RAPID3. Our results therefore suggest that the RAPID3 may be a less suitable measure to guide treatment in a treat-to-target framework than other disease activity measures, although our study was not specifically designed to address this issue. The CDAI and DAS28 more consistently classified patients into similar levels of disease severity and reported similar findings in the marginal and stratum-specific estimates of MTC on ACR20 response.

Patient visits without MTC were more likely to have had MTC in the previous 90 days and were also more likely to have had a bDMARD or csDMARD dispensed in the month prior to the eligible visit. These observations may explain the relatively high proportion of ACR20 response in the group that did not receive

an MTC on eligible visits, as these patients may be deemed to be still improving with the recent medication change. This highlights the time-varying nature of treatment decisions in a treat-to-target framework. We attempted to adjust for previous MTC, stability of disease activity, and DMARDs dispensed in the previous month but did not explicitly model the time-varying treatment process. Nevertheless, a recent study that explicitly modeled the relationship between MTC and low disease activity status found that adjustments within 90 days of active RA resulted in shorter times to low disease activity status (6).

A strength of the study was the data source; the VARA registry represents a national data set of patients with rheumatologist-confirmed RA. The registry benefits from the use of a common medical record, consistent pharmacy data, and universal DMARD capture (27). There is consistent documentation of core clinical measures required to compute disease activity scores across the 3 validated disease activity measures evaluated in the present study. Notably, participants in the VARA registry represent the spectrum of RA disease activity, which allowed us to evaluate the subset of patients with clinical features similar to subjects enrolled in clinical trials.

Since we modeled disease activity and MTC as a fixed exposure rather than a time-varying process, it is possible that some patient visits met criteria for active RA even though the patient may have been considered to have stable disease by their variation on disease activity. Further work is needed to better identify patients who are likely to respond to MTC because of persistent active RA compared to patients who have fluctuations in disease activity status that represent stable disease even though they meet criteria for active RA. We were unable to fit baseline established medication exposure for bDMARD or csDMARD in all models, so the line of therapy for each patient was unknown; this could affect response to the next medication. MTC criteria were selected as indicators of an intervention that could be interpreted objectively as a change in therapy rather than maintenance of a treatment program and does not attempt to model the ACR treat-to-target recommendations for established disease (1,28). Future work could determine the impact of specific treatment choices based on levels of disease activity using dynamic treatment models. ACR20 response was chosen because it allowed us to compare response across the 3 disease activity measures. The results of this study indicate that MTC, as defined in this study, can invoke an initial clinical response in VARA patients with active RA who have established disease. This study does not provide information on whether MTC resulted in low disease activity or sustained response to treatment modification. Additional limitations included the predominance of men and the long duration of RA; these results may not be applicable to the broader population of RA patients. Finally, the sample size of the subset of patients with active disease was small.

In summary, patients in the VARA registry with longstanding and active RA were still able to achieve clinical improvement with an MTC. These results support the use of the treat-to-target strategy and provide a real-world framework for conducting observational studies that can mimic clinical trial populations. Given the known clinical benefits of the treat-to-target strategy for patients with RA, further studies to identify barriers and to enhance the use of this strategy are warranted.

ACKNOWLEDGMENTS

The authors thank Dr. Yi Zhang (Medical Technology and Practice Patterns Institute) for providing initial guidance and training on the use of direct standardization.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Sauer, Chen, Accortt, Collier, Cannon.
Acquisition of data. Sauer, Cannon.
Analysis and interpretation of data. Sauer, Chen, Shen, Accortt, Collier, Cannon.

ROLE OF THE STUDY SPONSOR

Amgen authors collaborated in the study design, provided writing assistance for the manuscript, and reviewed and approved the manuscript prior to submission. The authors independently collected the data, interpreted the results, and had the final decision to submit the manuscript for publication. Writing assistance was provided by Dikran Toroser (formerly for Amgen) and Julia R. Gage (on behalf of Amgen). Publication of this article was not contingent upon approval by Amgen.

REFERENCES

1. Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheumatol 2016;6–8:126.
2. Brinkmann GH, Norvang V, Norli ES, Grovle L, Haugen AJ, Lexberg AS, et al. Treat to target strategy in early rheumatoid arthritis versus routine care: a comparative clinical practice study. Semin Arthritis Rheum 2019;48:808–14.
3. Stoffer MA, Schoels MM, Smolen JS, Aletaha D, Breedveld FC, Burmester G, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search update. Ann Rheum Dis 2016;75:16–22.
4. Mikuls TR, Fay BT, Michaud K, Sayles H, Thiele GM, Caplan L, et al. Associations of disease activity and treatments with mortality in men with rheumatoid arthritis: results from the VARA registry. Rheumatology (Oxford) 2011;50:101–9.
5. Stever JR, Cannon GW, Teng CC, Accortt NA, Collier DH, Sauer BC. Utility of administrative and clinical data to predict major change in medical treatment in US veterans enrolled in the Veterans Affairs Rheumatoid Arthritis (VARA) registry. Clin Exp Rheumatol 2019;37:552–60.
6. Shaw Y, Chang CH, Levesque MC, Donohue JM, Michaud K, Roberts MS. Timing and impact of decisions to adjust disease-modifying antirheumatic drug therapy for rheumatoid arthritis patients with active disease. Arthritis Care Res (Hoboken) 2018;70:834–41.
7. Zak A, Corrigan C, Yu Z, Bitton A, Fraenkel L, Harrold L, et al. Barriers to treatment adjustment within a treat to target strategy in

rheumatoid arthritis: a secondary analysis of the TRACTION trial. Rheumatology (Oxford) 2018;57:1933–7.

8. DAS28: Radboud University Nijmegen Medical Centre; [March 21, 2018]. URL: https://www.das-score.nl/en/.
9. Aletaha D, Smolen JS. The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. Best Pract Res Clin Rheumatol 2007;21:663–75.
10. Pincus T, Swearingen CJ, Bergman M, Yazici Y. RAPID3 (Routine Assessment of Patient Index Data 3), a rheumatoid arthritis index without formal joint counts for routine care: proposed severity categories compared to disease activity score and clinical disease activity index categories. J Rheumatol 2008;35:213647.
11. Munoz JG, Giraldo RB, Santos AM, Bello-Gualteros JM, Rueda JC, Saldarriaga EL, et al. Correlation between rapid-3, DAS28, CDAI and SDAI as a measure of disease activity in a cohort of Colombian patients with rheumatoid arthritis. Clin Rheumatol 2017;36:1143–8.
12. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727–35.
13. Mikuls TR, Kazi S, Cipher D, Hooker R, Kerr GS, Richards JS, et al. The association of race and ethnicity with disease expression in male US veterans with rheumatoid arthritis. J Rheumatol 2007;34:1480–4.
14. Mikuls TR, Reimold A, Kerr GS, Cannon GW. Insights and implications of the VA Rheumatoid Arthritis Registry. Federal Practitioner 2015;32:24–9.
15. Fihn SD, Francis J, Clancy C, Nielson C, Nelson K, Rumsfeld J, et al. Insights from advanced analytics at the Veterans Health Administration. Health Affairs 2014;33:1203–11.
16. Verhoeven AC, Boers M, van Der Linden S. Responsiveness of the core set, response criteria, and utilities in early rheumatoid arthritis. Ann Rheum Dis 2000;59:966–74.
17. Fisher MC, Collier DS. Rheumatoid arthritis stable follow up visits - 3 month vs 6 month intervals. ACR/ARHP Annual Meeting; November 14-19; Boston, MA, 2014.
18. England BR, Sayles H, Mikuls TR, Johnson DS, Michaud K. Validation of the rheumatic disease comorbidity index. Arthritis Care Res (Hoboken) 2015;67:865–72.
19. Miriovsky BJ, Michaud K, Thiele GM, O'Dell JR, Cannon GW, Kerr G, et al. Anti-CCP antibody and rheumatoid factor concentrations predict greater disease activity in men with rheumatoid arthritis. Ann Rheum Dis 2010;69:1292–7.
20. Fransen J, van Riel PL. The disease activity score and the EULAR response criteria. Clin Exp Rheumatol 2005;23 Suppl 39:S93–9.
21. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44–8.
22. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. Am J Epidemiol 2011;173:731–8.
23. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period: application to control of the healthy worker survivor effect. Mathematical Modelling 1986;7:1393–512.
24. Greenland S, Holland P. Estimating standardized risk differences from odds ratios. Biometrics 1991;47:319–22.
25. Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. Am J Epidemiol 2004;160:301–5.
26. Liang K, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13–22.
27. Schwab P, Sayles H, Bergman D, Cannon GW, Michaud K, Mikuls TR, et al. Utilization of care outside the Veterans Affairs Health Care System by US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2017;69:776–82.
28. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010; 69:631–7.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 975–982
DOI 10.1002/acr.24223

American College of Rheumatology
Empowering Rheumatology Professionals

Association of Rheumatoid Arthritis in Pregnancy With School Performance of Offspring: A Danish Nationwide Register-Based Study

Signe S. Knudsen,[1] Julia F. Simard,[2] Jakob Christensen,[3] Thomas M. Laursen,[4] Bent W. Deleuran,[3] and Bodil H. Bech[4]

Objective. To examine the overall cognitive development of children exposed to maternal rheumatoid arthritis (RA) in utero by comparing their school test scores to those of their peers.

Methods. Children born in Denmark during 1995–2008 and listed in the National School Test Register were included (n = 738,862). Children exposed to maternal RA were identified through linkage of national registers. In separate analyses, exposure was subdivided according to maternal serostatus. Preclinical maternal RA was included as a separate exposure. The Danish national school tests are mandatory standardized tests. Results from all reading tests (grades 2, 4, 6, and 8) and mathematics tests (grades 3 and 6) from 2010–2017 were included. Test scores were compared according to maternal RA exposure for each test separately using linear regressions.

Results. We identified 934 children exposed to maternal RA in utero. There were no differences in reading test scores between maternal RA exposed and unexposed children. RA exposed children scored poorer in both mathematics tests (adjusted differences of mean score –0.14 SD (95% confidence interval [95% CI] –0.23, –0.06) and –0.16 SD (95% CI –0.26, –0.07). There was no appreciable difference between children by maternal RA serostatus. Children exposed to preclinical RA (n = 589) showed the same pattern of performance as children exposed to RA.

Conclusion. RA-exposed children scored slightly poorer in mathematics tests but performed as well as their unexposed peers in the reading tests. The results do not suggest that RA in pregnancy has a major impact on offspring school performance.

INTRODUCTION

Rheumatoid arthritis (RA) is one of the most prevalent autoimmune rheumatic diseases affecting young women, and an increasing number of children are born to women with RA (1). RA in pregnancy is a risk factor for adverse pregnancy and child outcomes, such as preterm birth and low birth weight (2). Children born to women with RA also have higher risks of adverse neurodevelopmental outcomes in childhood; recent studies found that in utero exposure to maternal RA was associated with increased risk of childhood epilepsy and autism spectrum disorders (3,4). Women with RA have raised levels of inflammatory cytokines and circulating autoantibodies (5), and they often receive medical treatment in pregnancy, factors potentially affecting the growing fetus' neurologic development. Increased risks of autism and learning disabilities have been reported among exposed offspring in other autoimmune rheumatic diseases (systemic lupus erythematosus and antiphospholipid syndrome) (6–9).

We therefore hypothesized that in utero exposure to maternal RA can affect children's neurocognitive development as measured by school performance. In this study, the academic performance of children born to mothers with RA was assessed using Danish nationwide registers and results from the Danish national school tests (10–13).

MATERIALS AND METHODS

Study population. All live-born, singleton children born in Denmark between January 1, 1995 and December 31, 2008

Supported by The Danish Rheumatism Association and Aarhus University.
[1]Signe S. Knudsen, MD, PhD: Aarhus University and Aarhus University Hospital, Aarhus, Denmark, and Stanford Medicine, Stanford, California; [2]Julia F. Simard, ScD: Stanford Medicine, Stanford, California; [3]Jakob Christensen, MD, PhD, DMSc, Bent W. Deleuran, MD, DMSc: Aarhus University Hospital and Aarhus University, Aarhus, Denmark; [4]Thomas M. Laursen, MSc, PhD, Bodil H. Bech, MD, PhD: Aarhus University, Aarhus, Denmark.

Dr. Christensen has received consulting fees, speaking fees, and/or honoraria from Eisai and UCB (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Signe S. Knudsen, MD, PhD, Department of Public Health, Aarhus University, Bartholins allé 2, 8000 Aarhus C, Denmark. Email: sig@ph.au.dk.

Submitted for publication November 5, 2019; accepted in revised form April 14, 2020.

SIGNIFICANCE & INNOVATIONS

- Children born to mothers with rheumatoid arthritis (RA) perform worse in Danish standardized mathematics tests on average than their peers.
- Children born to mothers with RA perform as well in Danish standardized reading tests as their peers.
- Children whose mothers have preclinical RA during pregnancy have a tendency to perform worse in Danish standardized mathematics tests than their peers.
- Children whose fathers have RA perform as well in Danish standardized reading and mathematics tests as their peers.

(n = 887,811) were identified from the Central Persons Register (14). Children who were deceased or had emigrated from Denmark before their eighth birthday, i.e., before the first school test, were excluded (n = 19,871), as well as children not listed in the Danish National School Test Register (DNT) (i.e., most children attending private schools or home schooled [n = 129,078]). The children were followed until December 31, 2017.

Maternal RA exposure. Data on maternal RA were collected from the Danish National Patient Register (15), which contains data on all inpatient visits since 1977 and outpatient visits since 1995 to Danish public hospitals. Diagnoses from each visit are recorded according to the International Classifications of Disease (ICD), using the eighth revision (ICD-8) before 1995 and the tenth revision (ICD-10) thereafter (16,17).

A child was considered exposed if the mother had been diagnosed with RA before the child's date of birth and had at least 2 separate visits or hospitalizations with a diagnosis of RA (ICD-8 codes 71259, 71239, and 71219, and ICD-10 codes M05 and M06). To assess possible effects of specific autoantibodies, maternal RA was also categorized into seropositive and seronegative RA on the basis of the mother's most frequently coded ICD-10 code (M05 or M06).

Exposure to maternal preclinical RA, defined as a diagnosis of RA within 3 years after the birth of the child, was also assessed. Paternal RA status before the child's birth was assessed using a similar strategy as for maternal RA. Exposure to paternal RA served as a negative control for the (psychosocial) effects of growing up with a chronically ill parent, without the intrauterine exposure to RA.

Danish national school tests. Mandatory tests were introduced in the Danish public school system in 2010, and 85% of Danish children attend public schools. These are standardized, adaptive tests based on a Rasch model, described in detail elsewhere (10,18).

Reading tests are performed 4 times (in grades 2, 4, 6, and 8), and mathematics tests twice (in grades 3 and 6). The reading tests contain 3 domains: language comprehension, decoding, and reading comprehension. Mathematics tests contain numbers and algebra, geometry, and mathematics in use. The test score is a measure of the child's skill level in the subject, and reported on a logit scale (–7; 7).

As the test scores for all tests followed normal distributions, the scores were standardized into Z scores, a measure of the relative position within the normal distribution of test results. Z scores have a mean value of 0, an SD of 1, and the units are SDs (i.e., a pupil who has a score of 0 is in the 50th percentile, and a pupil with a score of +1 is in the 84th percentile). Standardization was achieved by subtracting the mean and dividing by the SD for each specific test (subject/grade level/calendar year). Standardization makes it possible and appropriate to combine test results from multiple calendar years in the same analyses (18).

Exemptions from testing are given by teachers when it is considered meaningless or impossible for a child to perform a test. Reasons for exemption are recorded in the DNT in predefined categories, e.g., mental or physical impairment. Information on exemption was assessed according to exposure level to estimate the potential for selection bias caused by exemption.

Results from all 4 reading tests and both mathematics tests, from 2010 through 2017, were included, along with information on exemption from testing. Test data were provided by the Danish Agency for Information Technology and Learning (http: //www.stil.dk).

Covariates. Information on maternal educational level, maternal nationality (country of origin), and parental cohabitation status came from Statistics Denmark (19). The Central Persons Register (14) provided information on maternal age, and the Medical Birth Register provided information on maternal smoking in pregnancy as well as several birth outcomes (20). Information on children's birth defects and psychiatric diagnoses were collected from the Danish National Patient Register and the Psychiatric Central Register (21). Potential confounding covariates were chosen from a priori knowledge of the subject matter, incorporated into a directed acyclic graph of the study hypothesis (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24223/abstract) (22). Few covariates were considered true potential confounders and included in the models. This was due to the timing of the main exposure; the intrauterine exposure to maternal disease precedes exposures in the upbringing, thus these subsequent exposures cannot be confounders.

Statistical analyses. Low levels of missing data were present in several covariates (Table 1). Missing data were handled by multiple imputations using the R package "mice" (23) and using a chained equations approach, generating 15 complete data sets.

Multivariable linear regression models were used to compare the test Z scores of children exposed to maternal RA to those of unexposed children for each test separately (i.e., by grade level and subject). Exposure was further divided based on maternal

Table 1. Descriptive characteristics of 738,862 singleton children born in Denmark during 1995–2008 who were listed in the register for Danish National Tests, according to maternal rheumatoid arthritis (RA) at birth*

Child/parental characteristics	No maternal RA (n = 737,928)	Maternal RA (n = 934)
Sex		
Male	379,177 (51)	491 (53)
Female	358,751 (49)	443 (47)
Birth year		
1995–1999	264,337 (36)	247 (26)
2000–2003	218,611 (30)	241 (26)
2004–2008	254,980 (35)	446 (48)
Maternal age		
Age, mean ± SD years	30.01 ± 4.78	31.48 ± 4.67
Smoking in pregnancy		
Nonsmoker	447,651 (61)	636 (68)
Smoker	99,027 (13)	109 (12)
Missing	191,250 (26)	189 (20)
Child order		
First born	318,252 (43)	407 (44)
Second child or more	419,676 (57)	527 (56)
Maternal education		
≤9 years	153,760 (21)	183 (20)
10–14 years	358,451 (49)	416 (45)
≥15 years	208,149 (28)	324 (35)
Missing	17,568 (2)	11 (1)
Maternal nationality		
Danish	651,566 (88)	884 (95)
Other nationality	86,362 (12)	50 (5)
Maternal civil status		
Single	87,608 (12)	109 (12)
School tests		
Never tested	10,453 (1)	14 (1)
Birth variables		
BW, mean ± SD gm	3,535.6 ± 563.5	3,453.0 ± 603.2
BW <2,500 gm	24,778 (3)	44 (5)
BW, missing	6,330 (0.8)	≤5
GA, mean ± SD days	278.5 ± 12.6	276.4 ± 14.2
GA <37 weeks	36,044 (5)	62 (7)
GA, missing	5,354 (0.7)	≤5
Child diagnoses		
Malformations/ chromosomal abnormalities†	80,441 (10.9)	133 (14.2)
Any psychiatric diagnosis‡	95,163 (12.9)	123 (13.2)

* Values are the number (%) unless indicated otherwise. BW = birth weight; GA = gestational age.
† Defined as any Q diagnosis in the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10).
‡ Defined as any F diagnosis in the ICD-10.

RA serostatus, and Z scores from children exposed to seropositive maternal RA and seronegative maternal RA were compared to those of unexposed children separately. Z scores from children exposed to maternal preclinical RA (i.e., diagnosed within 3 years after birth) were also compared to those of unexposed children.

Results are reported as mean Z score differences with 95% confidence intervals (95% CIs), adjusted for maternal age, educational level, country of origin, and smoking in pregnancy. We accounted for dependence between siblings by clustering on maternal ID (using the "miceadds" package in R) (24). Bonferroni-corrected *P* values were calculated for all main analyses to account for multiple testing.

Nested mixed-effects models were used for analyzing Z scores from all grades and calendar years as repeated measurements, with Z scores nested within child and children within mothers, thus allowing for random effects at the child and mother levels, and performed separately for reading and mathematics (25).

Stratified analyses were performed, first stratifying by sex to assess possible differences between RA-exposed boys and girls, secondly on birth–year categories to evaluate time trends. Z scores from children exposed to paternal RA were compared to those of children unexposed to paternal RA using linear regression models, adjusted for the same factors as the main analyses.

To account for the potential selection bias inherent to the design, an inverse probability weighted model was fitted. A prediction model for "no school test result" was built using data from all singleton children born in Denmark between 1995 and 2008 (n = 887,811), including birth year, maternal education, age, smoking in pregnancy, nationality, parity, paternal nationality, family socioeconomic status, and any psychiatric referral of the child. Based on this prediction model, weights were calculated and applied to linear regression models for each test, inversely weighting each child's score by the likelihood of that child not having a test result. We also calculated adjusted odds ratios for being exempt from testing due to mental or physical impairment (as recorded by teachers in the DNT) to determine if risk of exemption was associated with exposure status, thus potentially leading to selection bias.

Data were collected and hosted at Statistics Denmark's servers. Owing to the privacy protection rules of Statistics Denmark, any values from an individual person or any calculated means based on <5 individuals cannot be reported. The study was approved by the Danish Data Protection Agency under Aarhus University comment agreement (j. number 2015-57-0002) and by Aarhus University (j. number 2016-051-000001, sequential number 737). All statistical analyses and graphs were made using R (26).

RESULTS

The final study population consisted of 738,862 children (i.e., all singleton children born in Denmark between 1995 and 2008 and listed in the DNT). A comparison of the final study population to the excluded children is provided in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24223/abstract. In total, 18% of children exposed to maternal RA were excluded due to not being listed in the DNT versus 15% of the unexposed children. Within the study population, 1.5% of children exposed to maternal RA were either exempt from testing or illegitimately missing (thus having no outcome measurement) versus 1.4% among unexposed children. The odds ratio for exemption specifically due to mental

Table 2. Results of Danish national school tests for 738,862 children from 2010–2017 by exposure to maternal rheumatoid arthritis (RA) in utero*

Test	Unexposed, no.	Exposed to maternal RA				
		No.	Crude	Adjusted†	*P*	*P*‡
Reading 2nd grade	394,979	592	0.06 (−0.02, 0.14)	−0.01 (−0.09, 0.08)§	0.88	1.00
Reading 4th grade	394,662	498	0.04 (−0.05, 0.12)	−0.03 (−0.12, 0.06)§	0.51	1.00
Reading 6th grade	390,741	437	−0.03 (−0.12, 0.08)	−0.08 (−0.17, 0.02)§	0.12	0.723
Reading 8th grade	355,796	351	0.06 (−0.04, 0.16)	−0.03 (−0.08, 0.12)§	0.62	1.000
Math 3rd grade	395,804	546	−0.08 (−0.17, −0.00)	−0.14 (−0.23, −0.06)§	0.001	0.006
Math 6th grade	390,015¶	431¶	−0.12 (−0.22, −0.03)¶	−0.16 (−0.26, −0.07)¶	0.001¶	0.003

* Values are the mean Z score difference (95% confidence interval) unless indicated otherwise.
† Adjusted for maternal age, maternal educational level, maternal nationality, and smoking in pregnancy.
‡ Bonferroni adjusted.
§ Mean Z score difference significant.
¶ Significant.

or physical disability (as stated by a teacher in the DNT) was 0.9 (95% CI 0.6, 1.4) for children exposed to maternal RA when adjusting for maternal age, education, nationality, and smoking in pregnancy compared to unexposed children.

In total, 934 (0.1%) children in our study population were born to mothers with RA. Mothers with RA were slightly older, had slightly longer education, were less likely to smoke during pregnancy, and more likely to be born in Denmark than mothers without RA (Table 1).

There was little difference in the unadjusted Z scores of children exposed to maternal RA compared to those of unexposed children; for all tests, see Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24223/abstract.

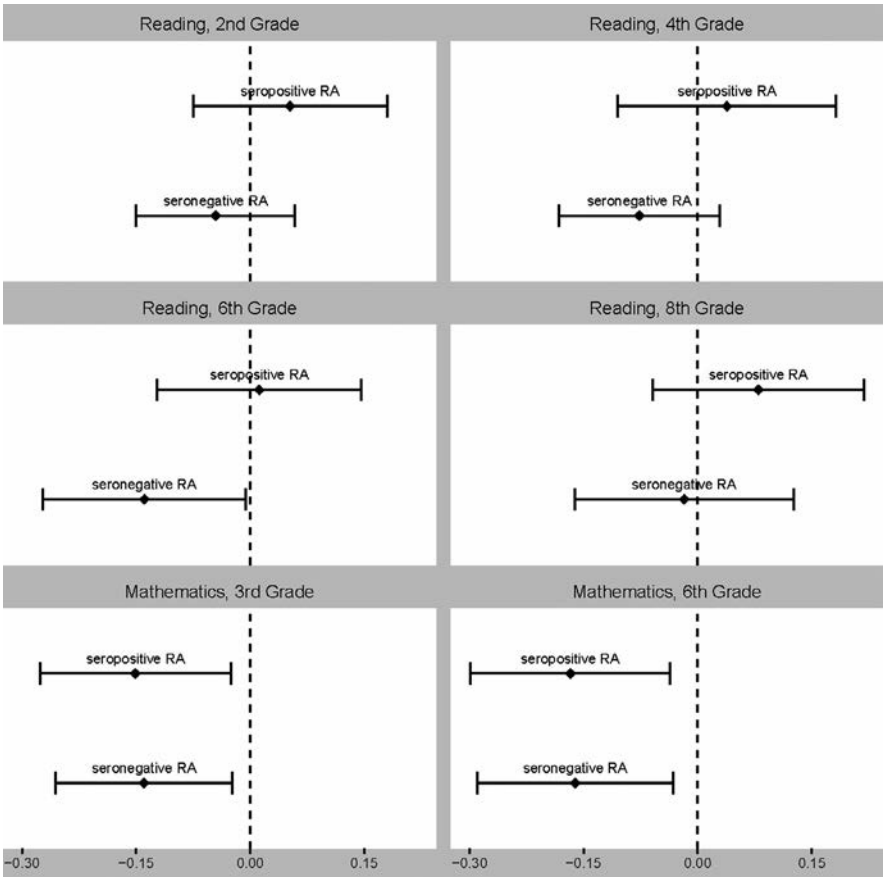

Figure 1. Adjusted differences of mean test scores in Danish national school tests by exposure to maternal seropositive rheumatoid arthritis (RA) or seronegative RA among 738,862 Danish children. Reference group was unexposed children (broken line). Adjusted for maternal age, maternal educational level, maternal nationality, and smoking in pregnancy. Diamonds represent the mean Z score difference. Error bars indicate the 95% confidence interval.

Adjustment for confounders did not appreciably change the interpretation for the reading tests; maternal RA was not associated with performance in any reading test (Table 2). However, adjusted analysis of the 2 mathematics tests found that maternal RA was associated with a poorer performance, with a difference of –0.14 SD (95% CI –0.23, –0.06) and –0.16 SD (95% CI –0.26, –0.07) for third and sixth grade, respectively (Table 2). Results did not change when combining the tests for all grade levels in mixed-effects models (adjusted differences for reading were 0.01 SD [95% CI –0.05, 0.08] and for mathematics –0.12 SD [95% CI –0.19, –0.05]).

When further classifying maternal RA into seropositive and seronegative RA, we found 397 children exposed to seropositive RA and 537 exposed to seronegative RA. Overall, no difference was observed between children exposed to seropositive RA and seronegative RA, although there was a trend toward children of mothers with seronegative RA performing worse in reading tests (Figure 1).

Among 589 children exposed to preclinical RA, we found a similar but attenuated pattern of test performance as for those exposed to maternal RA (Figure 2). There was no clear association between exposure to paternal RA and school test performance in any of the tests (see Supplementary Figure 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24223/abstract) (n = 424).

We found similar results for boys and girls in the analyses stratified by sex (see Supplementary Figure 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24223/abstract). Stratification by birth year categories did not change results (data not shown). The results from the inverse probability weighted analysis were consistent with the main results (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24223/abstract).

DISCUSSION

Danish children exposed to maternal RA performed similarly to their unexposed peers in standardized Danish reading tests but performed worse than their peers in mathematics tests regardless of maternal RA serostatus. However, the size of the difference was small, suggesting that the differences found may have little clinical

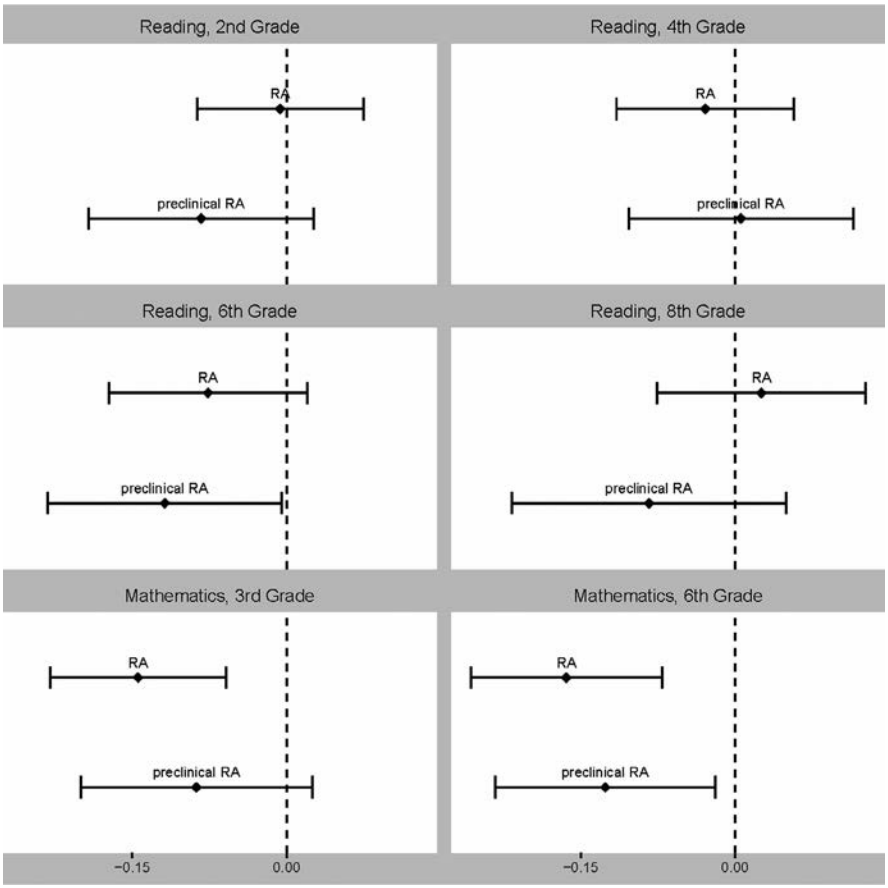

Figure 2. Adjusted differences of mean test scores in Danish national school tests by exposure to maternal rheumatoid arthritis (RA) or preclinical RA among 738,862 Danish children. Reference group was unexposed children (broken line). Adjusted for maternal age, maternal educational level, maternal nationality, and smoking in pregnancy. Diamonds represent the mean Z score difference. Error bars indicate the 95% confidence interval.

importance. On the percentile scale, the difference was roughly equivalent to the difference from the 50th percentile to the 45th percentile.

The children exposed to preclinical maternal RA likewise seemed to perform poorer in mathematics tests; however, the results were less clear, likely due to the smaller sample size. Exposure to paternal RA, which should not in itself influence the intrauterine environment, was not associated with poorer performance in either reading or mathematics tests. We found no indication of significant bias due to selection.

Very few other studies have assessed the long-term development of children exposed to maternal RA. In 2 other Danish population-based studies, Rom et al found a higher risk of childhood epilepsy and autism spectrum disorders in children exposed to maternal RA (3,27). The elevated risk of adverse neurodevelopmental outcomes did not seem to be explained by prematurity or low birth weight, suggesting that maternal RA may affect the intrauterine neurologic development of the growing fetus.

In a systematic review, Wojcik et al also found indications of higher risk of autism among children exposed to maternal RA; however, not all studies included supported this (28). No studies on neurocognitive development, IQ, or academic performance of children exposed to maternal RA were found.

In the current study, children exposed to maternal RA performed poorer on mathematics tests than their peers. This suggests a possible intrauterine effect on the child from either disease activity or medications or psychosocial effects in upbringing. The fact that children exposed to preclinical RA showed the same pattern of performance as children exposed to maternal RA suggests that intrauterine exposure to antirheumatic medications is not the main factor affecting mathematical performance, as women with preclinical RA are unlikely to receive any antirheumatic medication during pregnancy. Inflammatory cytokines and autoantibodies are known to precede manifest disease by several years (29), and the results may indicate that preclinical RA activity can affect the fetus in utero. However, given that these mothers were all diagnosed within the first 3 years of the child's life, potential biologic effects cannot readily be distinguished from the possible psychosocial effects of growing up with a chronically ill mother. We also cannot rule out that some preclinical RA patients already had symptoms causing them to take medications such as nonsteroidal antiinflammatory drugs and acetaminophen during pregnancy. Acetaminophen use in pregnancy especially has been associated with moderately lower IQ scores among 5 year-old children (30). Potential effects of prescription or over-the-counter medications warrants further investigations. If the association between preclinical maternal RA and school performance is an effect of growing up with a chronically ill parent, we might expect similar associations among children whose fathers have RA; however, that was not the case.

Previous studies have shown that Danish national test scores are strong predictors of final exam grades (10), which are in turn predictors of future education and income (31). IQ levels have been reported to be the strongest predictor of test scores in standardized school tests in international settings (32–34). However, this association has not been studied specifically for the Danish national tests, and any inference from test scores to psychometrics such as IQ levels must be made with caution. In a series of studies from Norway and Sweden, investigators found that another intrauterine exposure (radiation in the first trimester) was only associated with lower scores in mathematics tests, not in language tests, and also with lower scores in IQ tests and other neuropsychological tests (35–37). This might suggest that mathematics tests are more indicative of IQ levels than reading tests, but further research is needed to explore this.

This large, population-based study included >900 children exposed to RA from Danish national registers. The completeness and quality of the data are generally high (38). Several studies have validated the RA diagnoses in the Danish National Patient Register (39–41). Recently, the positive predictive value of a first ever hospital coding of an RA diagnosis was found to be ~70% among individuals <50 years of age (41). To further improve specificity, we required a minimum of 2 hospital contacts.

There was a risk of selection bias, as the study design inherently conditions on children having a measure of the outcome (i.e., being listed in the DNT). Approximately 15% of Danish children do not attend public schools and are not subjected to mandatory testing. The proportion of children exposed to maternal RA was roughly the same among children in the DNT as children not in the DNT. Also, the children exposed to maternal RA were no more likely to be exempt from testing due to mental or physical impairment, suggesting that they were no more prone to experience conditions that would result in exemption. Further, when we used inverse probability weighting to take potential sources for selection bias into account, we found the same results.

Another potential limitation of this study is the uncertainty of the outcome measurement, i.e., the test scores themselves. As for any test, any individual result might reflect more than just the student's skill, including factors such as motivation and concentration. However, although the test scores are uncertain at the individual level, they should be robust enough for comparison of large groups, as in the present study.

Unknown or residual confounding cannot be ruled out; however, we were able to include the potential confounders that we deemed most relevant to the association between maternal RA in utero and later school performance in offspring. No data were available on alcohol consumption in pregnancy; alcohol is a known risk factor for adverse cognition in offspring, but we adjusted for smoking and educational level, and may thereby indirectly have adjusted for other lifestyle factors and risk behaviors including alcohol.

The Danish national school tests are only used in Denmark; however, it is unlikely that Danish children exposed to maternal RA

are affected differently than children in other countries adhering to similar guidelines regarding care and treatment in pregnancy. We believe that the findings of this study are generalizable to other (high-income) populations.

In conclusion, children exposed to maternal RA in utero did not perform worse in Danish national reading tests when compared to their unexposed peers. However, a minor difference was found in mathematics tests, and further studies should focus on the importance of this finding in children exposed to RA. Overall, it is positive that there appear to be very little differences in the performance of children exposed to maternal RA and those unexposed, and no indications of major harmful effects from exposure to maternal RA were found.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Knudsen, Christensen, Laursen, Deleuran, Bech.

Acquisition of data. Knudsen, Bech.

Analysis and interpretation of data. Knudsen, Simard, Christensen, Laursen, Deleuran, Bech.

REFERENCES

1. Symmons D, Turner G, Webb R, Asten P, Barrett E, Lunt M, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology (Oxford) 2002;41:793–800.
2. Aljary H, Czuzoj-Shulman N, Spence AR, Abenhaim HA. Pregnancy outcomes in women with rheumatoid arthritis: a retrospective population-based cohort study. J Matern Neonatal Med 2020;33: 618–24.
3. Rom AL, Wu CS, Olsen J, Jawaheer D, Hetland ML, Christensen J, et al. Parental rheumatoid arthritis and childhood epilepsy: a nationwide cohort study. Neurology 2016;24:2510–6.
4. Atladottir HO, Pedersen MG, Thorsen P, Mortensen PB, Deleuran B, Eaton WW, et al. Association of family history of autoimmune diseases and autism spectrum disorders. Pediatrics 2009;124:687–94.
5. De Steenwinkel F, Hokken-Koelega A, De Man Y, De Rijke Y, De Ridder MA, Hazes J, et al. Circulating maternal cytokines influence fetal growth in pregnant women with rheumatoid arthritis. Ann Rheum Dis 2013;72:1995–2001.
6. Ross G, Sammaritano L, Nass R, Lockshin M. Effects of mothers' autoimmune disease during pregnancy on learning disabilities and hand preference in their children. Arch Pediatr Adolesc Med 2003;157:397–402.
7. Lahita RG. Systemic lupus erythematosus: learning disability in the male offspring of female patients and relationship to laterality. Psychoneuroendocrinology 1988;13:385–96.
8. Yousef Yengej F, van Royen-Kerkhof A, Derksen R, Fritsch-Stork R. The development of offspring from mothers with systemic lupus erythematosus: a systematic review. Autoimmun Rev 2017;16:701–11.
9. Nacinovich R, Galli J, Bomba M, Filippini E, Parrinello G, Nuzzo M, et al. Neuropsychological development of children born to patients with antiphospholipid syndrome. Arthritis Rheum 2008;59:345–51.
10. Beuchert L, Nandrup A. The Danish national tests at a glance. SSRN Electronic Journal. 2018. URL: https://doi.org/10.2139/ssrn.2996431.
11. Elkjær LS, Bech BH, Sun Y, Laursen TM, Christensen J. Association between prenatal valproate exposure and performance on standardized language and mathematics tests in school-aged children. JAMA Neurol 2018;75:663–71.
12. Dreier JW, Berg-Beckhoff G, Andersen PK, Andersen AM. Prenatal exposure to fever and infections and academic performance: a multilevel analysis. Am J Epidemiol 2017;186:29–37.
13. Skipper N, Gaulke A, Sildorf SM, Eriksen TM, Nielsen NF, Svensson J. Association of type 1 diabetes with standardized test scores of danish schoolchildren. JAMA 2019;321:484–92.
14. Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. Eur J Epidemiol 2014;29:541–9.
15. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT, et al. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol 2015;7:449.
16. Moriyama IM. International Classification of Diseases: eighth revision. WHO Chron 1968;22:43–7.
17. International Statistical Classification of Diseases and Related Health Problems, tenth revision. Geneva: World Health Organization; 2011.
18. Beuchert LV, Nandrup AB. The Danish national tests: a practical guide. Econ Work Pap 2014;25.
19. Baadsgaard M, Quitzau J. Danish registers on personal income and transfer payments. Scand J Public Health 2011;39:103–5.
20. Bliddal M, Broe A, Pottegård A, Olsen J, Langhoff-Roos J. The Danish Medical Birth Register. Eur J Epidemiol 2018;33:27–36.
21. Mors O, Perto GP, Mortensen PB. The Danish psychiatric central research register. Scand J Public Health 2011;39:54–7.
22. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology 1999;10:37–48.
23. Van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45:1–67.
24. Robitzsch A, Grund S, Henke T. Package "miceadds." R J 2019.
25. Twisk JW. Applied multilevel analysis: a practical guide for medical researchers. Cambridge: Cambridge University Press; 2006.
26. R Development Core Team. R: a language and environment for statistical computing. 2011. URL: http://www.R-project.org/.
27. Rom AL, Wu CS, Olsen J, Jawaheer D, Hetland ML, Mørch LS, et al. Parental rheumatoid arthritis and autism spectrum disorders in offspring: a Danish nationwide cohort study. J Am Acad Child Adolesc Psychiatry 2018;57:28–32.e1.
28. Wojcik S, Bernatsky S, Platt RW, Pineau CA, Clarke AE, Fombonne E, et al. Risk of autism spectrum disorders in children born to mothers with rheumatoid arthritis: a systematic literature review. Arthritis Care Res (Hoboken) 2017;69:1926–31.
29. Deane KD. Preclinical rheumatoid arthritis (autoantibodies): an updated review. Curr Rheumatol Rep 2014;16:1–16.
30. Liew Z, Ritz B, Virk J, Arah OA, Olsen J. Prenatal use of acetaminophen and child IQ: a Danish cohort study. Epidemiology 2016;27: 912–8.
31. Hvidtfeldt C, Tranæs T. Folkeskolekarakterer og succes på erhvervsuddannelserne. 2013. URL: https://www.rockwoolfonden.dk/app/uploads/2016/01/Arbpapir_29_Final.pdf.
32. Duckworth AL, Quinn PD, Tsukayama E. What no child left behind leaves behind: the roles of IQ and self-control in predicting standardized achievement test scores and report card grades. J Educ Psychol 2012;104:439–51.
33. Barton K, Dielman TE, Cattell RB. Personality and IQ measures as predictors of school achievement. J Educ Psychol 1972;63:398–404.
34. Mayes SD, Calhoun SL, Bixler EO, Zimmerman DN. IQ and neuropsychological predictors of academic achievement. Learn Individ Differ 2009;19:238–41.

35. Lie RT, Moster D, Strand P, Wilcox AJ. Prenatal exposure to Chernobyl fallout in Norway: neurological and developmental outcomes in a 25-year follow-up. Eur J Epidemiol 2017;32: 1065–73.
36. Heiervang KS, Mednick S, Sundet K, Rund BR. Effect of low dose ionizing radiation exposure in utero on cognitive function in adolescence. Scand J Psychol 2010;51:210–5.
37. Heiervang KS, Mednick S, Sundet K, Rund BR. The Chernobyl accident and cognitive functioning: a study of Norwegian adolescents exposed in utero. Dev Neuropsychol 2010;35:643–55.
38. Thygesen LC, Daasnes C, Thaulow I, Brønnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving. Scand J Public Health 2011;39:12–6.
39. Ibfelt EH, Sørensen J, Jensen DV, Dreyer L, Schiottz-Christensen B, Thygesen PH, et al. Validity and completeness of rheumatoid arthritis diagnoses in the nationwide DANBIO clinical register and the Danish national patient registry. Clin Epidemiol 2017;9:627–32.
40. Ibfelt EH, Jensen DV, Hetland ML. The Danish nationwide clinical register for patients with rheumatoid arthritis: DANBIO. Clin Epidemiol 2016;8:737–42.
41. Linauskas A, Overvad K, Johansen MB. Positive predictive value of first-time rheumatoid arthritis diagnoses and their serological subtypes in the Danish National Patient Registry. Clin Epidemiol 2018;10:1709–20.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 983–989
DOI 10.1002/acr.24236

American College of Rheumatology
Empowering Rheumatology Professionals

Precision Medicine With Leflunomide: Consideration of the *DHODH* Haplotype and Plasma Teriflunomide Concentration and Modification of Outcomes in Patients With Rheumatoid Arthritis

Michael D. Wiese,[1] Ashley M. Hopkins,[2] Catherine King,[1] Mihir D. Wechalekar,[3] Anita Lee,[4] Llewellyn Spargo,[5] Robert Metcalf,[5] Leah McWilliams,[5] Catherine Hill,[6] Leslie G. Cleland,[5] and Susanna M. Proudman[4]

Objective. Leflunomide is a commonly used disease-modifying drug in the treatment of rheumatoid arthritis (RA). Its effects are mediated via inhibition of dihydroorotate dehydrogenase (DHODH) by its active metabolite teriflunomide, and the pharmacokinetics of teriflunomide are highly variable. Our objective was to examine the association between the *DHODH* haplotype and plasma teriflunomide concentration with response to leflunomide in patients with RA where leflunomide was added to an existing disease-modifying drug regimen after failure to achieve an adequate response with conventional triple therapy.

Methods. Patients with RA who were taking, or were about to initiate, leflunomide were included. Participant characteristics, including the *DHODH* haplotype, were determined. Up to 5 plasma samples were collected after leflunomide was initiated for assays of total and free teriflunomide concentration. Disease activity was determined via the 28-joint Disease Activity Score (DAS28). The association between DAS28 scores and patient covariates was determined by linear mixed-effects modeling.

Results. A total of 67 patients were included in the study. The DAS28 score after initiation of leflunomide was associated with the baseline DAS28 score ($\beta = 0.70$, $P < 0.001$) and was higher in those who carried the *DHODH* haplotype 2 ($\beta = 0.56$. $P = 0.01$) and did not carry the shared epitope ($\beta = 0.56$, $P = 0.013$). As total and free plasma teriflunomide concentration increased, the DAS28 score was significantly lower ($P < 0.001$ and $P = 0.001$, respectively). When considering threshold concentrations, teriflunomide concentrations >16 mg/liter were associated with a DAS28 score that was 0.33 lower, and when free teriflunomide concentration was >35 μg/liter, the DAS28 score was 0.32 lower.

Conclusion. Teriflunomide concentration and carriage of the *DHODH* haplotype 2 are associated with response to leflunomide in patients with RA, and a total plasma teriflunomide concentration of at least 16 mg/liter is needed to maximize the likelihood of response.

INTRODUCTION

Rheumatoid arthritis (RA) is a potentially disabling form of arthritis associated with erosive joint destruction and subsequent deformity, pain, and increased mortality. Early referral and introduction of disease-modifying antirheumatic drugs (DMARDs) along with tight disease control using composite measures of disease activity and appropriate intensification of DMARDs (the treat-to-target approach) is the most effective treatment strategy, because rapid abrogation of inflammation is associated with long-term joint preservation and reduced disability (1–3).

Supported by Arthritis Australia (The Barbara Cameron Memorial Grant).
[1]Michael D. Wiese, B. Pharm, M. Clin. Pharm, PhD, Catherine King, M. Pharm, PhD: University of South Australia, Adelaide, South Australia, Australia; [2]Ashley M. Hopkins, B. Pharm (Hons), PhD: Flinders University of South Australia, Bedford Park, South Australia, Australia; [3]Mihir D. Wechalekar, MBBS, PhD, FRACP: Flinders University of South Australia and Flinders Medical Centre, Bedford Park, South Australia, Australia; [4]Anita Lee, MBBS, PhD, FRACP, Susanna M. Proudman, MBBS, FRACP: Royal Adelaide Hospital and The University of Adelaide, Adelaide Medical School, Adelaide, South Australia, Australia; [5]Llewellyn Spargo, MSc, Robert Metcalf, MNutrSc, PhD, Leah McWilliams, RN, BNg, Grad Dip NSc, Master of Nursing: Nurse Practitioner, Leslie G. Cleland, MBBS, PhD, FRACP: Royal Adelaide Hospital, Adelaide, South Australia, Australia; [6]Catherine Hill, MBBS, MD, MSc, FRACP: Royal Adelaide Hospital and The University of Adelaide, Adelaide, and The Queen Elizabeth Hospital, Woodville, South Australia, Australia.

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

Address correspondence to Michael D. Wiese, B. Pharm, M. Clin. Pharm, PhD, University of South Australia, UniSA: Clinical and Health Sciences, North Terrace, Adelaide, South Australia, Australia 5000. Email: Michael.wiese@unisa.edu.au.

Submitted for publication February 12, 2020; accepted in revised form April 21, 2020.

SIGNIFICANCE & INNOVATIONS

- After consideration of conventional patient factors such as age, seropositivity, and carriage of the shared epitope, response to leflunomide was associated with a patient's *DHODH* haplotype and plasma teriflunomide concentration.
- When considered in multivariate analysis, noncarriers of the *DHODH* haplotype 2 had a 28-joint Disease Activity Score (DAS28) that was 0.65 lower compared to carriers, and those with a plasma teriflunomide concentration >16 mg/liter had a DAS28 score that was 0.33 lower compared to those with concentrations below this threshold.
- The *DHODH* haplotype could be determined prior to disease-modifying antirheumatic drug selection as a method to identify patients more likely to benefit from leflunomide treatment. Once treatment has started, teriflunomide concentration should be measured to ensure that it is >16 mg/liter, which can inform decisions on whether to adjust the leflunomide dose when there is inadequate initial response or if it should be ceased due to futility.

A common first-line treatment option is single-agent methotrexate, but this treatment fails to achieve an adequate response in up to 85% of patients (4). In this setting, addition of a biologic DMARD (bDMARD) is advocated, particularly in individuals with poor prognostic features and high baseline disease activity (5,6). However, for patients who do not fall into this category, those where the cost of bDMARDs is prohibitive, or where methotrexate is contraindicated, addition of a second or alternative conventional synthetic DMARD (csDMARD) such as leflunomide is a common treatment strategy.

The therapeutic effects of leflunomide are mediated by its active metabolite teriflunomide, which is formed via both spontaneous conversion in blood and by hepatic metabolism via cytochrome P450 enzymes (CYP) 1A2 and 2C19 (7,8). Teriflunomide inhibits dihydroorotate dehydrogenase (DHODH), which reduces pyrimidine formation and therefore DNA synthesis in T lymphocytes, thereby decreasing their proliferation (9). Teriflunomide is highly bound to plasma protein (99.8%) (10) and has a half-life of 15 days (11). The lengthy half-life is likely due to enterohepatic recycling via ABCG2 transporter-mediated efflux into bile and reabsorption from the gastrointestinal tract (12).

Pharmacogenomic studies have shown that teriflunomide concentration is related to carriage of *CYP2C19*2* (13), *ABCG2* C421A genotype (14), and the *CYP1A2*1F* allele (15), while cessation of leflunomide due to toxicity is related to carriage of the *CYP1A2*1F* allele (16,17), the *CYP2C19* phenotype (18), and the *DHODH* genotype (19). Small studies have suggested that efficacy is related to the *DHODH* genotype (19) and haplotype (20). A concentration-efficacy relationship between total plasma teriflunomide concentration and selected disease activity metrics has been observed previously, although in each of these studies, leflunomide was used as a single agent in long-standing disease, response criteria were inconsistent, and the concentration associated with improved response varied between studies (13,21–23). None of these studies assessed the relationship between outcomes and concentration of free teriflunomide, the minor but biologically active fraction of plasma teriflunomide.

The aim of the current study was to determine, in a cohort of patients who were taking leflunomide after failure of csDMARD combination therapy, including methotrexate, the association between the *DHODH* haplotype and total and free plasma teriflunomide concentration with RA disease activity.

PATIENTS AND METHODS

Study population. Subjects enrolled in the RA inception cohort at the Royal Adelaide Hospital were eligible for inclusion in this study. At entry to the cohort, all participants were age >18 years with DMARD-naive RA, according to 2010 American College of Rheumatology criteria (24). RA was treated according to a previously published treat-to-target protocol, in which patients were initiated on triple csDMARD therapy (methotrexate, sulfasalazine, and hydroxychloroquine) (25). In those whose RA failed to respond to optimal dosing of triple therapy, leflunomide was added. When this study commenced, there were 2 groups of such patients: those who were already taking leflunomide and those who were yet to initiate leflunomide.

Data collection and follow-up. The following information was collected at entry to the cohort: sex, ethnicity, shared epitope and anti–citrullinated protein antibody (ACPA) and rheumatoid factor titers, age, height, weight, smoking status, and the dose of all DMARDs at the time of leflunomide initiation. The 28 swollen and tender joint count, patient assessment of disease activity (10-cm visual analog scale), and erythrocyte sedimentation rate were used to calculate the 28-joint Disease Activity Score (DAS28) (26).

Plasma samples were collected for determination of free and total teriflunomide concentrations. For subjects recruited into the study at the time of leflunomide initiation, samples were taken at up to 5 clinic visits, whereas for participants already taking leflunomide at recruitment, samples were taken at 2 consecutive clinic visits. The intent with this design was to obtain multiple samples from each participant and to achieve steady-state teriflunomide concentration at least once (steady state was assumed if the same leflunomide dose was taken for at least 8 weeks). At these plasma sample collection visits, the DAS28 score, leflunomide dose, and usage of other DMARDs were recorded.

Laboratory methods. Single-nucleotide polymorphism (SNP) genotypes for *DHODH* (rs3213422, rs3213423, rs2878404, rs8046916, rs11864453, and rs2288002) were determined using validated TaqMan SNP genotyping assays, carried out in

accordance with the manufacturer's instructions, and carriage of the *DHODH* haplotype 2 was determined as previously described (20). The total and unbound plasma concentration of teriflunomide was determined via liquid chromatograph–tandem mass spectrometry as previously described (10).

Statistical analysis. Analysis was performed to investigate the relationship between covariates and RA disease activity in response to leflunomide. Participants were included in the analysis if they had disease activity (as measured by the DAS28 score) determined at the commencement of leflunomide and at least once when a sample was taken to determine plasma teriflunomide concentration. Participants were excluded if they had a history of bDMARD usage (current or previous) and were withdrawn if leflunomide was ceased due to toxicity or if another DMARD was added due to active disease.

Correlations were represented with a Pearson's correlation coefficient (r^2) and 2-sided significance, represented by the *P* value. For analysis of the association between covariates and DAS28 scores, linear mixed-effects modeling was conducted that allowed for multiple observations per participant. Univariate analysis was initially conducted, and multivariate analysis was then performed via stepwise forward inclusion of variables, maintaining a *P* value less than 0.05. Since the baseline DAS28 score was so strongly associated with later measurements of DAS28 scores, the association between other variables was considered after baseline DAS28 scores were included in the model. Associations with teriflunomide concentration were represented by the effect size (β coefficient) and the statistical significance of the association (*P* value). For analysis of concentration cut points, at each of the observed concentrations, the data were dichotomized (into at or above the specified concentration or below the concentration), which were included in the linear mixed-effects model, and the effect size (with 95% confidence interval [95% CI]) was calculated. To determine whether a threshold effect was apparent, the effect size (and 95% CI) at each threshold was plotted against the total and free plasma teriflunomide concentrations. The threshold concentration associated with response was identified as the lowest concentration at which the entire 95% CI was associated with a reduction in disease activity. All statistical analyses were conducted using Microsoft Excel and Stata, version 14.2.

Ethics approval. Patients gave informed written consent for inclusion in the study, and ethics approval was obtained from the Human Research Ethics Committees of the Royal Adelaide Hospital and the University of South Australia.

RESULTS

A total of 67 patients, comprising 30 newly initiated patients and 37 patients who had been taking leflunomide prior to study entry, were included. At the time of leflunomide initiation, median disease activity was moderate, and approximately 60% of patients were taking concurrent triple csDMARD therapy (Table 1). Newly initiated and existing patients provided an average of 3.3 and 1.7 observations, respectively, for a total of 163 observations. Approximately 83% of observations occurred within a year of initiating leflunomide.

Of the 67 participants, 56 had ≥1 teriflunomide concentration that could be considered steady state. At the time of the first steady-state concentration, the average leflunomide dose was 15 mg/day, and the median steady-state total and free teriflunomide concentrations, respectively, were 19.3 mg/liter (interquartile range [IQR] 13.3–32.0) and 40.5 μg/liter (IQR 27.3–69.0). The total and free concentrations were closely correlated (r^2 = 0.874, $P < 0.0001$). There was a weak association between leflunomide dose and free (r^2 = 0.1066, P = 0.014) and total (r^2 = 0.0997, P = 0.0178) teriflunomide concentration (Figure 1).

Univariate analysis indicated that the DAS28 score after initiation of leflunomide ($DAS28_{lef}$) was positively correlated with the DAS28 score at the start of leflunomide treatment ($DAS28_{baseline}$) (β = 0.73, $P < 0.001$). After this effect was accounted for, $DAS28_{lef}$ was negatively correlated with the duration of DMARD treatment and was higher in those who carried the *DHODH* haplotype 2, were concurrently using methotrexate, and did not carry the shared epitope (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24236/abstract). No association was found with the duration of treatment with leflunomide (P = 0.19). Multivariate analysis demonstrated that $DAS28_{lef}$ was positively correlated with $DAS28_{baseline}$ (β = 0.70, $P < 0.001$), and $DAS28_{lef}$ was higher in

Table 1. Baseline participant characteristic (n = 67)*

Characteristic	Value
Age, median (IQR) years	53.0 (44.3–61.1)
Female	70.1
Height, median (IQR) cm	166 (160–173)
Weight, median (IQR) kg	73.3 (63.6–92.5)
Body mass index, median (IQR)	26.3 (23.3–29.1)
Caucasian ethnicity (n = 64)	82.8
History of smoking	44.8
Current smoker	26.9
ACPA positive (n = 63)	58.7
Rheumatoid factor positive (n = 64)	68.8
Shared epitope positive (n = 62)	67.7
DAS28 score, median (IQR)	4.2 (3.4–5.2)
DMARD treatment duration, median (IQR) weeks	88 (44–253)
Taking concurrent methotrexate	88.1
Taking concurrent sulfasalazine	71.6
Taking concurrent hydroxychloroquine	82.1
Taking concurrent triple therapy	59.7
Carrier of *DHODH* haplotype 2	59.7

* Values are the percentage unless indicated otherwise. ACPA = anti–citrullinated protein antibody; DAS28 = 28-joint Disease Activity Score; DMARD = disease-modifying antirheumatic drug; IQR = interquartile range.

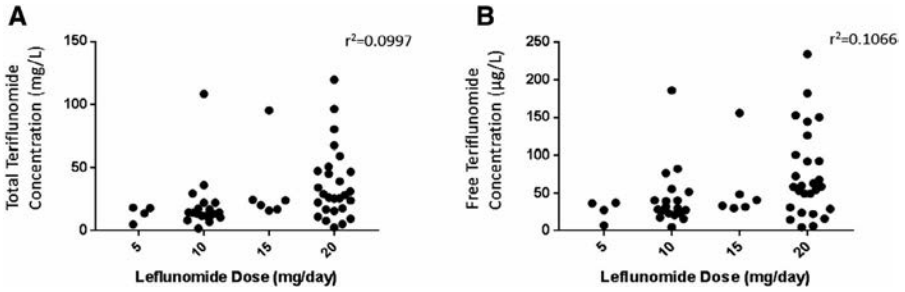

Figure 1. Total (**A**) and free (**B**) teriflunomide concentrations according to daily leflunomide dose. Circles represent individual participants.

those who carried the *DHODH* haplotype 2 (β = 0.56, *P* = 0.01) and did not carry the shared epitope (β = 0.56, *P* = 0.013).

When added to this model, the leflunomide dose was not associated with $DAS28_{lef}$ (β = 0.02 [95% CI –0.006, 0.046], *P* = 0.126), but there was a significant association of $DAS28_{lef}$ with total plasma teriflunomide concentration (β = –0.014 [95% CI –0.007, –0.021], *P* < 0.001) and free plasma teriflunomide concentration (β = –0.006 [95% CI –0.002, –0.009], *P* = 0.001). Sensitivity analysis that considered only steady-state total and free plasma teriflunomide concentrations (the participant was taking the same leflunomide for ≥8 weeks) showed similar associations with $DAS28_{lef}$ (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24236/abstract). These findings suggest that if total plasma teriflunomide concentration was to increase from 20 to 40 mg/liter, $DAS28_{lef}$ would be 0.28 lower, and if free plasma teriflunomide concentration was to increase from 40 to 80 µg/liter, $DAS28_{lef}$ would be 0.24 lower.

Every observed total and free teriflunomide concentration was investigated as a cut point, and all cut points for total teriflunomide concentration >16 mg/liter resulted in a significant relationship with $DAS28_{lef}$ (Figure 2). Above this cut point, $DAS28_{lef}$ was at least 0.33 lower, and if the cut point was increased to 35 mg/liter, $DAS28_{lef}$ was at least 0.56 lower. Of note, 47% and 80% of all observations had total teriflunomide concentrations below 16 and 35 mg/liter, respectively. Similarly, if a cut point for free teriflunomide concentration above 35 µg/liter was used, a significantly lower $DAS28_{lef}$ was always observed, whereby the $DAS28_{lef}$ was at least 0.32 lower. Furthermore, at cut points above 60 µg/liter, $DAS28_{lef}$ was improved by at least 0.5 compared to those patients who had a concentration <60 µg/liter. A total of 48% and 76% of all observations included free teriflunomide concentrations below 35 and 60 µg/liter, respectively.

DISCUSSION

The principal finding of this study is that both the *DHODH* haplotype 2 and teriflunomide plasma concentration (total and free) are associated with response to leflunomide in patients with RA that is refractory to treat-to-target therapy with combination csDMARDs. These associations were statistically significant after accounting for the effect of baseline disease activity and shared epitope positivity. The duration of treatment with leflunomide and a number of variables that are traditionally associated with response to DMARDs (e.g., seropositivity, tobacco smoking, sex) were not associated with $DAS28_{lef}$. Individuals who did not carry the *DHODH* haplotype 2 had a DAS28 score 0.65 lower than carriers, and the DAS28 score was at least 0.32 lower in those with total and unbound

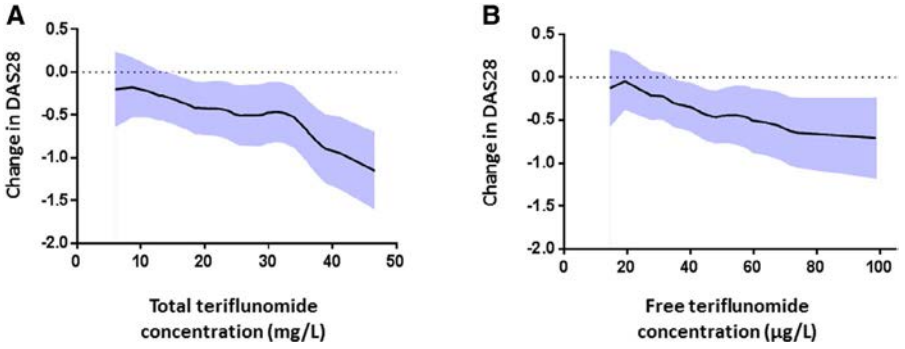

Figure 2. The association between the 28-joint Disease Activity Score (DAS28) after initiation of leflunomide ($DAS28_{lef}$) and various total (**A**) and free (**B**) teriflunomide concentration thresholds. At each observed concentration, data were dichotomized into at or above the concentration specified on the x-axis or below the concentration specified on the x-axis and included in a linear mixed-effects model. The y-axis represents the β coefficient, describing the effect size associated with concentrations at or above the specified threshold compared to concentrations below the threshold, with lower values indicating a greater reduction in $DAS28_{lef}$ relative to $DAS28_{baseline}$ scores. Black lines represent the estimated β coefficient, and shading indicates the 95% confidence interval.

plasma teriflunomide concentration above 16 mg/liter and 35 μg/liter, respectively. While the effect size describing the association between plasma teriflunomide concentration and response is modest, greater reductions in DAS28 scores were seen in individuals with higher plasma teriflunomide concentrations.

These findings suggest that patients with RA may be prospectively identified as being more (or less) likely to respond to treatment with leflunomide by considering the *DHODH* haplotype 2 carriage. The magnitude of effect in carriers compared to noncarriers is more than half the minimum clinically important improvement in DAS28 (27). This fact is particularly relevant, because we have previously reported no association between carriage of the *DHODH* haplotype 2 and leflunomide toxicity (17). The association that we have described between the *DHODH* haplotype 2 and $DAS28_{lef}$ is consistent with our prior report (20). The current study is substantially larger and is unique in that only time points where plasma teriflunomide concentration was measured were included in the analysis, and thus a comprehensive analysis of potential covariates was conducted.

Determining whether plasma teriflunomide concentration is above a proposed threshold could be useful in individuals who are taking leflunomide but have not experienced a satisfactory treatment response. This determination may be useful for a relatively large proportion of patients who take leflunomide, because in our cohort the first steady-state teriflunomide plasma concentrations were below this threshold in approximately 40% of patients. Leflunomide dose was only weakly associated with total and free teriflunomide concentration, and adherence to leflunomide therapy in these individuals was not determined. It is therefore unclear whether interventions to improve adherence or the use of previously defined genetic factors (such as carriage of *CYP2C19*2* [13], *ABCG2* C421A genotype [14], and the *CYP1A2*1F* allele [15]) would be more effective for selection of personalized doses of leflunomide that are more likely to achieve the threshold concentration. Equally important, plasma teriflunomide concentrations can help guide decisions on cessation of leflunomide due to futility. In patients who are responding well to leflunomide and are not experiencing toxicity but have a plasma teriflunomide concentration above the proposed threshold, our findings do not support reducing the leflunomide dose, because the observed reduction in DAS28 score increased further as the concentration rose even more above the proposed threshold (Figure 2).

To the best of our knowledge, this is the first study that has examined the effect of free teriflunomide concentration on disease activity in patients with RA who are taking leflunomide. Although we hypothesized that the free concentration would be more strongly associated with outcome than the total plasma teriflunomide concentration, both total and free plasma teriflunomide concentration were similarly related to $DAS28_{lef}$, which is likely due to the close relationship between total and free plasma teriflunomide concentration. Thus, the additional effort (and expense) needed to measure free concentration is unnecessary, and targeting a minimum total plasma teriflunomide concentration of 16 mg/liter is appropriate.

The current study has a number of important differences compared with the 4 prior studies that investigated a relationship between the concentration of teriflunomide in plasma and RA disease activity. The previous studies reported an association between a single concentration of teriflunomide and disease activity, and all but 1 were cross-sectional in design, whereas the current study was partly prospective, including data from the initiation of leflunomide in 45% of participants, and used up to 5 teriflunomide concentrations per patient. Cross-sectional study designs, particularly those that include patients who have received leflunomide for a long period of time, are limited because the cohort has become progressively enriched for those who are both responsive to and tolerant of leflunomide. This situation may bias the findings.

All prior studies were conducted in patients who were taking leflunomide as monotherapy and were conducted prior to common use of treat-to-target strategies, whereas in our study all participants took leflunomide in combination with other csDMARDs, and a treat-to-target strategy was employed. In 2 of the prior studies, only 1 or 2 markers were associated with teriflunomide concentration (i.e., C-reactive protein level in 1 study [13] and swollen joint count and the Short Form 36 health survey mental summary score in another [22]), whereas many other disease activity markers were not, and therefore multiple hypothesis testing may have contributed to the positive findings. Van Roon et al found that European Alliance of Associations for Rheumatology response criteria, but not DAS28, were associated with teriflunomide concentration (23), and Weber described an association between steady-state teriflunomide concentration and a Paulus response (21). While neither of these studies conducted multivariate analysis, both of these measures of response notably consider baseline disease activity when designating a patient as a responder, and our analysis found that $DAS28_{baseline}$ was strongly associated with $DAS28_{lef}$. Regardless, the total teriflunomide threshold concentration that we identified (16 mg/liter) was identical to that which was determined by van Roon et al (23), and similar to that determined by Weber (9) (13 mg/liter) (21).

In the current study, leflunomide was used in a contemporary manner according to treat-to-target principles (25). Therefore, the findings are likely to be relevant to clinical practice. However, our study did have some weaknesses. The results of our study represent associations rather than causality; for example, adherence to leflunomide was not assessed, and there is a possibility of confounding, because adherence is a behavior that extends beyond taking tablets and could result in both higher plasma teriflunomide concentrations and improved disease control via improved self-efficacy strategies and implementation of nonpharmacologic management strategies (28). Only 45% of participants were recruited at initiation of leflunomide, and thus the remaining 55% may have been enriched for responders to leflunomide when there was a substantial period since commencement of the

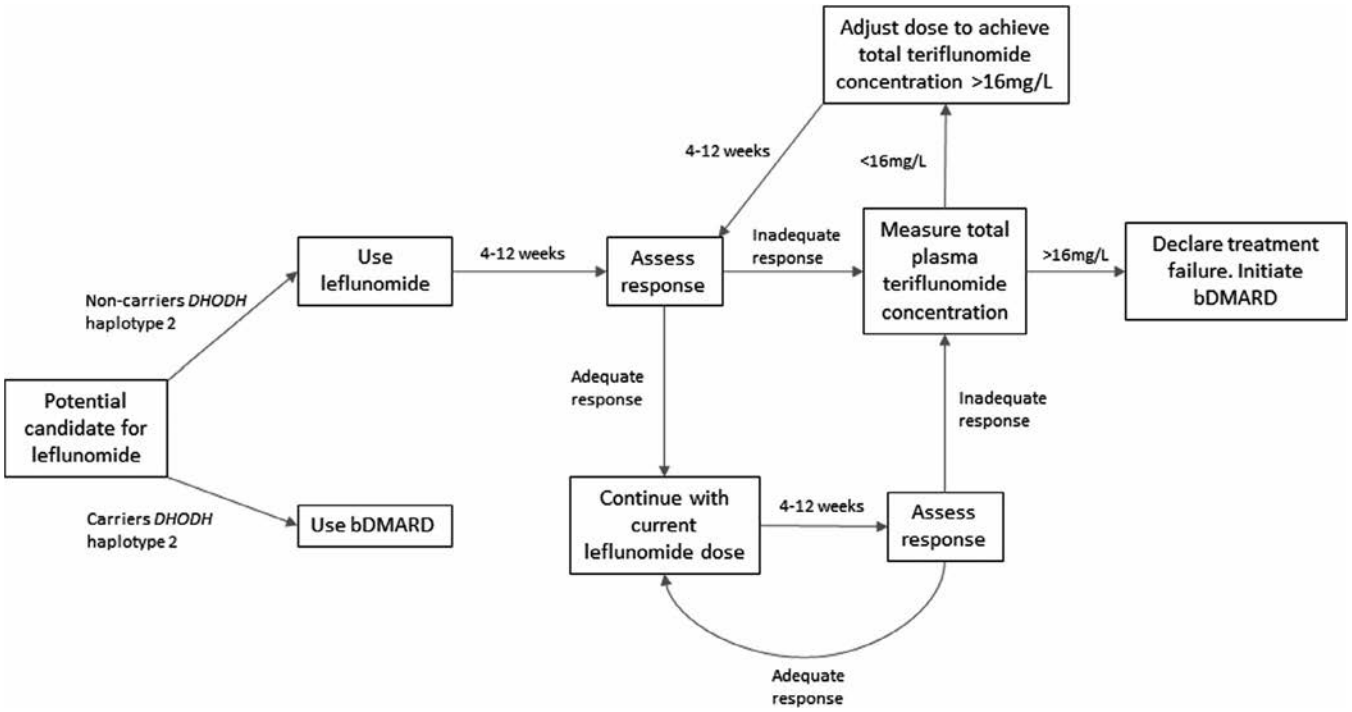

Figure 3. Potential predictive algorithm demonstrating the use of the dihydroorotate dehydrogenase (DHODH) haplotype status to select leflunomide or biologic disease-modifying antirheumatic drugs (bDMARDs) and for the use of plasma teriflunomide concentration to optimize response in patients for whom leflunomide treatment has been started.

drug. Recruiting patients who had taken leflunomide for a period of time may distort observed concentration-effect relationships. For example, patients with difficult-to-treat disease are more likely to require dose increases (and therefore they will have higher concentrations), whereas patients with highly responsive disease are less likely to require dose escalation, and therefore have lower plasma drug concentrations. Concentration-effect relationships are most likely to be apparent if patients are recruited shortly after leflunomide initiation (i.e., prior to commencement or at the time a first dose increase is scheduled).

In the current study >80% of samples were taken within 12 months of leflunomide initiation, and because no relationship of response to leflunomide treatment duration was apparent, any such effect appears to have been minimal. Furthermore, the statistical model assumes that any changes in teriflunomide concentration lead to an immediate change in disease activity, whereas more probably a lag to the observed treatment effect occurs after teriflunomide concentrations change. While this lag is likely to be most pronounced shortly after drug initiation or a change in dose, most patients provided multiple samples with a minimum of 3 weeks between sample collection, and so this effect should be minimized. Furthermore, analysis that included only samples that were at steady state showed a similar association between plasma teriflunomide concentration and disease activity.

While these findings suggest a more precise approach to leflunomide therapy, responses with leflunomide according to the *DHODH* haplotype 2 and teriflunomide concentration should be assessed in a prospective clinical trial, for example using the algorithm outlined in Figure 3. Such investigations could compare the efficacy of leflunomide in *DHODH* haplotype 2 noncarriers with bDMARDs in *DHODH* haplotype 2 carriers, and if the responses are similar, cost-effectiveness analysis could be conducted to determine cost savings that could be realized with widespread implementation. Furthermore, a prospective evaluation should be conducted to determine whether patients below the proposed concentration threshold will benefit from dose increases. A complementary approach may be to investigate the feasibility of selecting a personalized initial leflunomide dose to increase the likelihood of rapidly achieving the threshold concentration, and/or to examine leflunomide response and toxicity when teriflunomide concentrations are measured shortly after initiating leflunomide and doses are modified so that threshold concentrations are achieved more frequently and more rapidly.

In conclusion, our findings suggest that responses to leflunomide, as measured by the DAS28 score, can be predicted by considering baseline disease activity, shared epitope status, and carriage of the *DHODH* haplotype 2, and that further reductions may be realized by ensuring that all patients achieve a minimum total plasma teriflunomide concentration of 16 mg/liter.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Wiese, Hopkins, King, Cleland, Proudman.

Acquisition of data. Wiese, Hopkins, King, Wechalekar, Lee, Spargo, Metcalf, McWilliams, Hill, Cleland, Proudman.

Analysis and interpretation of data. Wiese, Hopkins, Proudman.

REFERENCES

1. Emery P, Breedveld FC, Dougados M, Kalden JR, Schiff MH, Smolen JS. Early referral recommendation for newly diagnosed rheumatoid arthritis: evidence based development of a clinical guide. Ann Rheum Dis 2002;61:290–7.
2. Schoels M, Knevel R, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas DT, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. Ann Rheum Dis 2010;69:638–43.
3. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Arthritis Rheum 2011;63:573–86.
4. Shammas RM, Ranganath VK, Paulus HE. Remission in rheumatoid arthritis. Curr Rheumatol Rep 2010;12:355–62.
5. Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, Kremer JM, et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64:625–39.
6. Smolen J, Landewe R, Bijlsma JW, Burmester G. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Ann Rheum Dis 2017;76:960–77.
7. Rozman B. Clinical pharmacokinetics of leflunomide. Clin Pharmacokinet 2002;41:421–30.
8. Kalgutkar AS, Nguyen HT, Vaz AD, Doan A, Dalvie DK, McLeod DG, et al. In vitro metabolism studies on the isoxazole ring scission in the anti-inflammatory agent leflunomide to its active alpha-cyanoenol metabolite A771726: mechanistic similarities with the cytochrome P450-catalyzed dehydration of aldoximes. Drug Metab Dispos 2003;31:1240–50.
9. Breedveld FC, Dayer JM. Leflunomide: mode of action in the treatment of rheumatoid arthritis. Ann Rheum Dis 2000;59:841–9.
10. Rakhila H, Rozek T, Hopkins A, Proudman S, Cleland L, James M, et al. Quantitation of total and free teriflunomide (A77 1726) in human plasma by LC-MS/MS. J Pharm Biomed Anal 2011;55:325–31.
11. Wiese MD, Rowland A, Polasek TM, Sorich MJ, O'Doherty C. Pharmacokinetic evaluation of teriflunomide for the treatment of multiple sclerosis. Expert Opin Drug Metab Toxicol 2013;9:1025–35.
12. Kis E, Nagy T, Jani M, Molnár E, Jánossy J, Ujhellyi O, et al. Leflunomide and its metabolite A771726 are high affinity substrates of BCRP: implications for drug resistance. Ann Rheum Dis 2009;68:1201–7.
13. Bohanec Grabar P, Grabnar I, Rozman B, Logar D, Tomsic M, Suput D, et al. Investigation of the influence of CYP1A2 and CYP2C19 genetic polymorphism on A771726 pharmacokinetics in leflunomide treated patients with rheumatoid arthritis. Drug Metab Dispos 2009;37:2061–8.
14. Kim KA, Joo HJ, Park JY. Effect of ABCG2 genotypes on the pharmacokinetics of A771726, an active metabolite of prodrug leflunomide, and association of A771726 exposure with serum uric acid level. Eur J Clin Pharmacol 2011;67:129–34.
15. Wiese MD, Hopkins A, Spargo L, et al. Patient, genetic and disease factors influence the response to the disease modifying anti-rheumatic drug leflunomide [abstract]. Arthritis Rheum 2014;66:S1102.
16. Bohanec Grabar P, Rozman B, Tomsic M, Suput D, Logar D, Dolzan V. Genetic polymorphism of CYP1A2 and the toxicity of leflunomide treatment in rheumatoid arthritis patients. Eur J Clin Pharmacol 2008;64:871–6.
17. Hopkins AM, Wiese MD, Proudman SM, O'Doherty CE, Upton RN, Foster DJ. Genetic polymorphism of CYP1A2 but not total or free teriflunomide concentrations is associated with leflunomide cessation in rheumatoid arthritis. Br J Clin Pharmacol 2016;81:113–23.
18. Wiese MD, Schnabl M, O'Doherty C, Spargo LD, Sorich MJ, Cleland LG, et al. Polymorphisms in cytochrome P450 2C19 enzyme and cessation of leflunomide in patients with rheumatoid arthritis. Arthritis Res Ther 2012;14:R163.
19. Grabar PB, Rozman B, Logar D, Praprotnik S, Dolzan V. Dihydroorotate dehydrogenase polymorphism influences the toxicity of leflunomide treatment in patients with rheumatoid arthritis. Ann Rheum Dis 2009;68:1367–8.
20. O'Doherty C, Schnabl M, Spargo L, Cleland LG, James M, Proudman SM, et al. Association of DHODH haplotype variants and response to leflunomide treatment in rheumatoid arthritis. Pharmacogenomics 2012;13:1427–34.
21. Weber WH. The population approach: measuring and managing variability in response, concentration and dose. In: Aaros LB, Danhof LP, Gex-Fabry M, Gundert-Remy M, Karlsson UA, editors. COST B1 medicine: European cooperation in the field of scientific and technical research. Brussels: European Commission; 1997. p. 238–44.
22. Chan V, Charles BG, Tett SE. Population pharmacokinetics and association between A77 1726 plasma concentrations and disease activity measures following administration of leflunomide to people with rheumatoid arthritis. Br J Clin Pharmacol 2005;60:257–64.
23. Van Roon EN, Jansen TL, van de Laar MA, Janssen M, Yska JP, Keuper R, et al. Therapeutic drug monitoring of A77 1726, the active metabolite of leflunomide: serum concentrations predict response to treatment in patients with rheumatoid arthritis. Ann Rheum Dis 2005;64:569–74.
24. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.
25. Proudman SM, Keen HI, Stamp LK, Lee AT, Goldblatt F, Ayres OC, et al. Response-driven combination therapy with conventional disease-modifying antirheumatic drugs can achieve high response rates in early rheumatoid arthritis with minimal glucocorticoid and nonsteroidal anti-inflammatory drug use. Semin Arthritis Rheum 2007;37:99–111.
26. Van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum 1998;41:1845–50.
27. Aletaha D, Funovits J, Ward MM, Smolen JS, Kvien TK. Perception of improvement in patients with rheumatoid arthritis varies with disease activity levels at baseline. Arthritis Rheum 2009;61:313–20.
28. Wabe N, Wiese MD. Treating rheumatoid arthritis to target: physician and patient adherence issues in contemporary rheumatoid arthritis therapy. J Eval Clin Pract 2017;23:486–93.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 990–997
DOI 10.1002/acr.24207

Relationship Between Pain and Sedentary Behavior in Rheumatoid Arthritis Patients: A Cross-Sectional Study

Helen O'Leary,[1] Louise Larkin,[2] Gráinne M. Murphy,[3] and Karen Quinn[3]

Objective. Despite the known benefits of physical activity, high numbers of individuals with rheumatoid arthritis (RA) remain physically inactive and sedentary. Little is known about the determinants of sedentary behavior (SB) in RA. This cross-sectional study was undertaken to examine a range of pain characteristics and RA-related symptoms and their relationship with objectively measured SB.

Methods. In total, 76 adults with RA wore an activPAL4 accelerometer (PAL Technologies) over a 7-day period. Pain characteristics (pain intensity, painful joint count, nonarticular pain), fatigue, sleep, depression, anxiety, and disease activity were assessed. Analyses were first conducted to evaluate correlations with sedentary time. The independent contribution of pain characteristics to variation in SB was analyzed with multivariable linear regression (adjusted for demographic data and disease activity).

Results. Participants with valid accelerometer data (n = 72) spent a mean ± SD of 533.7 ± 100.1 minutes/day in SB. Positive associations with daily SB were found for pain intensity (r = 0.31, $P < 0.01$) and number of painful joints (r = 0.24, $P < 0.05$) but not nonarticular pain. In multivariable analyses, pain characteristics were not independently associated with SB. Analyses indicated that disease activity had an indirect association with SB mediated by pain intensity. Other correlates of daily SB included anxiety and depression but not fatigue or sleep.

Conclusion. Results suggest that while pain and other RA-related factors do play a role in SB, they do not appear to have a significant influence after accounting for other variables. Future research should investigate SB and the role of factors unrelated to the symptoms of RA.

INTRODUCTION

Sedentary behavior (SB) is associated with poor health outcomes, including mortality, diabetes mellitus, and cardiovascular events in the general population (1). SB is defined as any waking behavior characterized by an energy expenditure of ≤1.5 metabolic equivalents and a sitting or reclining posture (e.g., television viewing, computer use, reading, and driving) (2). Evidence is emerging about the consequences of excessive sedentary time in the rheumatoid arthritis (RA) population, with indications that SB has a negative effect on patients' health (3). RA is a chronic, autoimmune, inflammatory condition that affects 0.5% of the adult population worldwide and affects 3 times more women than men (4). People with RA experience pain, stiffness, fatigue, and disability and are also at higher risk of cardiovascular and cerebrovascular disease compared to the general population due to the underlying inflammatory nature of the condition (5). Thus, the adverse consequences of excessive SB are likely to be even greater in individuals with RA, where SB has been reported to be ~10 hours per day (6), compared to an average of 8.7 hours for adults in the general population (7). Targeting individuals who spend the majority of their waking day in SB may have significant health benefits over and above those for the average population (8).

The majority of individuals with RA are physically inactive (9), having low aerobic capacity, and spending less time in vigorous activity compared to controls (10). Recent evidence suggests that SB and physical inactivity are separate constructs. Data on the factors that contribute to SB in individuals with RA are limited. However, the wider literature shows that the presence of RA-related symptoms, such as pain, stiffness, and fatigue, exert an important influence on daily activity levels (11,12). These clinical symptoms have been suggested as a potential explanation for more SB and less physical activity in individuals with RA (13). Pain is a key symptom in patients with RA and is a central component of diagnosis. Physical activity is known to alleviate symptoms of RA; a Cochrane review on exercise in RA has suggested that

[1]Helen O'Leary, BSc, MSc, PhD: Kerry University Hospital, Tralee, Ireland, and University of Limerick, Limerick, Ireland; [2]Louise Larkin, BSc, MSc, PhD: University of Limerick, Limerick, Ireland; [3]Gráinne M. Murphy, MB, BCh, BAO, MMedSC, PhD, Karen Quinn, BSc, MSc: Cork University Hospital, Cork, Ireland.

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

Address correspondence to Helen O'Leary, BSc, MSc, PhD, Physiotherapy Department, University Hospital Kerry, Tralee, County Kerry, Ireland. Email: Helen.oleary1@gmail.com.

Submitted for publication July 27, 2019; accepted in revised form March 31, 2020.

SIGNIFICANCE & INNOVATIONS

- It is known that individuals with rheumatoid arthritis (RA) spend the majority of their waking day in sedentary activities. No previous study has examined a range of pain characteristics and RA-related symptoms and their relationship with objectively measured sedentary behavior (SB).
- Pain intensity, self-reported number of painful joints, and the presence of foot/ankle pain were associated with daily sedentary time, but these relationships were relatively small and not maintained when other factors were accounted for in multivariable models.
- Other clinical correlates included mood and disease activity, while fatigue, sleep quality, or nonarticular pain were not related to SB. Pain intensity mediates the relationship between disease activity and daily sedentary time.
- Our findings suggest that while some clinical characteristics play a role, we must also look beyond RA-related symptoms in seeking to identify and understand the factors contributing to SB.

it can bring about moderate pain reductions (14). Furthermore, preliminary evidence indicates that reducing SB can achieve a reduction in pain levels in this population (8).

In order to address SB, a better understanding of how pain and other RA symptoms such as fatigue, sleep disturbance, and disease activity are related to SB is needed. Exploring related clinical factors could help target this behavior and enhance the delivery of focused and effective interventions for SB in patients with RA. To date, only 1 other study has explored the relationship between pain and SB in RA. Greene et al (15) focused exclusively on pain intensity and employed a self-report method of SB measurement. Such subjective measurements are known to be at risk of measurement bias due to misreporting and recall bias (1,16). Objective measurement of SB using accelerometers has become more feasible and affordable and, importantly, is more rigorous and addresses some of the limitations of self-report methods (1). Accelerometers are small, lightweight devices that record data on movement patterns continuously over several days (16).

No previous study in RA has investigated a range of patient-reported outcomes and has employed an objective measurement of SB. Thus, the aim of this study was to investigate the association between objectively measured SB and clinical pain characteristics and other patient-reported outcomes in RA, including sleep, fatigue, and mood. As the primary outcome measure, pain was explored under the dimensions of pain intensity, number of painful joints, and the presence of widespread pain. In addition to the primary aims, we also explored the extent to which pain intensity mediated the effect of disease activity on SB.

PATIENTS AND METHODS

Study design and participants. Potentially eligible participants for this cross-sectional study were identified consecutively from rheumatology clinics in a large acute public hospital serving a mix of urban and rural populations. Eligible participants were required to fulfill the following criteria: a diagnosis of RA by a rheumatologist according to the American College of Rheumatology criteria (17); age ≥18 and ≤80 years; ability to mobilize independently or aided by a stick; and ability to understand written and spoken English. Excluded participants had unstable disease (significant medication changes in past 3 months), a comorbidity interfering with their capacity to be physically active, recent surgery (in preceding 3 months), or were pregnant. Participating patients attended the rheumatology clinic, where a clinical assessment was carried out, questionnaires were completed, and an activPAL4 activity monitor (PAL Technologies) was fitted. Two clinical specialist physical therapists (HOL and KQ) undertook data collection between April and November 2018. Ethics approval was obtained from the Clinical Research Ethics Committee of the Cork Teaching Hospitals. All participants were provided with written and oral information about the research and gave written informed consent prior to study enrollment.

Clinical assessment. Demographic data on age, sex, disease duration, and medication were recorded. Body mass index (BMI) was calculated from height and weight measurements (kg/m^2). Participants' average arthritis pain intensity in the past week was quantified using the visual analog scale (VAS) pain scale (range 0–10 cm), which is reliable in patients with RA (18). The distribution of nonarticular pain was measured using the Widespread Pain Index (WPI), which assesses pain in 19 specific body areas (score range 0–19) (19). The distribution of joint pain was quantified using the joint score (range 0–48) from the Rheumatoid Arthritis Disease Activity Index (RADAI) (20). The Stanford Health Assessment Questionnaire determined participants' degree of functional disability (21). Anxiety and depression were assessed using the Hospital Anxiety and Depression Scale, which is a 14-item questionnaire validated in patients with physical health problems (22). A single item VAS (range 0–10) was used to assess the severity of fatigue over the past week (23). The Pittsburgh Sleep Quality Index measured sleep disturbance (score range 0–21) and is a validated measure of sleep disturbances among individuals with chronic pain (24). The Clinical Disease Activity Index (CDAI) (25) was used to reflect RA disease activity. This patient and provider composite tool can discriminate between low, moderate, and high disease activity states and is feasible to perform in clinical settings (26). The Charlson comorbidity index (27) was used to quantify the comorbidity burden by assessing the number and severity of 13 health conditions.

SB was measured over a 7-day period using the activPAL4 activity monitor. The activPAL4 is an objective measurement

device that incorporates an inclinometer to facilitate greater accuracy in classifying postures. Validated for time spent in SB, standing, and walking (28), the activPAL4 activity monitor is recommended for accurately recording SB in the RA population (29). There is good agreement ($R^2 > 0.94$) for sedentary time between activPAL4 data and direct observation in a free-living setting (30).

Participants were instructed to wear the monitor continuously 24 hours/day for 7 consecutive days (16) and only remove it if swimming. A 24-hour wear period is common in the published literature and may be associated with better device wear-time compliance (16). The device was wrapped in a flexible sleeve and attached by the physical therapist to the midline upper aspect of the anterior thigh using waterproof adhesive dressing (Tegaderm). Participants were provided with an instruction sheet, spare adhesive dressings (for reattachment if necessary), and a log sheet. Participants recorded any non-wear periods in the log sheet as well as daily times for going to bed and getting up, allowing for isolation of waking sitting/lying time from sleep time. When participants failed to complete the log sheet, non-wear or bedtimes were visually identified from the events files. Only participants who wore the activity monitor a minimum of 4 days for 24 hours were included in the analysis.

Data were extracted using the PAL software, version 8.1, and events files were created. The sampling frequency was 20 Hz, and the minimal sitting and upright period was defined as 10 seconds. SB characteristics of interest were as follows: time sitting or lying during the waking day; the percentage of waking hours spent sitting or lying; and sit-to-upright transitions and number of sitting bouts longer than 30 minutes. Subjectively, SB was assessed using the following item from the International Physical Activity Questionnaire short form (31): "During the past 7 days, how much time did you usually spend sitting on a weekday?" Responses were given in hours and minutes.

Data analysis. Data were analyzed for normality, homogeneity of variance, and multicollinearity with all required assumptions met. Descriptive statistics were calculated to describe participant characteristics. Relationships between clinical characteristics and SB variables were examined using Pearson's correlation coefficients. Correlation coefficients >0.5 were defined as high, those from 0.3 to 0.5 were defined as moderate, and those <0.3 were defined as fair. To explore the relevance of foot/ankle pain, an independent *t*-test was used to analyze differences in daily sedentary time between those with and without foot and/or ankle pain (assessed using the RADAI joint list).

To further examine any significant associations between pain characteristics and SB, the independent contributions of these pain variables to explain variance in daily sedentary time were analyzed in a series of multivariable regression models. Associations were analyzed in models that controlled for demographic data (age, sex, and BMI) and disease activity. Selection of these explanatory predictor variables was based on previous literature (32,33). Explanatory variables were entered first, with the pain variable included in the second step to assess any additional contribution to the models explaining SB. The variance in SB explained by pain variables was determined by examining change in R^2 values between the first and second steps. Where significant associations were observed in initial analyses, models were further adjusted for daily moderate–vigorous physical activity (MVPA). This final step (adjustment for MVPA) was carried out only where step 2 of the regression analysis revealed a statistically significant relationship between pain and SB. The levels of association were expressed as standardized β coefficients. This study required 85 participants (with 80% power at 5% level of significance) to detect an effect size of r = 0.3 for pain and other patient-reported outcomes.

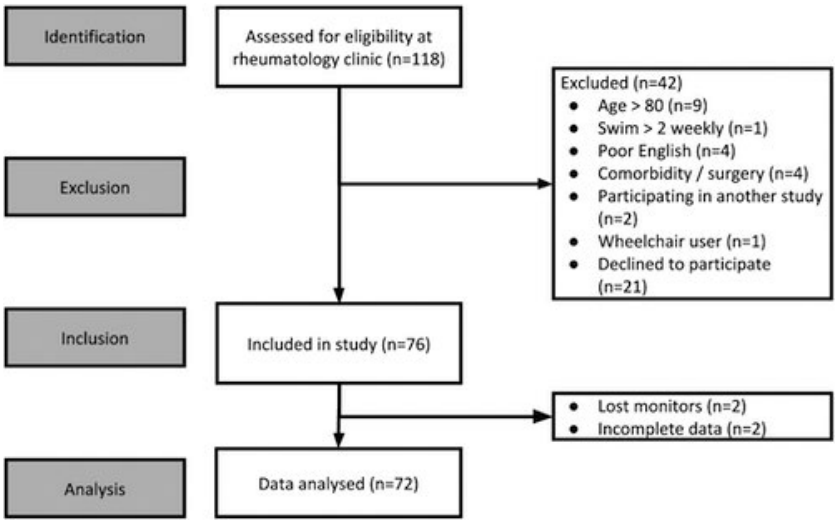

Figure 1. Flow diagram of study participants.

Mediation analysis was undertaken to examine associations between disease activity, pain intensity, and SB according to criteria described by Baron and Kenny (34). Disease activity was included as the independent variable, daily sedentary time as the dependent variable, and pain intensity as mediator. According to these criteria, linear regression models were used to assess whether 1) the independent variable was significantly associated with the dependent variable, 2) the independent variable was significantly associated with the proposed mediator, and 3) the proposed mediator was significantly associated with the dependent variable, with the independent variable as a control variable. To confirm these results, the proposed mediator was also assessed using Sobel's test and a bootstrapping approach utilizing 5,000 bootstrap samples.

RESULTS

Participants. RA patients with upcoming clinic appointments were screened for participation (n = 118). A review of medical notes and telephone-based screening for eligibility was conducted. In total, 76 patients were enrolled in the study (Figure 1), and 72 participants returned valid activPAL4 data. Participants were younger (60 versus 61 years), had longer disease duration (18 versus 10 years), and a greater proportion were men (35% versus 29%) compared to patients who declined to participate in this study (all $P > 0.05$). Of the 72 participants, 90% returned monitors with 7 days of valid data.

Participant characteristics (n = 72) are presented in Table 1. Participants' mean ± SD age was 61.5 years ± 10.5, and 65% (n = 47) were female. The average ± SD time since diagnosis was 17.8 ± 10.9 years. Disease activity score averaged 11.2 ± 8.7 on CDAI scores, representing borderline mild-to-moderate disease activity. The proportion taking biologic agents to treat their disease was 58.7%. The mean ± SD arthritis pain score on a VAS was 4.9 ± 2.9. Accelerometer data indicated that participants spent a mean ± SD 533.7 ± 100.1 minutes per day (8.9 hours, 59.9% of waking hours) in SB. Participants' subjective estimate of their sedentary time was 5 hours.

Associations between clinical characteristics and sedentary behavior. Table 2 outlines Pearson correlations between clinical and SB characteristics. Positive associations with daily SB (time) were found for pain intensity (r = 0.31, $P < 0.01$) and self-reported number of painful joints (r = 0.24, $P < 0.05$) but not nonarticular pain, as quantified by the WPI (r = 0.06, $P > 0.05$). Pain characteristics were not correlated with the number of sedentary bouts >30 minutes or the number of sedentary interruptions. Daily SB also had positive but fair associations with self-reported depression (r = 0.28, $P < 0.05$), anxiety (r = 0.31, $P < 0.01$), and disease activity (r = 0.24, $P < 0.05$). In an independent *t*-test, daily SB was higher among those with foot and/or ankle pain (mean ± SD 552.7 ± 104.0 minutes) compared to those without (mean ± SD 496.5 ± 83.8 minutes, $P < 0.05$).

Table 1. Participant demographic, clinical, and accelerometer data (n = 72)*

Characteristic	Value
Demographic and clinical characteristic	
Age, years	61.5 ± 10.5
Female, no. (%)	47 (65)
BMI, kg/m^2	28.9 ± 4.5
Time since diagnosis, years	17.8 ± 10.9
Disease activity (CDAI)	11.2 ± 8.7
DMARDs, no. (%)	48 (68)
Biologics, no. (%)	44 (58.7)
Comorbidity score (CCI)	2.3 ± 1.3
Activity limitation score (HAQ)	0.8 ± 0.5
Pain intensity score (VAS)	4.9 ± 2.9
Painful joint count (RADAI)	9.6 ± 7.0
Widespread pain score (WPI)	4.9 ± 3.9
Fatigue score (VAS)	5.7 ± 2.9
Anxiety score (HADS)	5.7 ± 3.8
Depression score (HADS)	5.4 ± 3.8
Sleep score (PSQI)	7.2 ± 5.0
Self-report sedentary time, hours/day	5.0 ± 2.0
Accelerometer data	
Daily sedentary time, minutes/day	533.7 ± 100.1
% waking time spent sedentary	59.9 ± 11.0
Sedentary bouts >30 minutes, no./day	6.7 ± 1.7
Sedentary interruptions, no./day	52.3 ± 17.9
Standing time, minutes	271.8 ± 86.0
Daily total step count	7,210.1 ± 3,684.0

* Values are mean ± SD unless indicated otherwise. BMI = body mass index; CCI = Charlson comorbidity index; CDAI = Clinical Disease Activity Index; DMARDs = disease-modifying antirheumatic drugs; HADS = Hospital Anxiety and Depression Scale; HAQ = Health Assessment Questionnaire; PSQI = Pittsburgh Sleep Quality Index; RADAI = Rheumatoid Arthritis Disease Activity Index; VAS = visual analog scale; WPI = Widespread Pain Index.

Regression analyses. Significant positive associations between pain variables and daily sedentary time were no longer significant in regression models adjusted for demographic factors and disease activity (Table 3). The addition of pain variables to the adjusted models accounted for a small increase (≤5%) in the variance (R^2) of daily SB.

Mediation analysis. Tests for mediation found that higher disease activity was associated with daily SB ($\beta = 0.24$, $P = 0.046$) and higher pain intensity ($\beta = 0.622$, $P < 0.001$). Higher pain intensity was significantly associated with greater SB (time) while controlling for the disease activity score ($\beta = 0.314$, $P < 0.036$). The Sobel's test result was 2.05 ($P = 0.041$), indicating that pain intensity mediates the relationship between disease activity and total sedentary time. This result was confirmed by the bootstrapping approach, where the 95% confidence interval (95% CI) for the indirect effect of disease activity on SB level, mediated by pain levels, did not overlap with zero (bootstrap 95% CI 0.37–4.66).

Table 2. Correlations between clinical and accelerometer-assessed sedentary behavior*

	Daily sedentary time, minutes	Waking time spent sedentary, %	Sedentary bouts >30 minutes	Sedentary time interruptions	Total standing time
Pain intensity score (VAS)	0.31†	0.35†	0.16	0.18	−0.41†
Painful joint count (RADAI)	0.24‡	0.22	0.17	0.09	−0.20
Widespread pain score (WPI)	0.06	0.10	0.03	0.08	−0.12
Fatigue score (VAS)	0.17	0.19	0.06	0.08	−0.28‡
Sleep score (PSQI)	0.13	0.14	0.12	−0.01	−0.21
Anxiety score (HADS)	0.31‡	0.23	0.14	−0.09	−0.23
Depression score (HADS)	0.28‡	0.24‡	0.09	−0.01	−0.22
Disease activity score (CDAI)	0.24‡	0.26‡	0.05	0.01	−0.34†

* CDAI = Clinical Disease Activity Index; HADS = Hospital Anxiety and Depression Scale; PSQI = Pittsburgh Sleep Quality Index; RADAI = Rheumatoid Arthritis Disease Activity Index; VAS = visual analog scale; WPI = Widespread Pain Index.
† $P < 0.01$.
‡ $P < 0.05$.

DISCUSSION

This cross-sectional study aimed to examine the relationship between SB and pain in patients with RA. We found that pain intensity, self-reported painful joint count, and the presence of foot and/or ankle pain were correlated with a greater volume of time spent in sedentary activities. Multivariate models, when adjusted for potential explanatory variables, found that these pain characteristics were not independently associated with daily SB. Other clinical correlates of daily SB included depression, anxiety, and disease activity. Mediation analysis revealed that pain intensity mediates the relationship between disease activity and daily sedentary time.

Evaluating the links between SB and disease-specific outcomes was identified by the European Alliance of Associations for Rheumatology (EULAR) as a key aspect of the future research agenda (35). To our knowledge, this is the first study to investigate the relationship between several dimensions of pain in patients with RA and objectively measured SB. Two previous studies in RA have examined the relationship between pain intensity and SB. Pain intensity was not independently associated with time spent sitting and lying in a sample of predominantly African American female patients with RA (15). Multivariate analysis suggested that sedentary time was best explained by a combination of variables that included pain; however, the amount of variance explained by the model was small, at just 10%. Comparisons with the current study are limited, as sedentary time was determined by individual interview rather than objective measures. Huffman et al objectively measured SB in patients with established RA using triaxial accelerometry. In contrast to our findings, there was no correlation with pain intensity; however, participants' pain levels were relatively low (mean intensity 25 of 100 on a VAS) (36). In the current study, while pain intensity and self-reported joint count were associated with sedentary time, the strength of correlations was only fair. This weak link between pain and sedentary time was somewhat unexpected. Qualitative research has linked higher pain levels with more bad days and, consequently, more sitting (13). Notably, individuals with RA have also described other aspects of life unrelated to RA as playing a role in SB. For some, being sedentary was described as "simply a way of living" (13). These accounts are in line with our findings, suggesting that there are other more influential determinants of SB than clinical and pain characteristics.

Table 3. Regression analyses investigating association between pain variables and daily sedentary time*

	Independent variable	β (*P*)	R^2
Model 1	Pain intensity	0.29 (0.08)	0.05
Model 2	Painful joint count	0.09 (0.57)	0.01
Model 3	Ankle/foot pain	0.18 (0.19)	0.03

* β = beta coefficient. R^2 represents the variance explained in the dependent variable (daily sedentary time) by the independent variable of interest (i.e., pain intensity, painful joint count, ankle/foot pain). Models were adjusted for age, sex, body mass index, and disease activity.

Individuals with RA experience inflammatory joint pain but also report higher levels of nonarticular chronic widespread pain (37). No previous study has differentiated between articular and nonarticular pain when examining the relationship between pain and SB in RA. While fibromyalgia was not specifically examined, 13 participants had widespread nonarticular pain (score ≥7 on the WPI), which is suggestive of this condition; however, WPI score was not a correlate of any aspect of SB. As centralized chronic widespread pain tends to coexist with mood disorders and fatigue (38), it might be expected that such a clinical profile would be more sedentary, but this was not the case. Pain is frequently used as a proxy for inflammation and disease activity in the assessment of RA (39). While the measure of disease activity in this study (the CDAI) did not specifically incorporate a self-reported painful joint count or pain severity, patients' global assessment of their arthritis likely reflects, to some extent, pain currently experienced. Our mediation analysis confirmed what might be assumed clinically: that the indirect effect of disease activity on SB is mediated by arthritis pain levels. Relations between SB and clinical factors such as pain and mood are complex, and additional mediated

effects not examined in this study are likely. However, this mediating role for arthritis pain intensity, in combination with the finding that nonarticular pain was not a correlate of daily SB, could suggest that treatment aimed at reducing disease activity, thereby indirectly reducing the severity of arthritis pain, is more likely to influence SB than intervening with nonarticular pain.

Participants reporting foot or ankle pain were significantly more sedentary compared to those without. Foot-related disability and pain are highly prevalent in individuals with RA (40), impacting on their ability to walk, with loss of social and leisure activities (41). To date, studies investigating foot pain and activity in patients with RA have focused primarily on self-reported limitations (41). This is the first study to explore the association between accelerometer-derived data and foot/ankle symptoms. It is worth noting that the foot/ankle pain group also had significantly higher disease activity scores and higher pain levels. Foot pain may be associated with a more severe clinical presentation and worse physical functioning (42), and these factors could also potentially account for increased SB. In our adjusted model, the presence of foot and/or ankle pain did not independently predict sedentary time after the addition of covariates; thus, further investigation is warranted in an appropriately powered study.

This study found higher levels of depression and anxiety to be associated with more SB. While no comparable evidence exists in RA, the positive association between SB and psychological factors such as depression and anxiety have been reported in other population groups (43,44). Longitudinal research is needed to determine the direction of these relationships. Interestingly, fatigue was unrelated to any aspect of SB measured in this study. This is unexpected given that physical activity interventions are known to reduce symptoms of RA-related fatigue (45), and in bivariate analysis, physical inactivity has been shown to be significantly associated with fatigue (46). Patients with RA report that fatigue also limits their daily activities (13). Other research suggests that particular dimensions of fatigue in RA are more closely related to physical activity and inactivity, namely fatigue-related reduced activity and physical fatigue (12). Our study focused on fatigue severity, and we may have found different results had we examined multiple dimensions of fatigue.

Discrepancies between subjective estimates of sedentary time and accelerometer-derived data have been previously observed in patients with RA, with a tendency to underreport sedentary time and overreport physical activity on questionnaires (6). In the current study, participants' subjective estimate of their sedentary time correlated weakly with the objective measurement. Self-reporting of sedentary time has known susceptibility to recall bias and measurement error (47). Use of objective measurement avoids these subjective biases, but there can be issues around patient compliance. In this study, the activPAL4 monitor was well tolerated, with 90% of participants returning 7 days of valid data.

The pattern of accumulation of SB time is also important, with prolonged bouts of uninterrupted sedentary time conferring the greatest cardiometabolic risk (48). The current study assessed several aspects of sedentary pattern, including the number of longer sedentary bouts and sedentary bout interruptions, but found no association with patient-reported outcomes. While these long bouts were considered more harmful in observational studies, replacing with shorter sedentary bouts was not helpful in lowering mortality risk. Instead, substituting with physical activity of any intensity was necessary to mitigate the mortality risk (49).

Due to the cross-sectional nature and the mainly weak relationships, it is not possible to make clinical recommendations based on this study alone. Nonetheless, our results suggest that disease-related factors such as low mood and multiple painful joints warrant consideration. Findings also suggest that we must look beyond RA symptoms in seeking to identify and understand the factors contributing to SB in this at-risk population. Research on the implications of reducing SB is limited; however, Thomsen and colleagues (8) have demonstrated the positive impact that behavior change interventions can have on SB in patients with RA. Future research should seek to meet the research agenda set out by EULAR (35) and prospectively evaluate the relationship between SB and disease-specific outcomes.

This is one of the first studies to examine the relationships between SB, pain characteristics, and other patient-reported symptoms, such as depression, anxiety, sleep, and fatigue, using a reliable and valid objective measure. Other strengths of this study include excellent compliance rates, with monitor wear for 7 days. This study also has some limitations. The cross-sectional design precludes us from establishing the direction of relationships. SB could represent both a consequence and a cause of increased pain in RA. While our regression analysis investigated pain as a predictor of SB, there is evidence from animal research and other pain populations that increased pain may also be a consequence of SB (50). The study was not adequately powered for multivariable analyses; however, given that pain characteristics were not independent predictors in partially adjusted models between pain and SB, inclusion of a full range of confounders was unlikely to change these results.

This study assessed a range of pain characteristics and patient-reported outcome measures and examined their relationship with SB. The various dimensions of pain in RA and their role in SB have not previously been examined. We found that several pain characteristics correlated with daily sedentary time, but these relationships were relatively small and not maintained when other factors were accounted for in multivariable models. Other clinical correlates included mood and disease activity, while fatigue, sleep, or nonarticular pain were not related to SB. These results indicate that other non-RA–related factors beyond pain and clinical characteristics are likely to be important in determining SB. This may have implications for developing and delivering future interventions.

ACKNOWLEDGMENTS

The authors thank Dr. James Green (University of Limerick) for his assistance with data analysis and interpretation. The authors also thank the participants, who so generously gave their time, and the outpatient staff at Cork University Hospital, who facilitated this project.

AUTHOR CONTRIBUTIONS

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

Study conception and design. O'Leary, Larkin, Murphy, Quinn.
Acquisition of data. O'Leary, Murphy, Quinn.
Analysis and interpretation of data. O'Leary, Larkin, Quinn.

REFERENCES

1. Powell C, Herring MP, Dowd KP, Donnelly AE, Carson BP. The cross-sectional associations between objectively measured sedentary time and cardiometabolic health markers in adults: a systematic review with meta-analysis component. Obes Rev 2018;19:381–95.
2. Sedentary Behaviour Research Network. Letter to the editor: standardized use of the terms "sedentary" and "sedentary behaviours". Appl Physiol Nutr Metab 2012;37:540–2.
3. Fenton SA, Veldhuijzen van Zanten JJ, Kitas GD, Duda JL, Rouse PC, Yu CA, et al. Sedentary behaviour is associated with increased long-term cardiovascular risk in patients with rheumatoid arthritis independently of moderate-to-vigorous physical activity. BMC Musculoskelet Disord 2017;18:1–12.
4. Carmona L, Cross M, Williams B, Lassere M, March L. Rheumatoid arthritis. Best Pract Res Clin Rheumatol 2010;24:733–45.
5. Metsios GS, Kitas GD. Physical activity, exercise and rheumatoid arthritis: effectiveness, mechanisms and implementation. Best Pract Res Clin Rheumatol 2018;32:669–82.
6. Yu CA, Rouse PC, Veldhuijzen Van Zanten JJ, Ntoumanis N, Kitas GD, Duda JL, et al. Subjective and objective levels of physical activity and their association with cardiorespiratory fitness in rheumatoid arthritis patients. Arthritis Res Ther 2015;17:59.
7. Van Dyck D, Cerin E, De Bourdeaudhuij I, Hinckson E, Reis RS, Davey R, et al. International study of objectively measured physical activity and sedentary time with body mass index and obesity: IPEN adult study. Int J Obes (Lond) 2015;39:199–207.
8. Thomsen T, Aadahl M, Beyer N, Hetland ML, Løppenthin K, Midtgaard J, et al. The efficacy of motivational counselling and SMS reminders on daily sitting time in patients with rheumatoid arthritis: a randomised controlled trial. Ann Rheum Dis 2017;76:1603–6.
9. Sokka T, Häkkinen A, Kautiainen H, Maillefert JF, Toloza S, Hansen TM, et al. Physical inactivity in patients with rheumatoid arthritis: data from twenty-one countries in a cross-sectional, international study. Arthritis Rheum 2008;59:42–50.
10. Munsterman T, Takken T, Wittink H. Are persons with rheumatoid arthritis deconditioned? A review of physical activity and aerobic capacity. BMC Musculoskelet Disord 2012;13:202.
11. Hewlett S, Choy E, Kirwan J. Furthering our understanding of fatigue in rheumatoid arthritis. J Rheumatol 2012;39:1775–7.
12. Loppenthin K, Esbensen BA, Ostergaard M, Jennum P, Tolver A, Aadahl M, et al. Physical activity and the association with fatigue and sleep in Danish patients with rheumatoid arthritis. Rheumatol Int 2015;35:1655–64.
13. Thomsen T, Beyer N, Aadahl M, Hetland ML, Løppenthin K, Midtgaard J, et al. Sedentary behaviour in patients with rheumatoid arthritis: a qualitative study. Int J Qual Stud Health Well-being 2015;10:28578.
14. Hurkmans E, van der Giesen FJ, Vliet Vlieland TP, Schoones J, van den Ende EC. Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis [review]. Cochrane Database Syst Rev 2009;2009:CD006853.
15. Greene BL, Haldeman GF, Kaminski A, Neal K, Lim SS, Conn DL. Factors affecting physical activity behavior in urban adults with arthritis who are predominantly African-American and female. Phys Ther 2006;86:510–9.
16. Edwardson CL, Winkler EA, Bodicoat DH, Yates T, Davies MJ, Dunstan DW, et al. Considerations when using the activPAL monitor in field-based research with adult populations. J Sport Heal Sci 2017;6:162–78.
17. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
18. Ferraz MB, Quaresma MR, Aquino LR, Atra E, Tugwell P, Goldsmith CH. Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. J Rheumatol 1990;17:1022–4.
19. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10.
20. Stucki G, Liang MH, Stucki S, Brühlmann P, Michel BA. A self-administered Rheumatoid Arthritis Disease Activity Index (RADAI) for epidemiologic research: psychometric properties and correlation with parameters of disease activity. Arthritis Rheum 1995;38:795–8.
21. Kirwan JR, Reeback JS. Stanford Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. Br J Rheumatol 1986;25:206–9.
22. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–70.
23. Wolfe F. Fatigue assessments in rheumatoid arthritis: comparative performance of visual analog scales and longer fatigue questionnaires in 7760 patients. J Rheumatol 2004;31:1896–902.
24. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213.
25. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther 2005;7:R796–806.
26. Anderson J, Caplan L, Yazdany J, Robbins ML, Neogi T, Michaud K, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. Arthritis Care Res (Hoboken) 2012;64:640–7.
27. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83.
28. Larkin L, Nordgren B, Purtill H, Brand C, Fraser A, Kennedy N. Criterion validity of the activPAL™ activity monitor for sedentary and physical activity patterns in people who have rheumatoid arthritis. Phys Ther 2016;96:1093–101.
29. Fenton SA, Veldhuijzen van Zanten JJ, Duda JL, Metsios GS, Kitas GD. Sedentary behaviour in rheumatoid arthritis: definition, measurement and implications for health. Rheumatology (Oxford) 2018;57:213–26.
30. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. Med Sci Sports Exerc 2011;43:1561–7.
31. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011;8:115.

32. Khoja SS, Almeida GJ, Chester Wasko M, Terhorst L, Piva SR. Association of light-intensity physical activity with lower cardiovascular disease risk burden in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:424–31.
33. Chastin SF, Buck C, Freiberger E, Murphy M, Brug J, Cardon G, et al. Systematic literature review of determinants of sedentary behaviour in older adults: a DEDIPAC study. Int J Behav Nutr Phys Act 2015;12:127.
34. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986;51:1173–82.
35. Rausch Osthoff AK, Niedermann K, Braun J, Adams J, Brodin N, Dagfinrud H, et al. 2018 EULAR recommendations for physical activity in people with inflammatory arthritis and osteoarthritis. Ann Rheum Dis 2018;77:1251–60.
36. Huffman KM, Pieper CF, Hall KS, St Clair EW, Kraus WE. Self-efficacy for exercise, more than disease-related factors, is associated with objectively assessed exercise time and sedentary behaviour in rheumatoid arthritis. Scand J Rheumatol 2015;44:106–10.
37. Christensen AW, Rifbjerg-Madsen S, Christensen R, Dreyer L, Tillingsoe H, Seven S, et al. Non-nociceptive pain in rheumatoid arthritis is frequent and affects disease activity estimation: cross-sectional data from the FRAME study. Scand J Rheumatol 2016;45:461–9.
38. Lee YC, Frits ML, Iannaccone CK, Weinblatt ME, Shadick NA, Williams DA, et al. Subgrouping of patients with rheumatoid arthritis based on pain, fatigue, inflammation, and psychosocial factors. Arthritis Rheumatol 2014;66:2006–14.
39. Fransen J, Langenegger T, Michel BA, Stucki G. Feasibility and validity of the RADAI, a self-administered rheumatoid arthritis disease activity index. Rheumatology (Oxford) 2000;39:321–7.
40. Borman P, Ayhan F, Tuncay F, Sahin M. Foot problems in a group of patients with rheumatoid arthritis: an unmet need for foot care. Open Rheumatol J 2012;6:290–5.
41. Otter SJ, Lucas K, Springett K, Moore A, Davies K, Young A, et al. Identifying patient-reported outcomes in rheumatoid arthritis: the impact of foot symptoms on self-perceived quality of life. Musculoskeletal Care 2012;10:65–75.
42. Wechalekar MD, Lester S, Hill CL, Lee A, Rischmueller M, Smith MD, et al. Active foot synovitis in patients with rheumatoid arthritis: unstable remission status, radiographic progression, and worse functional outcomes in patients with foot synovitis in apparent remission. Arthritis Care Res (Hoboken) 2016;68:1616–23.
43. Schuch F, Vancampfort D, Firth J, Rosenbaum S, Ward P, Reichert T, et al. Physical activity and sedentary behavior in people with major depressive disorder: a systematic review and meta-analysis. J Affect Disord 2017;210:139–50.
44. Vancampfort D, Stubbs B, Herring MP, Hallgren M, Koyanagi A. Sedentary behavior and anxiety: association and influential factors among 42,469 community-dwelling adults in six low- and middle-income countries. Gen Hosp Psychiatry 2018;50:26–32.
45. Rongen-van Dartel SA, Repping-Wuts H, Flendrie M, Bleijenberg G, Metsios GS, van den Hout WB, et al. Effect of aerobic exercise training on fatigue in rheumatoid arthritis: a meta-analysis. Arthritis Care Res (Hoboken) 2015;67:1054–62.
46. Katz P, Margaretten M, Gregorich S, Trupin L. Physical activity to reduce fatigue in rheumatoid arthritis: a randomized controlled trial. Arthritis Care Res (Hoboken) 2018;70:1–10.
47. Atkin AJ, Gorely T, Clemes SA, Yates T, Edwardson C, Brage S, et al. Methods of measurement in epidemiology: sedentary behaviour. Int J Epidemiol 2012;41:1460–71.
48. Saunders TJ, Larouche R, Colley RC, Tremblay MS. Acute sedentary behaviour and markers of cardiometabolic risk: a systematic review of intervention studies. J Nutr Metab 2012;2012:1–12.
49. Diaz KM, Duran AT, Colabianchi N, Judd SE, Howard VJ, Hooker SP. Potential effects on mortality of replacing sedentary time with short sedentary bouts or physical activity: a national cohort study. Am J Epidemiol 2019;188:537–44.
50. Park S, Thogersen-Ntoumani C, Veldhuijzen van Zanten JJ, Ntoumanis N. The role of physical activity and sedentary behavior in predicting daily pain and fatigue in older adults: a diary study. Ann Behav Med 2018;52:19–28.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 998–1003
DOI 10.1002/acr.24487

Management of Rheumatic Diseases During the COVID-19 Pandemic: A National Veterans Affairs Survey of Rheumatologists

Jasvinder A. Singh,[1] John S. Richards,[2] Elizabeth Chang,[3] Amy Joseph,[4] and Bernard Ng[5]

Objective. To assess the experience, views, and opinions of rheumatology providers at Veterans Affairs (VA) facilities about rheumatic disease health care issues during the COVID-19 pandemic.

Methods. We performed an anonymized cross-sectional survey, conducted from April 16 to May 18, 2020, of VA rheumatology providers. We assessed provider perspectives on COVID-19 issues and resilience.

Results. Of the 153 eligible VA rheumatologists, 103 (67%) completed the survey. A significant proportion of providers reported a ≥50% increase related to COVID-19 in visits by telephone (53%), video-based VA video connect (VVC; 44%), and clinical video telehealth with a facilitator (29%). A majority of the responders were somewhat or very comfortable with technology for providing health care to established patients during the COVID-19 pandemic using telephone (87%), VVC (64%), and in-person visits (54%). A smaller proportion were comfortable with technology providing health care to new patients. At least 65% of rheumatologists considered telephone visits appropriate for established patients with gout, osteoporosis, polymyalgia rheumatica, stable rheumatoid arthritis, stable spondyloarthritis, or osteoarthritis; 32% reported a rheumatology medication shortage. Adjusted for age, sex, and ethnicity, high provider resilience was associated with significantly higher odds ratios (ORs) of comfort with technology for telephone (OR 3.1 [95% confidence interval (95% CI) 1.1–9.7]) and VVC visits for new patients (OR 4.7 [95% CI 1.4–15.7]).

Conclusion. A better understanding of COVID-19 rheumatic disease health care issues using a health-system approach can better inform providers, improve provider satisfaction, and have positive effects on the care of veterans with rheumatic disease.

INTRODUCTION

COVID-19 is highly infectious, with significant associated mortality (1). Not surprisingly, its effects on people and societies are multiple. To combat this pandemic, several measures for infection prevention have been implemented. Stay at home (shelter-in-place), social distancing, and other measures to reduce transmission have been adopted by many countries worldwide, including the US (2).

The COVID-19 pandemic has had a significant impact on health care and health care delivery systems. Three major changes have involved the conversion of regular in-person clinic visits to telephone/video health care visits, the use of personal protective equipment by both patients and health care providers during in-person health care visits, and the performance of some work duties by health care providers while working from home (3). The reduced in-person access to health care providers and health information is worrisome for people with rheumatic diseases, who require close long-term monitoring. Provision of optimal health care in these suboptimal circumstances is very challenging.

The Veterans Affairs (VA) is the largest integrated health care system in the US, with 1,255 health facilities that provide care to

Supported by the University of Alabama at Birmingham and the Birmingham Veterans Affairs Medical Center, Birmingham, Alabama.

[1]Jasvinder A. Singh, MBBS, MPH: Birmingham Veterans Affairs Medical Center and University of Alabama at Birmingham; [2]John S. Richards, MBBS: Veterans Affairs Pittsburgh Healthcare System and University of Pittsburgh, Pittsburgh, Pennsylvania; [3]Elizabeth Chang, MD, MSCI: Phoenix Veterans Affairs Medical Center, Phoenix, Arizona; [4]Amy Joseph, MD: Veterans Affairs St. Louis Health Care System and Washington University School of Medicine, St. Louis, Missouri; [5]Bernard Ng, MBBS, MSc: Veterans Affairs Puget Sound Health Care System and University of Washington, Seattle.

Dr. Singh has received consulting fees from Medisys, Fidia, UBM LLC, Trio Health, Medscape, WebMD, Clinical Care Options, Clearview Healthcare Partners, Putnam Associates, Spherix, and Practice Point Communications (less than $10,000 each) and from Crealta/Horizon (more than $10,000), and owns stock options in Amarin Pharmaceuticals and Viking Therapeutics. No other disclosures relevant to this article were reported.

Address correspondence to Jasvinder A. Singh, MBBS, MPH, University of Alabama at Birmingham, Faculty Office Tower 805B, 510 20th Street S, Birmingham, AL 35294-0022. Email: Jasvinder.md@gmail.com.

Submitted for publication June 4, 2020; accepted in revised form October 8, 2020.

SIGNIFICANCE & INNOVATIONS

- A majority of the rheumatologists were somewhat or very comfortable with technology for providing health care to established patients during the COVID-19 pandemic, but not to new patients.
- Rheumatologists reported some shortages of hydroxychloroquine and of interleukin-6 inhibitors for their patients with rheumatic diseases.
- At least 65% of rheumatologists considered telephone visits appropriate for established patients with gout, osteoporosis, polymyalgia rheumatica, stable rheumatoid arthritis, stable spondyloarthritis, and osteoarthritis.
- High provider resilience was independently associated with significantly higher odds of more comfort with technology for telephone-assisted or video-assisted telemedicine visits.

>9 million veterans annually (4). The VA has had a state-of-the-art electronic health care record system since 1998 that has helped to improve quality of health care. The VA pioneered telehealth more than a decade ago (5). VA telemedicine visits, including using telephone or video (with a facilitator for examination [clinical video telehealth (CVT)] or without a facilitator [VA video connect (VVC)], direct-to-patient), were performed for 702,000 veterans in the fiscal year 2016 (6). In 2019, more than 900,000 veterans received care through VA telemedicine (5).

Most VA facilities switched from in-person outpatient visits to telemedicine, using telephone or video (CVT or VVC) visits, between March 16 and 20, 2020, with many facilities prohibiting routine in-person outpatient visits. Our study objective was to conduct a cross-sectional survey of a nationally representative sample of rheumatologists at the VA during the first few months of the COVID-19 pandemic, to assess their experience, views, and opinions about rheumatic disease health care issues, and to understand the impact of the pandemic on VA rheumatologists and their patients.

MATERIALS AND METHODS

This study was approved by the human ethics committee at the University of Alabama at Birmingham. We obtained a list of email addresses of VA rheumatologists from the VA Rheumatology Consortium (VARC). VARC is a volunteer work group of VA rheumatologists who practice across the US. These data are available from the authors after appropriate approvals have been obtained from the Ethics Committee at the University of Alabama at Birmingham and meeting all privacy policies and regulations. After prepiloting with 6 rheumatologists, we finalized the survey. We used Qualtrics survey software to send an anonymous survey to all VA rheumatologists who were VARC members on April 16, 2020. Nonresponders received reminders to complete the survey from April 21 to May 18, 2020.

The survey assessed providers' views and opinions about the new health care delivery methods, including the best health care delivery modality (in-person, telephone, or video visit) for the management of each rheumatic disease, diseases appropriate for alternative methods, the perceived risk of COVID-19 in rheumatic diseases, rheumatic disease medication shortages, and the safety of a future COVID-19 vaccine with rheumatic disease medications. Only a few, but not all questions included CVT with a facilitator, since CVT is used much less frequently compared to a telephone visit or VVC. Provider resilience, or stress coping ability, was measured with a validated 2-item Connor-Davidson Resilience Scale (7), scored from 0 to 8, higher scores corresponding with higher resilience, with a general population mean of 6.9. Physicians have higher resilience scores compared to the general employed population (8).

Summary statistics were assessed as proportions. Since the number of people completing the surveys was close to 100, the actual numbers in the tables were close to the percentages, which are presented in the Results section. Logistic regression assessed whether provider age, sex, years of experience, and provider resilience (categorized as high resilience, score of 7 or 8, i.e., scores at par with the general population or higher) were independently associated with comfort with technology in providing virtual care to new or established clinic patients. We obtained the information on sex and age for all potential participants from Healthgrades and other publicly available search websites. Analyses were done using IBM SPSS, version 25. The University of Alabama at Birmingham's Institutional Review Board (IRB) approved this study, and all investigations were conducted in conformity with ethical principles of research (UAB X120207004). The IRB waived the need for an informed consent for this anonymized study.

RESULTS

Of the 153 eligible VA rheumatologists, 103 completed the survey (67% response rate). Of these, 26% each were in the age groups 45–54 years and 55–64 years; 56% were White, 27% Asian, 6% African American, and 5% Hispanic; and 63% were female. More than two-thirds had practiced rheumatology for 10 years or more (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24487/abstract). Nonresponders were slightly older (16% versus 11% were age ≥65 years) and more likely to be male compared to the survey responders (45% versus 38%).

Rheumatic diseases and adjudicated appropriateness of health care delivery methods early in the COVID-19 pandemic. Two-thirds or more of the rheumatologists chose a telephone follow-up visit as the best modality for gout, osteoporosis, polymyalgia rheumatica, stable rheumatoid arthritis, stable spondyloarthritis, and osteoarthritis (Figure 1). One-third or more chose a video-based VVC follow-up visit as

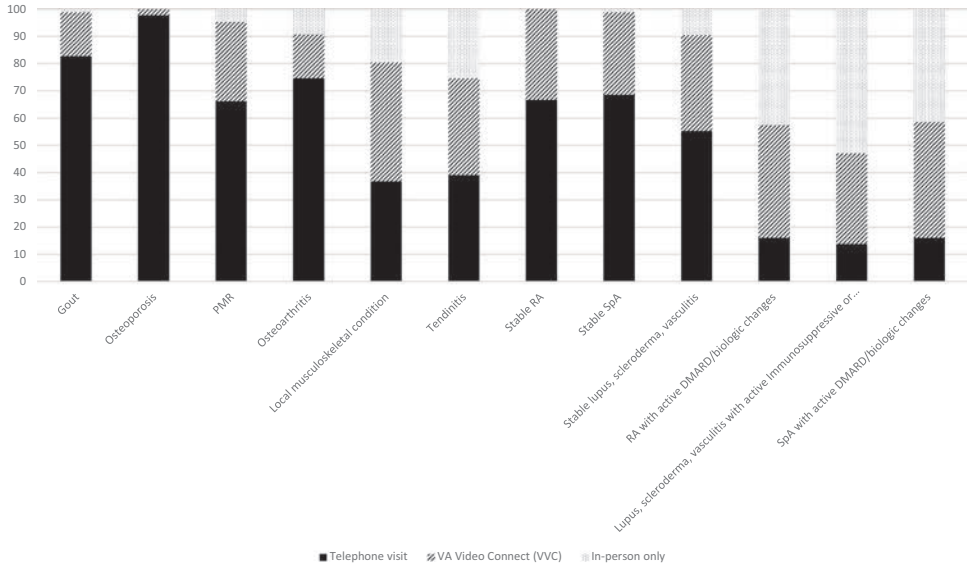

Figure 1. Provider-preferred clinic follow-up appointment modality for established patients due to COVID-19 by the type of rheumatic disease. The y-axis represents the percent of all valid nonmissing responses. The number of missing responses for each condition varied (n =16 to 18). Providers responded to the question: "Which of the following conditions in established patients do you feel are best suited for telephone or video-based visits during follow-up during the COVID-19 pandemic? Choose the single best response." This was followed by listing each rheumatic condition in a separate row. Response options included telephone, Veterans Administration (VA) video connect, and in-person visit. DMARD = disease-modifying antirheumatic drug; PMR = polymyalgia rheumatica; RA = rheumatoid arthritis; SpA = spondyloarthritis.

the best modality for local musculoskeletal conditions, tendinitis, rheumatoid arthritis with active medication (disease-modifying antirheumatic drug [DMARD]/biologic) changes, and patients with stable lupus, scleroderma, or vasculitis (Figure 1). In contrast, 41–53% of responders selected an in-person follow-up visit as the best modality for people with lupus, scleroderma, or vasculitis with immunosuppressive or glucocorticoid dose changes and rheumatoid arthritis or spondyloarthritis with active medication (DMARD/biologic) changes (Figure 1). A total of 43% of responders agreed or strongly agreed that they were able to provide health care efficiently, 68% were able to provide it safely, and >50% spent a lot of extra time providing this care.

Provider technology use and comfort for VA health care delivery methods early in the COVID-19 pandemic. Of the responders, 50% reported using their personal desktop and laptop, 69% were using a VA desktop, and 18% were using a VA laptop for providing VA health care during the COVID-19 pandemic (providers could choose multiple responses). Of these, 31% were working entirely from the VA hospital or VA clinic, 14% from a non-VA location (or home), and the rest were working from both non-VA and VA locations. Survey responders reported providing VA health care to veterans with rheumatic disease using multiple methods during the COVID-19 pandemic: 91% used telephone visits, 59% used video-based VVC visits, 7% used CVT visits with a facilitator, and 59% used in-person visits. A significant proportion of providers reported a 50% or more increase in the following types of visits related to COVID-19: telephone visits (53%), video-based VVC visits (44%), and CVT visits with a facilitator (29%) (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24487/abstract).

The proportion of responders who were somewhat or very comfortable while providing health care to established clinic patients using each of these methods was as follows: telephone visits (87%), video-based VVC visits (64%), and in-person visits (54%) (Figure 2). The proportion of responders who were somewhat or very comfortable providing health care to new clinic patients was as follows: telephone visits (25%), video-based VVC visits (34%), and in-person visits (58%) (Figure 2). More than two-thirds of responders reported that evaluating a new patient scheduled in their clinic was feasible during the COVID-19 pandemic.

Risk of COVID-19 infection in veterans with rheumatic diseases. Among respondents, a majority agreed or strongly agreed that veterans with autoimmune rheumatic diseases were at a higher risk of COVID-19 infection even in the absence of immunosuppressive drugs (54%) and when currently using immunosuppressive drugs (71%). Similarly, only a small proportion (23%) agreed or strongly agreed that veterans with nonautoimmune rheumatic diseases were at a higher risk of COVID-19 infection.

Rheumatic disease medications: shortages and risks with a future COVID-19 vaccine or convalescent sera. Approximately 32% of responders reported a medication shortage. Responders indicated some (little or extreme) shortage for the following: hydroxychloroquine (45%), interleukin (IL)-6

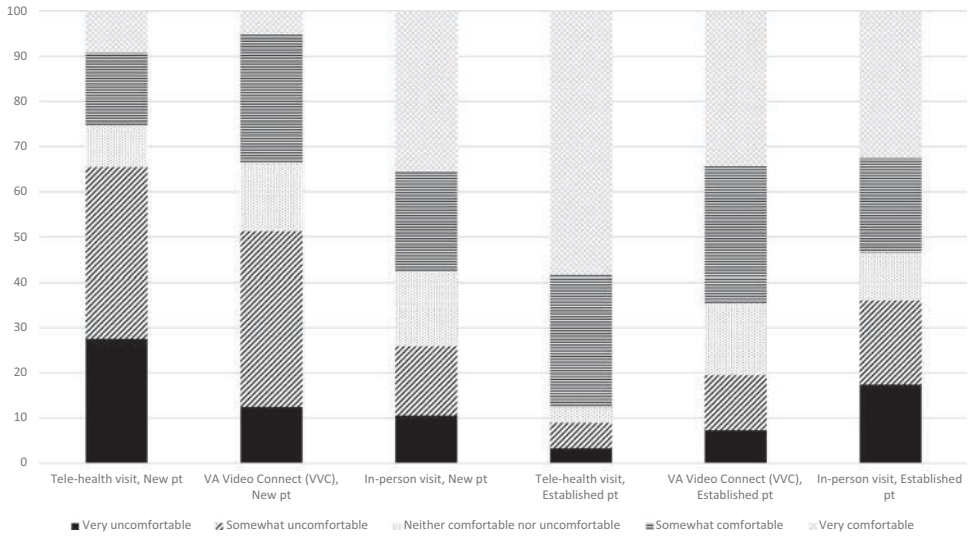

Figure 2. Provider comfort with technology in providing care to new or established patients using each of the modalities during the COVID-19 pandemic. The y-axis represents the percent of all valid nonmissing responses. Providers responded to 2 questions: "What is your level of comfort with technology with providing health care to new patients in your clinic during the COVID-19 pandemic? What is your level of comfort with technology with providing health care to established patients in your clinic during the COVID-19 pandemic?" Each question was followed by listing telephone, Veterans Administration (VA) video connect, and in-person visit in a separate row. The response option was a 5-point ordinal scale: very uncomfortable, somewhat uncomfortable, neither comfortable nor uncomfortable, somewhat comfortable, and very comfortable. As an example, 87% of respondents were somewhat or very comfortable while providing health care to established clinic patients with telephone visits versus only 25% of respondents for new patient evaluations. pt = patient.

inhibitors (15%), non–tumor necrosis factor (TNF) biologics (1%), TNF-biologics (0%), Janus-inhibitors (1%), and other immunosuppressives (1%).

Most responders would not withhold hydroxychloroquine (95%) or sulfasalazine (74%) for a future, live attenuated COVID-19 vaccine. A majority would withhold methotrexate or leflunomide (66%) and glucocorticoids of 20 mg/day or higher (52%) for 2 weeks or less, and would withhold anti TNF-biologics (85%), anti-IL-17/23 biologics (82%), Janus-kinase inhibitors (78%), belimumab (77%), non-TNF biologics (76%), and immunosuppressive drugs such as azathioprine (64%), for 3–8 weeks for administering a future, live attenuated COVID-19 vaccine. A majority of responders (55–100%) would not withhold these drugs for administering a killed COVID-19 vaccine; another 5–30% would hold them off for <2 weeks. A majority of responders (≥50%) would not withhold any of these drugs for a convalescent sera treatment of COVID-19.

Perceived increase in health care disparities in veterans during COVID-19. A significant proportion of responders perceived an increase in health care disparities during the COVID-19 pandemic in the following groups: African Americans (40%), Hispanics (31%), other racial minorities (19%), low socioeconomic groups (47%), females (8%), rural residents (23%), and those with nonservice-connected illnesses (12%). Of responders, 24% had had a family member or friend with a COVID-19–positive test. Three responders had been tested for COVID-19, 1 reported a negative test result, and 2 received care for COVID-19 at home.

Responder resilience and comfort with technology for virtual health care visits. Resilience was high among responders. The mean ± SD Connor-Davidson Resilience Scale score was 6.35 ± 1.260; scores were 6 or higher for >80% of people: 46%, 16%, and 23% of the responders had high resilience scores of 6, 7, and 8, respectively.

Adjusted for age, sex, and ethnicity, a high provider resilience score was independently associated with a significantly higher odds ratio (OR) of more comfort with technology (somewhat or very comfortable) for telephone health care visits (OR 3.1 [95% confidence interval (95% CI) 1.1–9.7]) and video-based VVC visits (OR 4.7 [95% CI 1.4–15.7]) for new patients, with no difference for in-person visits (OR 1.8 [95% CI 0.7–5.0]). No significant associations of provider resilience were noted with comfort with technology for established patients for telephone (OR 1.7 [95% CI 0.3–8.0]), VVC (OR 1.7 [95% CI 0.6–5.0]), or in-person visits (OR 2.8 [95% CI 1.0–7.8]).

DISCUSSION

We performed a national cross-sectional study of rheumatologists at the VA, the largest integrated health care system in the US. The survey response rate was 67%, higher than the average 61% response rate for physician surveys (9). Survey responders were similar in age and sex distribution to all potential participants, with slight differences. The survey was conducted 1 month after COVID-19–associated outpatient health care delivery changes at

the VA. Therefore, findings mostly represent provider experience and practice patterns early in the COVID-19 pandemic. Several study findings deserve further discussion.

During the COVID-19 pandemic, telemedicine has emerged as one of the main ways to deliver health care. Telemedicine is an acceptable alternative to an in-person visit from the patient perspective. It can ameliorate the economic burden of clinic visits for people traveling long distances, and patients are satisfied with telemedicine visits in these situations (10). In an observational study of 85 patients with inflammatory arthritis at a single VA Medical Center, patient-reported outcomes for care delivered via telemedicine were similar to usual care, with a significant cost and distance savings (11).

In a meta-analysis of telemedicine studies in rheumatology, feasibility and patient satisfaction rates were high or very high for various telemedicine interventions, and effectiveness was similar to a standard in-person approach (12). On the other hand, the majority of people preferred an in-person over telemedicine visit for pediatric rheumatology care, despite travel and inconvenience (13). These articles highlight the contrast in patient preference for telemedicine versus in-person visits. High rates of patient satisfaction with telemedicine care (when offered and provided to selected patients) and higher patient preference for in-person over telemedicine visits can coexist in an ideal world. Telemedicine is a viable alternative to in-person rheumatology follow-up visits during the COVID-19 pandemic.

Our study is the first national study of VA rheumatologists to examine which rheumatic conditions were considered appropriate for virtual visits using the telephone during the COVID-19 pandemic. Survey responders made a clear distinction between conditions that were appropriate for telemedicine versus in-person visits. More than 90% of rheumatologists surveyed considered gout, osteoporosis, osteoarthritis, and polymyalgia rheumatica to be appropriate for telephone visits or video-based health care visits for established patients. Active systemic autoimmune rheumatic conditions (rheumatoid arthritis, spondyloarthritis, lupus, vasculitis, scleroderma, etc.) with ongoing changes to disease-modifying, immunosuppressive, or biologic medications were considered most appropriate for in-person visits during the COVID-19 pandemic by the majority of responders. However, one-third favored video visits. These patterns may change over time.

Our study found that most respondents were comfortable with telemedicine technology to provide health care to established patients with rheumatic diseases. In contrast, less than one-third of responders were comfortable with telemedicine technology in providing care to new patients. Previous studies have shown that physicians are satisfied with telemedicine when providing care in specific specialties, including cardiology (14), neurology (15), and primary care (16). Our national study is among the first to assess this comfort for various rheumatic diseases. Our study describes VA rheumatology providers' views and preferences 4–8 weeks after the switch from in-person regular outpatient visits to telemedicine at VA facilities due to the COVID-19 outbreak in the US.

We found that a high provider resilience score was associated with a 3- to 5-fold higher odds of comfort with technology for telephone and video visits for new patients, with no difference for in-person visits. To our knowledge, there are no published studies of the relationship between provider resilience and higher comfort levels with using telemedicine. Our study provides new data that will need confirmation in other studies. A mean resilience score of 6.35 for VA rheumatologists was similar to 6.49 for US physicians from a recent survey (8).

We found that 55–100% of VA rheumatology providers would not withhold 1 or more treatments for rheumatic diseases or withhold it for <2 weeks to administer an inactivated/killed COVID-19 vaccine. In contrast, >75% would withhold biologic therapy for 2–8 weeks for administering a live attenuated COVID-19 vaccine. The VA rheumatology providers reported some shortages of hydroxychloroquine (45%) and IL-6 inhibitors (16%) for their VA patients with rheumatic diseases. Due to the potential for hydroxychloroquine and IL-6 inhibitors to be treatments for COVID-19, shortages have been reported by patients with rheumatic diseases (17). Poor outcomes in African Americans with COVID-19 point to racial health care disparities in the US (18). VA rheumatologists, however, perceived a potential increase in health care disparities not only in African Americans and Hispanics, but also in people in the low socioeconomic groups and those living in the rural areas.

Our study findings must be interpreted considering limitations. These findings cannot be generalized to non-VA settings without an additional similar study. Even though our study responders were similar to the overall sample in age and sex, we do not have information on other characteristics (since the survey was anonymous), and therefore nonresponse bias is a limitation. Several outcomes represent VA rheumatology provider views and opinions, which might change as the COVID-19 epidemic evolves. However, given the nature of health care delivery changes related to COVID-19, examining provider views and opinions was our study goal. Prior experience with telemedicine was not assessed, which might have influenced comfort with technology and the likelihood of using telemedicine visits. Providers could only choose 1 best modality for follow-up visits; for some conditions, 2 modalities could perhaps be equally good, which our survey is unable to detect.

In conclusion, we conducted a study of experiences, views, and opinions of VA rheumatology providers. The VA is the largest integrated health care system in the US; therefore these findings are important and have implications for the VA system. The knowledge of barriers to the use of telemedicine, medication shortage, increasing health care disparities, and considerations for future COVID-19 vaccines can inform future delivery of health care to patients with rheumatic diseases.

ACKNOWLEDGMENTS

The authors thank all of the participating Veterans Affairs Rheumatology Consortium members who responded to the survey. We also thank the members of the Veterans Affairs Field Advisory Committee for their support.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Singh, Richards, Chang, Joseph, Ng.
Acquisition of data. Singh.
Analysis and interpretation of data. Singh, Richards, Chang, Joseph, Ng.

REFERENCES

1. Weiss P, Murdoch DR. Clinical course and mortality risk of severe COVID-19. Lancet 2020;395:1014–5.
2. Gostin LO, Wiley LF. Governmental public health powers during the COVID-19 pandemic: stay-at-home orders, business closures, and travel restrictions. JAMA 2020;323:2137–8.
3. Hollander JE, Carr BG. Virtually perfect? Telemedicine for Covid-19. N Engl J Med 2020;383:1679–81.
4. Perlin JB, Kolodner RM, Roswell RH. The Veterans Health Administration: quality, value, accountability, and information as transforming strategies for patient-centered care. Am J Manag Care 2004;10 Pt 2:828–36.
5. Department of Veterans Affairs. VA telehealth. 2020. URL: https://telehealth.va.gov/.
6. Department of Veterans Affairs. VA telehealth services: fact sheet. 2020. URL: https://www.va.gov/COMMUNITYCARE/docs/news/VA_Telehealth_Services.pdf.
7. Vaishnavi S, Connor K, Davidson JR. An abbreviated version of the Connor-Davidson Resilience Scale (CD-RISC), the CD-RISC2: psychometric properties and applications in psychopharmacological trials. Psychiatry Res 2007;152:293–7.
8. West CP, Dyrbye LN, Sinsky C, Trockel M, Tutty M, Nedelec L, et al. Resilience and burnout among physicians and the general US working population. JAMA Netw Open 2020;3:e209385.
9. Cummings SM, Savitz LA, Konrad TR. Reported response rates to mailed physician questionnaires. Health Serv Res 2001;35:1347–55.
10. Poulsen KA, Millen CM, Lakshman UI, Buttner PG, Roberts LJ. Satisfaction with rural rheumatology telemedicine service. Int J Rheum Dis 2015;18:304–14.
11. Wood PR, Caplan L. Outcomes, satisfaction, and costs of a rheumatology telemedicine program: a longitudinal evaluation. J Clin Rheumatol 2019;25:41–4.
12. Piga M, Cangemi I, Mathieu A, Cauli A. Telemedicine for patients with rheumatic diseases: systematic review and proposal for research agenda. Semin Arthritis Rheum 2017;47:121–8.
13. Bullock DR, Vehe RK, Zhang L, Correll CK. Telemedicine and other care models in pediatric rheumatology: an exploratory study of parents' perceptions of barriers to care and care preferences. Pediatr Rheumatol Online J 2017;15:55.
14. Oliveira GL, Cardoso CS, Ribeiro AL, Caiaffa WT. Physician satisfaction with care to cardiovascular diseases in the municipalities of Minas Gerais: Cardiosatis-TEAM Scale. Rev Bras Epidemiol 2011;14:240–52.
15. Harper K, McLeod M, Brown SK, Wilson G, Turchan M, Gittings EM, et al. Teleneurology service provided via tablet technology: 3-year outcomes and physician satisfaction. Rural Remote Health 2019;19:4743.
16. Waterman AD, Banet G, Milligan PE, Frazier A, Verzino E, Walton B, et al. Patient and physician satisfaction with a telephone-based anticoagulation service. J Gen Intern Med 2001;16:460–3.
17. Peschken CA. Possible consequences of a shortage of hydroxychloroquine for patients with systemic lupus erythematosus amid the COVID-19 pandemic. J Rheumatol 2020;47:787–90.
18. Yancy CW. COVID-19 and African Americans. JAMA 2020;323:1891–2.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 1004–1012
DOI 10.1002/acr.24547

BRIEF REPORT

Pregnancy and Rheumatic Disease: Experience at a Single Center in New York City During the COVID-19 Pandemic

Medha Barbhaiya,[1] Bessie Stamm,[2] Gregory Vitone,[2] Marianna B. Frey,[2] Deanna Jannat-Khah,[1] Jonah Levine,[2] JoAnn Vega,[2] Candace H. Feldman,[3] Jane E. Salmon,[1] Mary K. Crow,[1] Vivian Bykerk,[1] Michael D. Lockshin,[1] Lisa Sammaritano,[1] and Lisa A. Mandl[1]

Objective. The present study was undertaken to evaluate the pregnancy experiences of women receiving care in the division of rheumatology at a major academic center in New York City during the COVID-19 pandemic.

Methods. A web-based COVID-19 survey was emailed to 26,045 patients who were followed in the division of rheumatology at a single center in New York City. Women ages 18–50 years were asked about their pregnancy. We compared the COVID-19 experience between pregnant and nonpregnant women and also explored the impact of the pandemic on prenatal care and perinatal outcomes.

Results. Among 7,094 of the 26,045 respondents, 1,547 were women ages 18–50 years, with 61 (4%) reporting being pregnant during the pandemic. The prevalence of self-reported COVID-19 was similar in pregnant and nonpregnant women (8% versus 9%, respectively; $P = 0.76$). Among women with COVID-19, pregnant women had a shorter duration of symptoms ($P < 0.01$) and were more likely to experience loss of smell or taste ($P = 0.02$) than nonpregnant women. Approximately three-fourths of women had a systemic rheumatic disease, with no differences when stratified by pregnancy or COVID-19 status. In all, 67% of pregnant women noted changes to prenatal care during the pandemic, and 23% of postpartum women stated that the pandemic affected delivery.

Conclusion. Among women followed in the division of rheumatology at a major center in New York City, pregnancy was not associated with increased self-reported COVID-19. Pregnancy was associated with a shorter duration of COVID-19 symptoms and a higher prevalence of loss of smell or taste. The COVID-19 pandemic impacted prenatal care for the majority of pregnant patients.

INTRODUCTION

COVID-19, caused by the SARS–CoV-2 infection, is a worldwide public health crisis. Within the US, New York City was an early hot spot from March through May 2020. Pregnant women experience immunologic and physiologic changes shown to increase the risk for more severe illness from infections (1,2). In addition, patients with rheumatic diseases may be at increased risk of severe illness due to immune dysfunction and use of immunomodulatory or immunosuppressive medications (1,2). It is not known if pregnant patients with rheumatic disease have greater risks associated with SARS–CoV-2.

In previous coronavirus outbreaks (SARS-CoV and Middle East respiratory syndrome coronavirus) and the H1N1 influenza outbreak, pregnant women were at increased risk for endotracheal intubation, admission to an intensive care unit (ICU), renal failure, and death (3,4). In contrast, the first study describing clinical characteristics in 9 pregnant women with laboratory-confirmed COVID-19 suggested that the severity of COVID-19 was similar to that in nonpregnant adults (5). To date, studies of the effect of

Supported by the Barbara Volcker Center for Women and Rheumatic Diseases at Hospital for Special Surgery. Dr. Barbhaiya's work was supported by the Rheumatology Research Foundation (Investigator award).

[1]Medha Barbhaiya, MD, MPH, Deanna Jannat-Khah, DrPH, MSPH, Jane E. Salmon, MD, Mary K. Crow, MD, Vivian Bykerk, MD, Michael D. Lockshin, MD, Lisa Sammaritano, MD, Lisa A. Mandl, MD, MPH: Hospital for Special Surgery and Weill Cornell Medicine, New York, New York; [2]Bessie Stamm, BA, Gregory Vitone, BS, Marianna B. Frey, BA, Jonah Levine, BA, JoAnn Vega, BA: Hospital for Special Surgery, New York, New York; [3]Candace H. Feldman, MD, ScD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Dr. Jannat-Khah owns stock or stock options in AstraZeneca, Walgreens Boots Alliance, and CytoDyn. Dr. Mandl has received grant support from Regeneron Pharmaceuticals. No other disclosures relevant to this article were reported.

Address correspondence to Medha Barbhaiya, MD, MPH, 535 East 70th Street, New York, NY 10021. Email: barbhaiyam@hss.edu.

Submitted for publication September 19, 2020; accepted in revised form December 17, 2020.

SIGNIFICANCE & INNOVATIONS

- This is the first study to evaluate the impact of COVID-19 on pregnant patients with rheumatic disease in an early COVID-19 hot spot in the US.
- During the peak of the COVID-19 pandemic in New York City, pregnant women evaluated in our division of rheumatology reported similar COVID-19 prevalence and disease severity compared to nonpregnant patients, and 67% reported changes to their prenatal care.
- Our results demonstrating shorter total COVID-19 symptom duration and more frequent loss of taste and smell in pregnant patients with COVID-19 are novel, hypothesis-generating findings that deserve further study.

COVID-19 in pregnancy focus on the general population, not on patients with rheumatic disease.

Our study evaluates the impact of COVID-19 on women of reproductive age followed at a major rheumatology center in New York City. We hypothesized that pregnant women were at increased risk of COVID-19 and had worse outcomes than nonpregnant women and that the pandemic would impact perinatal care.

PATIENTS AND METHODS

Study population. We identified patients ages ≥18 years with at least 1 visit to a rheumatologist between April 1, 2018 and April 21, 2020 at the Hospital for Special Surgery, an academic hospital in New York City. A secure, web-based survey was sent to all English-speaking patients with an email address. Between April 24, 2020 and May 17, 2020, patients received up to 3 invitations, and a subset with missing or incorrect emails was contacted by telephone. We asked women ages 18–50 years to indicate their pregnancy status on January 1, 2020 and at the time of survey completion. Women who answered the pregnancy questions within the general medical history questionnaire are included in this analysis (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24547/abstract). Data collection was locked on July 1, 2020.

Data collection. Data on sociodemographic factors and body mass index (BMI) were collected from the last physician visit on all patients. From respondents, we collected detailed information on COVID-19 exposures, symptoms, and outcomes occurring between January 1, 2020 and the date of survey completion. COVID-19 status was defined as a composite of confirmed COVID-19 (self-report of a positive nasopharyngeal polymerase chain reaction [PCR] test) and suspected COVID-19 (being told by a health care provider of a COVID-19 diagnosis). The latter was included because PCR testing was not readily available early in the pandemic. Self-reported demographic data, rheumatic disease diagnoses, and use of immunomodulatory and/or immunosuppressive medications in the previous 6 months were also collected. Information on pain interference, anxiety, depression, fatigue, and sleep disturbance was elicited from Patient-Reported Outcomes Measurement Information System 29 (PROMIS-29). Women between 18 and 50 years of age were also asked their pregnancy status, pregnancy outcomes, and perceived impact of COVID-19 on their prenatal and perinatal care.

Statistical analysis. In our primary analysis, we compared COVID-19 status, severity, and exposure history, as well as sociodemographic factors and patient-reported outcome measures between pregnant and nonpregnant women. We also evaluated pregnancy outcomes and prenatal/perinatal care during the COVID-19 pandemic. We conducted 2 secondary analyses comparing pregnant versus nonpregnant women with COVID-19, and pregnant versus nonpregnant women without COVID-19. We performed bivariate comparisons using Fisher's exact test and chi-square tests for categorical variables and *t*-tests for continuous variables. Statistical significance was determined using a *P*-value threshold of 0.05. All analyses were performed using Stata, version 14.0. This study was approved by the Hospital for Special Surgery Institutional Review Board.

RESULTS

Among the 26,045 patients who were emailed the COVID-19 survey, 7,094 (27%) responded, of whom 1,547 (22%) were women ages 18–50 years (see Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24547/abstract). In total, 34 eligible women did not complete the pregnancy questions; therefore, 1,513 women are included in this analysis. Compared to 4,297 nonrespondents ages 18–50 years, participants were slightly older (mean ± SD age 38.1 ± 8.0 years versus 37.1 ± 8.4 years), more likely to be White (77% versus 62%), and slightly less likely to reside in New York state (73% versus 76%); no significant difference in ethnicity or BMI was noted (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24547/abstract).

In total, 61 study participants (4%) reported a pregnancy. Compared to nonpregnant women, pregnant women were younger (mean ± SD 36.1 ± 4.9 years versus 38.2 ± 8.1 years) and significantly more likely to be married, have a household income >$150,000, and hold a master's, professional, or doctorate degree (Table 1). Race and ethnicity were similar, with the majority of women in both groups being White and non-Hispanic/Latino (Table 1). Systemic rheumatic disease occurred similarly in both pregnant and nonpregnant women (67.2% versus 74.0%; $P = 0.24$). Inflammatory arthritis was most common overall, and undifferentiated connective tissue disease occurred more frequently in nonpregnant

Table 1. Characteristics of female rheumatology outpatients, ages 18–50 years, from an academic hospital in New York City, stratified by pregnancy status during the COVID-19 pandemic*

Characteristic	Nonpregnant women (n = 1,452)	Pregnant women (n = 61)	*P*
Sociodemographic factors			
Age, mean ± SD years†	38.2 ± 8.1	36.1 ± 4.9	0.04‡
Race§			0.57
Asian/Indian subcontinent	121 (8.3)	8 (13.1)	
Black	98 (6.7)	4 (6.6)	
White	1,122 (77.3)	46 (75.4)	
Other¶	97 (6.7)	2 (3.3)	
Unknown#	14 (1.0)	1 (1.6)	
Ethnicity§			0.10
Hispanic or Latino	183 (12.6)	3 (4.9)	
Not Hispanic or Latino	1,203 (82.9)	53 (86.9)	
Unknown#	66 (4.5)	5 (8.2)	
Marital status			<0.01‡
Married or partnered	762 (52.5)	54 (88.5)	
Separated, divorced, single, widowed	642 (44.2)	6 (9.8)	
Unknown#	48 (3.3)	1 (1.6)	
Household income as of January 1, 2020			<0.01‡
<$75,000	285 (19.6)	6 (9.8)	
$75,000–$150,000	395 (27.2)	13 (21.3)	
>$150,000	555 (38.2)	40 (65.6)	
Unknown#	217 (14.9)	2 (3.3)	
Education level			0.047‡
High school graduate or below	38 (2.6)	2 (3.3)	
Any amount of college (including college graduate post-college courses)	727 (50.1)	21 (34.4)	
Master's, professional, or doctorate degree	653 (45.0)	37 (60.7)	
Unknown#	34 (2.3)	1 (1.6)	
Employment status as of January 1, 2020			0.13
Employed	1,110 (76.4)	52 (85.2)	
Unemployed, retired, other	304 (20.9)	8 (13.1)	
Unknown#	38 (2.6)	1 (1.6)	
General medical history (self-reported)			
Blood clotting problem	95 (6.5)	4 (6.6)	1.00
Cancer	54 (3.7)	1 (1.6)	0.40
Chronic kidney disease	39 (2.7)	1 (1.6)	0.62
Diabetes mellitus	34 (2.3)	0 (0.0)	0.23
Heart attack	5 (0.3)	0 (0.0)	0.65
Hypertension	126 (8.7)	2 (3.3)	0.14
Lung disease	331 (22.8)	7 (11.5)	0.04‡
Obesity	136 (9.4)	2 (3.3)	0.11
Rheumatic disease history (by self-report)**			
Any systemic rheumatic disease	1,074 (74.0)	41 (67.2)	0.24
Systemic inflammatory arthritis (including adult-onset Still's disease, inflammatory arthritis, JIA, PsA, AS, RA)	597 (41.1)	22 (36.1)	0.43
Vasculitis, scleroderma, myositis (including dermatomyositis, polymyositis, other inflammatory muscle diseases)	104 (7.2)	3 (4.9)	0.50
Undifferentiated connective tissue disease	128 (8.8)	1 (1.6)	0.049‡
Systemic lupus erythematosus	268 (18.5)	16 (26.2)	0.13
Antiphospholipid syndrome	52 (3.6)	5 (8.2)	0.06
Medication history in previous 6 months (self-reported)**			
Any immunomodulatory medication use††	823 (56.7)	31 (50.8)	0.37
Any antimalarial use (hydroxychloroquine, chloroquine)	484 (33.3)	23 (37.7)	0.48
Any biologic use	296 (20.4)	7 (11.5)	0.09
Abatacept	12 (0.8)	1 (1.6)	0.50
Belimumab	31 (2.1)	0 (0.0)	0.25
TNF inhibitors (infliximab, etanercept, adalimumab, golimumab, certolizumab)	171 (11.8)	4 (6.6)	0.21
IL-12/23 inhibitors (ustekinumab, guselkumab)	4 (0.3)	0 (0.0)	0.68
IL-17 inhibitors (secukinumab, ixekizumab)	33 (2.3)	2 (3.3)	0.61

(Continued)

Table 1. (*Cont'd*)

Characteristic	Nonpregnant women (n = 1,452)	Pregnant women (n = 61)	*P*
Cyclophosphamide	3 (0.2)	0 (0.0)	0.72
Rituximab	31 (2.1)	0 (0.0)	0.25
IL-6 inhibitors (tacrolimus, sirolimus)	20 (1.4)	1 (1.6)	0.86
IL-1 inhibitors (anakinra, canakinumab, rilonacept)	5 (0.3)	0 (0.0)	0.65
Any JAK inhibitor use (tofacitinib, baricitinib, upadacitinib)	27 (1.9)	0 (0.0)	0.28
Any conventional DMARDs use	299 (20.6)	8 (13.1)	0.15
Leflunomide	16 (1.1)	1 (1.6)	0.70
Methotrexate	125 (8.6)	1 (1.6)	0.054
Mycophenolate mofetil, mycophenolic acid	79 (5.4)	0 (0.0)	0.06
Azathioprine, 6-mercaptopurine	42 (2.9)	5 (8.2)	0.02‡
Sulfasalazine	50 (3.4)	1 (1.6)	0.44
Any glucocorticoid use	285 (19.6)	9 (14.8)	0.35

* Values are the number (%) unless indicated otherwise. *T*-tests and chi-square tests were used to calculate *P* values as appropriate. AS = ankylosing spondylitis; DMARDs = disease-modifying antirheumatic drugs; IL = interleukin; JIA = juvenile idiopathic arthritis; PsA = psoriatic arthritis; RA = rheumatoid arthritis; TNF = tumor necrosis factor.
† Age was calculated from date of birth to April 24, 2020, the first date that the survey was emailed.
‡ Significant (*P* threshold < 0.05).
§ Data on race and ethnicity were obtained by self-report from survey (when available) or from self-report in the electronic medical record (EMR).
¶ Includes other race, in addition to American Indian/Alaskan Native and Native Hawaiian/Pacific Islander.
# Unknown includes those with missing values or who preferred not to answer.
** Systemic rheumatic diseases and medication history were not mutually exclusive.
†† Includes any use of antimalarials, biologics, JAK inhibitors, or conventional DMARDs.

patients (Table 1 and Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24547/abstract). Immunomodulatory medication and glucocorticoid use in the previous 6 months were similar in the 2 groups, but pregnant women had higher use of azathioprine or 6-mercaptopurine (Table 1). Approximately one-third of patients in both groups reported taking antimalarials in the previous 6 months ($P = 0.48$). Two pregnant patients reported discontinuing at least 1 immunomodulatory medication due to pregnancy.

Pregnant and nonpregnant women had similar patterns of self-isolation and potential COVID-19 exposures, both at home and at work (Table 2). Although there were similar frequencies of any COVID-19 symptoms in both groups, significantly fewer pregnant women had a total duration of symptoms ≥10 days, and pregnant women more frequently reported loss of smell or taste (Table 2). Nonpregnant women more frequently reported chest pain and joint pain (Table 2). Higher chest pain frequency did not appear to be due to underlying lung disease, as the prevalence of lung disease in nonpregnant women with chest pain was similar to the prevalence in pregnant women with chest pain ($P = 0.11$). In patients without COVID-19, there was no significant difference in loss of smell or taste between nonpregnant and pregnant women ($P = 0.22$). Similarly, no significant difference in chest pain ($P = 0.36$) or total duration of COVID-19 symptoms ≥10 days ($P = 0.17$) was demonstrated in nonpregnant versus pregnant patients without COVID-19. The only clinically significant difference in PROMIS-29 scores was less depression in pregnant women versus nonpregnant women (Table 2).

COVID-19 infection was reported by 5 pregnant and 136 nonpregnant women (8.2% versus 9.4%; $P = 0.76$). Pregnant women with COVID-19 tended to be younger than nonpregnant women with COVID-19 (mean ± SD 32.8 ± 3.5 years versus 38.5 ± 7.9 years; $P = 0.11$), with no significant differences in age, race, ethnicity, or presence of a systemic rheumatic disease (Table 3). All patients with COVID-19 endorsed at least 1 COVID-19 symptom (Table 3). Loss of smell or taste was significantly higher in pregnant versus nonpregnant patients with COVID-19 (100% versus 44%; $P = 0.02$). Two of 5 pregnant women with COVID-19 experienced loss of smell or taste for >30 days, of whom 1 had a systemic rheumatic disease (scleroderma) and used immunomodulatory medication. In all, 8.9% of pregnant women without COVID-19 reported loss of taste or smell versus 100% of pregnant women with COVID-19 ($P < 0.01$). No pregnant patients with COVID-19 reported any emergency room (ER) visits or hospitalization. Among 28 (20%) nonpregnant women with COVID-19 who reported an ER visit or hospitalization, 1 patient was mechanically ventilated.

Among all pregnant women, 67% reported changes to prenatal obstetric/gynecologic (OB/GYN) care during the COVID-19 pandemic; in-person OB/GYN visits were rescheduled to telemedicine visits (15%) or canceled (18%). At survey completion, 22 of 61 pregnancies were complete (mean ± SD gestational age 38.1 ± 1 weeks). Pregnancy outcomes included the following: 10 (45%) vaginal deliveries, 5 (23%) cesarean sections, and 3 (14%) miscarriage/terminations (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24547/abstract). Only 1 of 5 patients

Table 2. Experiences of pregnant versus nonpregnant rheumatology outpatients from an academic hospital in New York City during the COVID-19 pandemic*

Characteristic	Nonpregnant women (n = 1,452)	Pregnant women (n = 61)	*P*
COVID-19 exposure			
Cohabitants (not including oneself), mean ± SD	2.26 ± 1.59	2.26 ± 1.30	0.97
Close contact with confirmed/suspected COVID-19 case	438 (30.2)	17 (27.9)	0.70
% of 24-hour day spent at home over past 2 weeks			0.45
>95	1,100 (75.8)	50 (82.0)	
76–95	249 (17.1)	9 (14.8)	
50–75	52 (3.6)	2 (3.3)	
<50	51 (3.5)	0 (0.0)	
High-risk of exposure at work (n = 1,159)	371 (33.4)	24 (46.2)	0.06
COVID-19 status			
Confirmed or suspected COVID-19‡	136 (9.4)	5 (8.2)	0.76
COVID-19 symptoms since January 1, 2020			
Any COVID-19 symptoms	894 (61.6)	37 (60.7)	0.89
Abdominal/belly pain	180 (12.4)	3 (4.9)	0.08
Chest pain	202 (13.9)	3 (4.9)	0.04†
Chills	289 (19.9)	7 (11.5)	0.10
Confusion/irritability	64 (4.4)	0 (0.0)	0.09
Cough	415 (28.6)	13 (21.3)	0.22
Diarrhea	265 (18.3)	7 (11.5)	0.18
Dizziness/lightheadedness	202 (13.9)	7 (11.5)	0.59
Fatigue or malaise	436 (30.0)	13 (21.3)	0.14
Fever	282 (19.4)	8 (13.1)	0.22
Headache or migraine	459 (31.6)	15 (24.6)	0.25
Joint pain	303 (20.9)	3 (4.9)	<0.01†
Loss of smell or taste	130 (9.0)	10 (16.4)	0.049†
Muscle aches	281 (19.4)	7 (11.5)	0.12
Runny nose	327 (22.5)	14 (23.0)	0.94
Shortness of breath	210 (14.5)	5 (8.2)	0.17
Sore throat or scratchy throat	458 (31.5)	15 (24.6)	0.25
Vomiting or nausea	137 (9.4)	2 (3.3)	0.10
Other	48 (3.3)	1 (1.6)	0.47
None of the above	557 (38.4)	24 (39.3)	0.88
Total duration of COVID-19 symptoms, days			0.03†
0–4	190 (21.3)	9 (24.3)	
5–9	220 (24.6)	16 (43.2)	
10–14	153 (17.1)	6 (16.2)	
>15	331 (37.0)	6 (16.2)	
Total COVID-19 symptom duration ≥10 days (n = 931)	484 (54.1)	12 (32.4)	0.01†
Health-related quality of life measures using PROMIS-29 score, mean ± SD			
Pain interference	53.17 ± 9.82	49.29 ± 8.39	<0.01†
Depression	53.57 ± 8.86	48.67 ± 7.59	<0.01†
Fatigue	55.02 ± 10.85	53.31 ± 9.41	0.23
Anxiety	60.47 ± 8.41	58.94 ± 8.06	0.17
Sleep disturbance	53.38 ± 6.71	53.52 ± 5.91	0.87

* Values are the number (%) unless indicated otherwise. *T*-tests and chi-square tests were used to calculate *P* values as appropriate. PROMIS-29 = Patient-Reported Outcomes Measurement Information System 29.
† Significant (*P* value threshold < 0.05).
‡ Confirmed COVID-19 case by nasopharyngeal polymerase chain reaction test; suspected COVID-19 case by health care provider.

with suspected or confirmed COVID-19 delivered by the time of survey completion, and she underwent a cesarean section. Neither of the 2 pregnant women with a systemic rheumatic disease and COVID-19 had delivered by the time of survey completion. In all, 23% of women who delivered stated that the pandemic affected their deliveries, most commonly by no visitors being permitted in the hospital or spouses/partners needing to depart soon after delivery (see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24547/abstract).

DISCUSSION

During the peak of the COVID-19 pandemic in New York City, pregnant women evaluated in our division of rheumatology reported similar COVID-19 prevalence and disease severity compared to nonpregnant women. Approximately 75% of patients had a systemic rheumatic disease, with no increased prevalence in pregnant patients or in those reporting COVID-19. In this cohort, 25% of patients seen by rheumatologists did not report having a systemic rheumatic disease. This relatively high number likely reflects the fact that patients are routinely evaluated at our tertiary referral center to rule out systemic rheumatic disease or to

Table 3. Characteristics of female rheumatology outpatients, ages 18–50 years, at an academic hospital in New York City with confirmed or suspected COVID-19, stratified by pregnancy status during the COVID-19 pandemic*

Characteristic	Nonpregnant women with COVID-19 (n = 136)	Pregnant women with COVID-19 (n = 5)	*P*
Sociodemographic factors			
Age, mean ± SD years†	38.48 ± 7.88	32.78 ± 3.50	0.11
Race‡			0.40
Asian/Indian subcontinent	5 (3.7)	1 (20.0)	
Black	12 (8.8)	0 (0.0)	
White	107 (78.7)	4 (80.0)	
Other§	10 (7.4)	0 (0.0)	
Unknown¶	2 (1.5)	0 (0.0)	
Ethnicity‡			0.67
Hispanic or Latino	22 (16.2)	0 (0.0)	
Not Hispanic or Latino	108 (79.4)	5 (100.0)	
Unknown¶	6 (4.4)	0 (0.0)	
Rheumatic disease history (by self-report)#			
Any systemic rheumatic disease	96 (70.6)	2 (40.0)	0.17
Systemic inflammatory arthritis (including adult-onset Still's disease, inflammatory arthritis, JIA, PsA, AS, RA)	55 (40.4)	1 (20.0)	0.65
Vasculitis, scleroderma, myositis (including dermatomyositis, polymyositis, other inflammatory muscle diseases)	9 (6.6)	1 (20.0)	0.31
Systemic lupus erythematosus	27 (19.9)	0 (0.0)	0.58
Antiphospholipid syndrome	6 (4.4)	0 (0.0)	1.00
COVID-19 exposures			
Close contact with confirmed/suspected COVID-19 case**	88 (64.7)	5 (100.0)	0.17
Percentage of 24-hour day spent at home over past 2 weeks			0.54
>95	100 (73.5)	4 (80.0)	
76–95	23 (16.9)	0 (0.0)	
50–75	10 (7.4)	1 (20.0)	
<50	3 (2.2)	0 (0.0)	
COVID-19 symptoms since January 1, 2020			
Any COVID-19 symptom	134 (98.5)	5 (100.0)	1.00
Abdominal/belly pain	41 (30.1)	0 (0.0)	0.32
Chest pain	68 (50.0)	0 (0.0)	0.06
Chills	83 (61.0)	1 (20.0)	0.16
Confusion/irritability	25 (18.4)	0 (0.0)	0.59
Cough	98 (72.1)	4 (80.0)	1.00
Diarrhea	68 (50.0)	1 (20.0)	0.37
Dizziness/lightheadedness	59 (43.4)	0 (0.0)	0.08
Fatigue or malaise	104 (76.5)	4 (80.0)	1.00
Fever	88 (64.7)	4 (80.0)	0.66
Headache or migraine	105 (77.2)	2 (40.0)	0.09
Joint pain	64 (47.1)	1 (20.0)	0.23
Loss of smell or taste	61 (44.9)	5 (100.0)	0.02††
Muscle aches	76 (55.9)	3 (60.0)	1.00
Runny nose	69 (50.7)	2 (40.0)	0.68
Shortness of breath	69 (50.7)	0 (0.0)	0.06
Sore throat or scratchy throat	91 (66.9)	1 (20.0)	0.05
Vomiting or nausea	35 (25.7)	1 (20.0)	1.00
Other	12 (8.8)	0 (0.0)	1.00
None of the above	2 (1.5)	0 (0.0)	1.00
Total COVID-19 symptom duration, days			<0.01††

(Continued)

Table 3. (*Cont'd*)

Characteristic	Nonpregnant women with COVID-19 (n = 136)	Pregnant women with COVID-19 (n = 5)	*P*
0–4	4 (2.9)	1 (20.0)	
5–9	18 (13.4)	3 (60.0)	
10–14	18 (13.4)	0 (0.0)	
>15	94 (70.1)	1 (20.0)	
Total COVID-19 symptom duration ≥10 days (n = 139)	112 (83.6)	1 (20.0)	<0.01††
COVID-19 disease severity			
Emergency room or hospitalization due to COVID-19	28 (20.6)	0 (0.0)	0.58
Hospitalization due to COVID-19	6 (4.4)	0 (0.0)	1.00

* Values are the number (%) unless indicated otherwise. *T*-tests and Fisher's exact tests were used to calculate *P* values as appropriate. AS = ankylosing spondylitis; JIA = juvenile idiopathic arthritis; PsA = psoriatic arthritis; RA = rheumatoid arthritis.
† Age was calculated from date of birth to April 24, 2020, the first date that the survey was emailed.
‡ Data on race and ethnicity were obtained by self-report from survey (when available) or from self-report in the electronic medical record.
§ Includes other race, in addition to American Indian/Alaskan Native and Native Hawaiian/Pacific Islander.
¶ Unknown includes those with missing values or who preferred not to answer.
# Systemic rheumatic diseases were not mutually exclusive.
** Confirmed COVID-19 case by nasopharyngeal polymerase chain reaction test; suspected COVID-19 case by health care provider.
†† Significant (*P*-value threshold < 0.05).

be medically evaluated prior to an orthopedic procedure. Pregnant women with COVID-19 had a shorter total duration of COVID-19 symptoms. Loss of taste or smell was significantly higher (in fact, observed uniformly) in pregnant women with COVID-19. Although both pregnant and nonpregnant women experienced loss of smell or taste, we did not observe a significant difference in loss of smell or taste in pregnant versus nonpregnant women without COVID-19, suggesting the clinical importance of this symptom among pregnant women infected with SARS–CoV-2. Furthermore, the majority of pregnant women reported that the pandemic influenced their prenatal OB/GYN care or delivery experience.

Our findings were consistent with 2 relatively small case series from China, in which no pregnant women with COVID-19 required mechanical ventilation or ICU admission (5,6). Additionally, the majority of 43 pregnant women (86%) with COVID-19 presenting to a hospital in New York City in March 2020 had mild disease (7). However, our findings contrast with recent population-based data from the Centers for Disease Control and Prevention demonstrating more mechanical ventilation and ICU admissions in pregnant versus nonpregnant women with COVID-19 (8). Similarly, Sweden's public health agency also reported that pregnant women with COVID-19 were 4–5 times more likely to be admitted to the ICU or receive mechanical ventilation (9). Our study may have been underpowered to demonstrate differences in COVID-19 prevalence and disease severity in pregnant versus nonpregnant women due to the small number of pregnant women positive for COVID-19 and because our patients were at lower risk of COVID-19, being predominantly non-Hispanic and nonobese.

Our results demonstrate some novel, hypothesis-generating findings. A shorter total duration of COVID-19 symptoms in pregnant patients with COVID-19, suggesting a milder disease course, is contrary to previous pandemics (3). How and whether the intersection of rheumatic disease and pregnancy may mitigate the severe inflammatory response documented in COVID-19 deserves further study. Loss of smell or taste is an early and specific neurologic manifestation of COVID-19 (10,11). In a study from 12 European centers of 417 patients with mild-to-moderate COVID-19, >85% of patients reported olfactory and gustatory dysfunctions, with higher prevalence in women (10). Estimates were lower in prospective data from 3,191 Korean patients with COVID-19, in whom acute loss of smell or taste occurred in 15% of those with asymptomatic-to-mild disease, also with higher prevalence in women (12). While some reports demonstrate symptomatic resolution within 3 weeks and a median time to recovery of taste and smell of 7 days (12), in our study, 2 of 5 pregnant rheumatology patients with COVID-19 had loss of smell or taste for at least 30 days. Furthermore, although olfactory changes have been observed in longitudinal pregnancy cohorts, increased smell sensitivity or abnormal smells, as opposed to loss of smell or taste, are typically reported (13,14). Whether pregnant patients with COVID-19 are more likely to experience loss of taste and smell and persistence of this loss warrants further investigation.

Of the 15 deliveries, 10 (67%) were vaginal deliveries, and 5 (33%) were cesarean sections, similar to the findings of a report from another New York hospital on 18 deliveries of patients from the general population who were positive for COVID-19 (55.5% uncomplicated normal vaginal deliveries and 44.4% cesarean sections) (4). We were underpowered to compare perinatal outcomes in those with or without COVID-19. Our study found that 67% of pregnant women experienced changes to their prenatal care, including cancelation and rescheduling of in-person visits to telemedicine visits. Nearly one-fourth of those who delivered indicated that the pandemic restrictions on visitors affected their delivery experience. Anxiety was higher than population norms in both pregnant and nonpregnant patients, which may be a reflection of

living in a pandemic hot spot. Pregnancy was associated with less depression, which has been shown in other US populations (15).

Our study has strengths and limitations. Using a patient self-report survey to assess COVID-19 prevalence and disease severity may not capture patients too ill to respond or those who died. Our relatively small sample from a single hospital may not be generalizable, and our patients were older, predominantly White, non-Hispanic, and with high household income and high educational attainment. Although we used telephone calls to increase representation of those without email access or with low literacy, our results should be confirmed in larger, more diverse cohorts. Given the lack of availability of COVID-19 testing in New York City at the time of survey administration, including only PCR-positive cases would have led to significant undercounting; however, using suspected COVID-19 cases may have resulted in some misclassification. We were also underpowered to perform subgroup analyses by race/ethnicity, systemic rheumatic disease history, medication-related factors, or positive PCR test status. Additionally, although we collected information on immunomodulatory and immunosuppressive medication use in the previous 6 months, medication dosages and the timing of medication modifications, including discontinuations, in relation to pregnancy could not be ascertained.

Strengths of our study include focusing exclusively on patients seen at a large division of rheumatology at an early US epicenter of the SARS–CoV-2 pandemic. Owing to the high COVID-19 prevalence in New York City at the time, we are able to provide the first study of the impact of COVID-19 on pregnant patients diagnosed with a systemic rheumatic disease by a rheumatologist. By timing our survey administration to coincide with the COVID-19 peak in New York City, we collected detailed self-report data in real time and minimized participant recall bias. We also included pregnant patients without COVID-19, providing a valid comparator group. Ongoing data collection in our longitudinal cohort will assess maternal and fetal outcomes in our pregnant rheumatology patients with and without COVID-19.

The COVID-19 pandemic offers a rare opportunity to study the impact of a serious infection on pregnant patients with rheumatic disease, which only few studies have previously addressed. Our finding of similar COVID-19 prevalence and disease severity in pregnant versus nonpregnant patients must be interpreted with caution but provides helpful data to women with systemic rheumatic diseases contemplating pregnancy during the pandemic. The shortened overall symptom duration and universal loss of smell or taste in pregnant women with COVID-19 also provides preliminary information to guide clinicians taking care of pregnant patients during the pandemic. As universal COVID-19 testing for pregnant patients who are admitted for delivery is increasingly performed (7), future studies may better assess the impact of the pandemic on perinatal and postpartum outcomes in this unique population.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Barbhaiya, Vitone, Frey, Feldman, Mandl.

Acquisition of data. Barbhaiya, Vitone, Frey, Jannat-Khah, Vega, Bykerk, Mandl.

Analysis and interpretation of data. Barbhaiya, Stamm, Vitone, Frey, Jannat-Khah, Levine, Feldman, Salmon, Crow, Bykerk, Lockshin, Sammaritano, Mandl.

REFERENCES

1. Cérbulo-Vázquez A, Figueroa-Damián R, Arriaga-Pizano LA, Hernández-Andrade E, Mancilla-Herrera I, Flores-Mejía LA, et al. Pregnant women infected with pandemic H1N1pdm2009 influenza virus displayed overproduction of peripheral blood CD69+ lymphocytes and increased levels of serum cytokines. PLoS One 2014;9:e107900.
2. Liu H, Wang LL, Zhao SJ, Kwak-Kim J, Mor G, Liao AH. Why are pregnant women susceptible to COVID-19? An immunological viewpoint. J Reprod Immunol 2020;139:103122.
3. Rasmussen SA, Jamieson DJ, Bresee JS. Pandemic influenza and pregnant women. Emerg Infect Dis 2008;14:95–100.
4. Assiri AM, Biggs HM, Abedi GR, Lu X, Bin Saeed A, Abdalla O, et al. Increase in Middle East respiratory syndrome–coronavirus cases in Saudi Arabia linked to hospital outbreak with continued circulation of recombinant virus, July 1–August 31, 2015. Open Forum Infect Dis 2016;3:ofw165.
5. Chen H, Guo J, Wang C, Luo F, Yu X, Zhang W, et al. Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records. Lancet 2020;395:809–15.
6. Liu D, Li L, Wu X, Zheng D, Wang J, Yang L, et al. Pregnancy and perinatal outcomes of women with coronavirus disease (COVID-19) pneumonia: a preliminary analysis. AJR Am J Roentgenol 2020;215:127–32.
7. Breslin N, Baptiste C, Gyamfi-Bannerman C, Miller R, Martinez R, Bernstein K, et al. COVID-19 infection among asymptomatic and symptomatic pregnant women: two weeks of confirmed presentations to an affiliated pair of New York City hospitals. Am J Obstet Gynecol MFM 2020;2:100118.
8. Ellington S, Strid P, Tong VT, Woodworth K, Galang RR, Zambrano LD, et al. Characteristics of women of reproductive age with laboratory-confirmed SARS-CoV-2 infection by pregnancy status: United States, January 22–June 7, 2020. MMWR Morb Mortal Wkly Rep 2020;69:769–75.
9. Collin J, Byström E, Carnahan A, Ahrne M. Public Health Agency of Sweden's brief report: pregnant and postpartum women with severe acute respiratory syndrome coronavirus 2 infection in intensive care in Sweden. Acta Obstet Gynecol Scand 2020;99:819–22.
10. Lechien JR, Chiesa-Estomba CM, De Siati DR, Horoi M, Le Bon SD, Rodriguez A, et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. Eur Arch Otorhinolaryngol 2020;277:2251–61.
11. Mao L, Jin H, Wang M, Hu Y, Chen S, He Q, et al. Neurologic manifestations of hospitalized patients with coronavirus disease 2019 in Wuhan, China. JAMA Neurol 2020;77:1–9.
12. Lee Y, Min P, Lee S, Kim SW. Prevalence and duration of acute loss of smell or taste in COVID-19 patients. J Korean Med Sci 2020;35:e174.
13. Nordin S, Broman DA, Olofsson JK, Wulff M. A longitudinal descriptive study of self-reported abnormal smell and taste perception in pregnant women. Chem Senses 2004;29:391–402.

14. Ochsenbein-Kölble N, von Mering R, Zimmermann R, Hummel T. Changes in olfactory function in pregnancy and postpartum. Int J Gynaecol Obstet 2007;97:10–4.

15. Mota N, Cox BJ, Enns MW, Calhoun L, Sareen J. The relationship between mental disorders, quality of life, and pregnancy: findings from a nationally representative sample. J Affect Disord 2008;109:300–4.

DOI 10.1002/acr.24726

Erratum

In the article by Jorge et al published in the January 2017 issue of *Arthritis Care & Research* (Depression and Progression of Subclinical Cardiovascular Disease in Systemic Lupus Erythematosus [pages 5–11]), the Results section of the Abstract was incorrect. The corrected Results are shown below. These changes do not affect any of the conclusions of the study.

Results. The SLE group had a higher rate of depression at both baseline and 5-year follow-up: 21% compared with 3% in the control group ($P < 0.0001$). When controlling for traditional CVD risk factors, the presence of depression at both baseline and 5-year follow-up correlated with increased progression of CIMT in the SLE group, but not in the control group. The adjusted mean increase in CIMT was 0.029 mm in the SLE group without depression versus 0.071 mm in the depressed SLE group ($P = 0.007$). There was no association between depression and carotid plaque in either group, with a calculated odds ratio for plaque progression in the depressed SLE group of 1.17 (95% confidence interval 0.45, 3.03) in the adjusted model.

In the same article by Jorge et al, there are also corrected values for Tables 1–5. These changes do not affect any of the conclusions of the study.

For Table 1 (Baseline demographics of participants with SLE and controls), the corrected values for SLE cases (n = 149) and controls (n = 126), respectively, are mean ± SD age: 43.2 ± 10.1 and 46.6 ± 10.0 ($P = 0.006$); mean ± SD systolic blood pressure (mm Hg): 118.0 ± 15.1 and 119.2 ± 15.7 ($P = 0.524$); no. (%) with hypertension: 78 (52.0) and 30 (24.0) ($P < 0.0001$); no. (%) with diabetes mellitus: 14 (9.0) and 5 (4.0) ($P = 0.077$); no. (%) statin use: 11 (7.0) and 6 (5.0) ($P = 0.369$); no. (%) aspirin use: 31 (21.0) and 9 (7.0) ($P = 0.0014$); mean ± SD SLICC/ACR DI 1.64 ± 1.8 (SLE cases); mean ± SD SLEDAI-2K 3.98 ± 3.63 (SLE cases); no. (%) CES-D ≥16, baseline only: 57 (38.0) and 14 (11.1.0) ($P < 0.0001$); no. (%) mean ± SD CES-D, baseline only: 13.95 ± 11.64 and 6.86 ± 6.44 ($P < 0.0001$); no. (%) CES-D ≥16, baseline and 5-year follow-up: 32 (21.5) and 4 (3.2) ($P < 0.0001$).

For Table 2 (Baseline demographics of SLE cases with and without depression at both baseline and follow-up visits), the corrected values for the not depressed SLE group (n = 117) and the depressed SLE group (n = 32), respectively, are mean ± age: 43.25 ± 10.6 years and 43.34 ± 8.38 years ($P = 0.996$); no. (%) statin use: 8 (6.84) and 3 (9.38) ($P = 0.703$); no. (%) aspirin use: 22 (18.8) and 9 (28.1) ($P = 0.325$); prednisone use: 38 (33) and 18 (56.3) ($P = 0.017$); mean ± SD prednisone dose (of 18 depressed and 38 not depressed patients taking prednisone at baseline visit): 12.74 ± 9.53 and 10.64 ± 6.45 ($P = 0.402$); mean ± SD SLICC/ACR DI: 1.52 ± 1.83 and 2.06 ± 1.79; mean ± SD SLEDAI-2K: 3.87 ± 3.76 and 4.38 ± 3.11 ($P = 0.489$); mean ± SD CES-D: 9.97 ± 8.79 and 28.53 ± 8.85 ($P < 0.0001$).

For Table 3 (CIMT at baseline and follow-up), the corrected values for sample sizes are n = 17 for not depressed SLE group, n = 33 for depressed SLE group, n = 122 for not depressed control group, and n = 4 for depressed control group. Depression status is defined as depressed at both baseline and 5-year follow-up.

For Table 4 (Progression of carotid plaque and CIMT at 5 years), the corrected OR (95% CI) values of carotid plague progression comparing depressed vs. not depressed cases, respectively, are 1.67 (0.74, 3.76) for unadjusted and 1.17 (0.45, 3.03) for adjusted. Since n = 4 for depressed controls, the OR was not able to be estimated. The corrected IMT mean change (mm) (95% CI) values, from baseline to 5 years, for depressed vs. not depressed cases, respectively, are 0.069 (0.043, 0.095) and 0.029 (0.015, 0.043) ($P = 0.008$) for unadjusted; 0.071 (0.044, 0.097) and 0.029 (0.015, 0.042) ($P = 0.007$) for adjusted; the corrected IMT mean change (mm) (95% CI) values, from baseline to 5 years, for depressed vs. not depressed controls, respectively, are 0.014 (–0.083, 0.111) and 0.025 (0.008, 0.043) ($P = 0.820$) for unadjusted; 0.023 (–0.076, 0.123) and 0.018 (–0.007, 0.044) ($P = 0.921$) for adjusted. For all values, n = 32 depressed, n = 117 not depressed for cases, and n = 4 depressed, n = 122 not depressed for controls. Depression status is defined as depressed at both baseline and 5-year follow-up.

For Table 5 (Presence of carotid plague at baseline and follow-up), the corrected values for no. of participants, no. (%) with plaque at baseline, no. (%) with plaque at 5 years, and no. (%) total participants with progression at 5 years are, respectively, n = 117; 43 (36.8), 50 (42.6), and 34 (29.1) for the not depressed SLE group; n = 32, 12 (37.5), 17 (53.1), and 13 (40.6) for the depressed SLE group; n = 122, 48 (39.3), 41 (33.6), and 19 (15.6) for the not depressed control group; and n = 4, 3 (75), 3 (75), and 0 (0.0) for the depressed control group. Depression status is defined as depressed at both baseline and 5-year follow-up.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 1013–1022
DOI 10.1002/acr.24365

American College of Rheumatology
Empowering Rheumatology Professionals

Use of Physical Therapy in Patients With Osteoarthritis in Germany: An Analysis of a Linkage of Claims and Survey Data

Hannes Jacobs,[1] Johanna Callhoff,[2] Katinka Albrecht,[2] Anne Postler,[3] Joachim Saam,[4] Toni Lange,[5] Jens Goronzy,[3] Klaus-Peter Günther,[3] and Falk Hoffmann[1]

Objective. To examine the utilization of physical therapy (PT) and predictors for its use in individuals with osteoarthritis (OA) while focusing on sociodemographic and disease-related factors.

Methods. For this cross-sectional study, 657,807 patients (age 30–79 years) diagnosed with hip, knee, or polyarticular OA were identified in claims data. In 2016, a questionnaire including information on disease status, demography, and socioeconomics was sent to a random sample of 8,995 patients stratified by sex, age, and type of diagnosis. Claims data from 2016 included the utilization and type of PT, as well as the prescribing medical specialist, and were linked to questionnaire data. Multivariable logistic regression was conducted to determine variables associated with the use of PT.

Results. In total, 3,564 (40%) patients completed the questionnaire and agreed to linking questionnaire and claims data (69% female, mean age 66.5 years). In 2016, 50% of the study population received PT at least once, and women received it more frequently than men (53% versus 43%). Most PT was prescribed by orthopedists (45%) and general practitioners (32%). Multivariable logistic regression showed that women, higher household income, having both hip and knee OA, lower functional status, higher disease activity, and individuals living in the eastern, southern, and western states of Germany were associated with an increased utilization of PT.

Conclusion. Considering current guideline recommendations and that more than one-third of OA patients with high functional impairment and/or pain did not receive PT in the last 12 months, there is considerable potential for improvement. This is especially true for men and individuals with a low income.

INTRODUCTION

Osteoarthritis (OA) is the most common joint disorder and a leading cause of disability in older adults (1,2). Worldwide prevalence in individuals age >60 years is estimated at 10% in men and 18% in women (3). Overweight individuals are more frequently affected by OA (4,5). Because of both demographic changes and increasing populations with obesity in high-income countries, an increasing number of patients with this syndrome is to be expected (6).

OA is characterized by progressive, degenerative changes in the joints associated with pain, restriction in movement, and as a result, a diminished quality of life (7,8). As a consequence, primary goals in OA therapy are pain reduction and long-term preservation of joint function. Guidelines of OA management recommend education/self-management, analgesic and antiinflammatory medication, and low-impact behaviors (e.g., weight reduction if overweight or obese) as nonsurgical treatments (9–13). Furthermore, physical therapy (PT) is seen as one of the key elements of OA management (6,14). Exercise therapy as one of the main interventions in PT is particularly helpful in decreasing pain and preserving joint motion, for which high-quality evidence has been available for the past decade (15–18).

Data on the utilization of PT in patients with OA are scarce and vary across studies, and the populations included

Supported by the Federal Ministry of Education and Research within the PROCLAIR research network (grant 01EC1405).

[1]Hannes Jacobs, MSc, Falk Hoffmann, MPH: Carl von Ossietzky University, Oldenburg, Germany; [2]Johanna Callhoff, MD, Katinka Albrecht, MD: German Rheumatism Research Centre, Berlin, Germany; [3]Anne Postler, MD, Jens Goronzy, MD, Klaus-Peter Günther, MD: University Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany; [4]Joachim Saam, Dipl Sportwiss: Barmer Institute for Health System Research and Barmer Statutory Health Insurance Fund, Schwäbisch-Gmünd, Germany; [5]Toni Lange, MSc: University Medicine, University Hospital, and Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany.

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

Address correspondence to Hannes Jacobs, MSc, Department of Health Services Research, Carl von Ossietzky University Oldenburg, Ammerländer Heerstrasse 114-118, 26129 Oldenburg, Germany. Email: Hannes.Jacobs@uol.de.

Submitted for publication February 3, 2020; accepted in revised form June 23, 2020.

SIGNIFICANCE & INNOVATIONS

- The current study combined information on the utilization of physical therapy (PT) retrieved from claims data with information on disease-related and sociodemographic factors from patient-reported questionnaire data.
- In Germany, 50% of patients with osteoarthritis received PT at least once in the last 12 months, mostly prescribed by orthopedists and general practitioners.
- Female sex, higher household income, lower functional status, increased disease activity, and affection of knee and hip joints were associated with higher utilization of PT.
- More than one-third of patients with high functional limitations and/or pain (with a Western Ontario and McMaster Universities Osteoarthritis Index score ≥54.6) did not receive PT in the last year.

are diverse. Carter et al reported outpatient PT use in 7% of patients with musculoskeletal conditions, but only 2% of them had OA (19). In a large Taiwanese cohort of 25,000 incident OA patients, 25% received PT 12 months after diagnosis (20). In a Canadian cohort of adults with at least moderately severe hip or knee OA, 19% received PT within the past year; of those, 65% had total joint replacement (TJR) surgery in the past year, and only 17% did not have TJR (21). Iversen et al reported the use of PT in 52% of patients with symptomatic knee OA, but these data refer to self-reports from patients enrolled in an exercise trial (22). In Germany, claims data showed the use of PT in 49% of patients with knee or hip OA 1 year before TJR (23).

Factors associated with the utilization of PT in patients with OA have been evaluated in a few studies. Most previous studies show that women receive PT more frequently than men (20–24), whereas Carter et al observed no differences (19). Being of younger age was another factor reported with a higher utilization of PT (25). Other studies showed no association between younger age and PT but less frequent utilization in the elderly (>65 years) (19,20) or no correlation at all (22,23). Higher income or education were associated with higher PT utilization (19,20,22). However, results are inconsistent (19–26). Other important clinical and patient-reported factors such as functional status, comorbidities, and body mass index (BMI) have not been examined sufficiently. Taken together, most studies on the utilization of PT in OA include selected populations, and knowledge on the factors associated with its use is limited.

Therefore, this study aimed to evaluate the utilization of PT in a more representative sample of patients with OA as well as to identify factors associated with higher utilization by taking advantage of the linkage of claims data and self-reported patient outcomes.

MATERIALS AND METHODS

Study design, sample, and study population. For this cross-sectional study, data were obtained from Germany's second largest nationwide statutory health insurance fund (BARMER), which insured 9.5 million individuals in 2016 (27). Study participants had to be insured continuously in 2013 and 2014, and an outpatient claim with an OA diagnosis (International Classification of Diseases, Tenth Revision, German Modification [ICD-10-GM] code M15 [polyarticular], M16 [hip], M17 [knee]) was required in at least 2 quarters of the year 2014. We considered only these types of OA because they are highly prevalent in the general population and have great impact on an individual's quality of life. Of these 657,807 OA patients, a random sample (n = 9,734) was drawn stratified by sex (female, male), age (30–39, 40–49, 50–59, 60–69, 70–79 years), and type of diagnosis (M15, M16, M17). After exclusion of deceased persons or individuals who had changed insurance, 8,995 questionnaires were sent in June 2016, with a reminder sent in September 2016 to those who had not answered within 8 weeks. A positive ethics vote was issued in 2015 by the Ethics Committee of the Charité Universitätsmedizin Berlin (EA1 / 051/15).

Data collection. *Claims data.* All information from claims data referred to the year of the survey (2016). In Germany, PT is prescribed by a physician. The prescription (including the referring medical specialist, type of interventions, and number of treatments) is handed out to the patient, who then seeks a PT practice. After treatment, the PT practice transfers the prescription to the health insurance fund in order to get reimbursed, and then this procedure is recorded in claims data. The type of PT is coded according to a standardized index for therapeutic services (28) (therapist massage: X01; manual lymphatic drainage: X02; therapeutic exercise: X03–X10; traction therapy: X11; manual therapy: X12; electrotherapy: X13; and thermotherapy: X15).

The referring group of the prescribing medical specialist was analyzed based on specific national provider identifying numbers (29) (general practitioner [01–03], orthopedist [06, 10–12], rheumatologist [31], or any other medical specialist [any other national provider identifying number]). Physicians who have used special accounting rheumatologic codes (13690, 13691, 13692, 13700, 13701, 99012) were considered as rheumatologists. Specialist medical care was defined according to whether the patient had visited the corresponding specialist at least once in the last 12 months.

Independent of stratification, 4 mutually exclusive, hierarchical groups were defined for analyses: polyarticular OA, hip and knee OA, hip OA, and knee OA. Patients with hip and knee OA were only considered in the combined group. Comorbidities (at least 1 outpatient diagnosis) were identified via ICD-10-GM codes and classified by the Elixhauser comorbidity index (30). Data on demography (age, sex, residential area [North: Bremen, Hamburg, Lower Saxony, Schleswig-Holstein; East: Berlin, Brandenburg, Mecklenburg-Vorpommern, Saxony, Saxony-Anhalt, Thuringia;

South: Baden-Wuerttemberg, Bavaria; West: Hesse, North Rhine-Westphalia, Rhineland-Palatinate, Saarland]) were retrieved.

Survey data. The questionnaire included information on sociodemographic characteristics such as household income and size of town. To evaluate functional status, the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) was used. This consists of 24 questions with 3 subscales, and it reports the components of pain, stiffness, and joint functionality as a total score (range 0–100, 100 representing the worst outcome) (31). Because individuals might have >1 symptomatic joint, they were instructed to answer based on the most severely affected joint. Additionally, BMI (obtained from height and weight) was requested, and disease activity was determined by the number and location (0–29) of chronically painful joints (i.e., continuous pain for at least 3 months in the last 2 years). Psychological well-being/presence of depressive symptoms was assessed using the WHO-5 Well-Being Index (WHO-5) (32). Scores were transformed (range 0–100) and categorized as moderate-to-severe (0–28), mild (29–50), and no (>50) depressive symptoms, an approach that has been validated in previous studies (32–34).

Statistical analyses. Because we used stratified sampling, the total number of individuals returning questionnaires who gave their consent for linking questionnaire data to claims data was weighted according to sex, type of diagnosis, and age group distribution of the total OA population (n = 657,807) in the claims data for all analyses. Characteristics of the study population were analyzed descriptively (percentages and SEM) and stratified by sex. The SEM was used instead of the SD due to the stratified nature of the study sample. The proportions of PT utilized were calculated with a 95% confidence interval (95% CI) and stratified by sex, residential area, and sociodemographic and disease-related factors. The proportions of prescribing medical specialists were analyzed with a 95% CI as well as the mean number of prescriptions and PT treatments. Univariable logistic regression analysis was used to determine the main demographic (age, sex, residential area, size of town), disease-related (WOMAC score, number of chronically painful joints, type of diagnosis, comorbidities, psychological well-being [WHO5-score]), lifestyle factor (BMI), and socioeconomic (household income) characteristics that are associated with the utilization of PT. Finally, all variables were included in a multivariable model. Odds ratios were calculated with a 95%

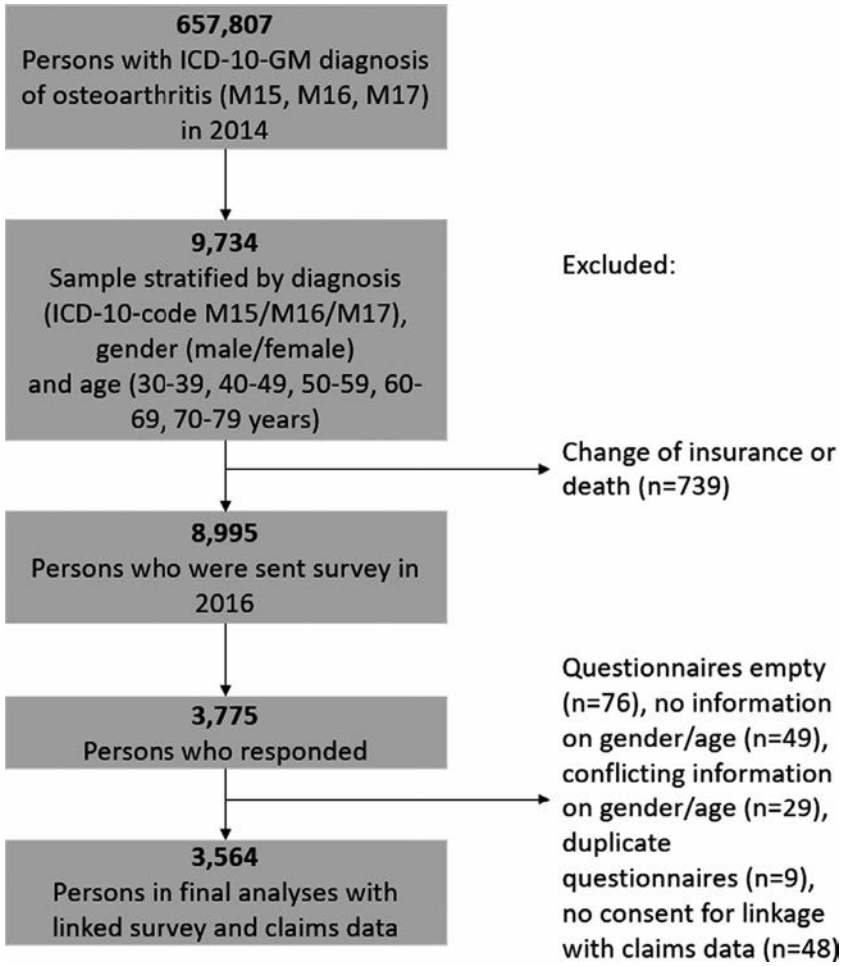

Figure 1. Flow chart of the study population. ICD-10-GM = International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, German version.

CI. Nonoverlapping 95% CIs were considered statistically significant. For univariable and multivariable logistic regression, missing values for the WOMAC score, number of painful joints, WHO-5 score, household income, BMI, and size of town were imputed with multiple imputations (n = 10 imputations) (PROC MI) using the fully conditional specification method, assuming that data were missing at random. Data analyses were performed with SAS Enterprise Guide, version 9.4, using complex survey designs (SURVEYMEANS, SURVEYFREQ, and SURVEYLOGISTIC) that incorporated the stratified design into the analyses.

RESULTS

Response. A total of 8,995 individuals received the questionnaire. Of those, 3,775 (42%) responded, and after excluding invalid questionnaires and those who did not give their consent for linking questionnaire data to claims data, a total of 3,564 individuals were included in the analyses (Figure 1).

Individuals who responded were slightly older than those who had not responded (67.2 years versus 65.8 years), and the proportion of women was slightly higher (71% versus 69%).

Table 1. Characteristics of the study population*

Characteristic	Women (n = 1,952, 69.4%)	Men (n = 1,612, 30.6%)	Total (n = 3,564)
Age, mean ± SEM years (n = 3,564)	66.8 ± 0.08	65.9 ± 0.08	66.5 ± 0.06
Diagnosis (n = 3,564)			
Knee OA	48.3 (45.5–51.2)	52.2 (49.3–55.1)	49.5 (47.4–51.7)
Hip OA	21.9 (19.6–24.1)	28.9 (26.2–31.5)	24.0 (22.2–25.8)
Polyarticular OA	16.3 (14.4–18.1)	7.3 (6.4–8.3)	13.5 (12.2.14.9)
Hip and knee OA	13.5 (11.6–15.5)	11.6 (9.7–13.5)	12.9 (11.5–14.4)
Medical consultation (n = 3,564)†			
General practitioner	97.9 (97.2–98.6)	98.0 (97.2–98.7)	97.9 (97.4–98.4)
Orthopedist	65.3 (62.6–68.0)	59.9 (57.0–62.7)	63.6 (61.6–65.7)
Rheumatologist	8.1 (6.6–9.5)	4.4 (3.3–5.5)	7.0 (5.9–8.0)
Other	96.4 (95.4–97.4)	95.8 (94.8–96.9)	96.2 (95.5–97.0)
Household income, € (n = 3,313)			
<1,500	35.0 (32.2–37.8)	23.0 (20.4–25.5)	31.3 (29.2–33.4)
1,500–3,200	54.0 (51.1–56.9	63.0 (60.1–65.9)	56.8 (54.6–59.0)
>3,200	11.0 (9.4–12.6)	14.0 (12.0–16.0)	11.9 (10.6–13.2)
WOMAC score, mean ± SEM (n = 2,879)	39.6 ± 0.69	36.7 ± 0.70	38.7 ± 0.52
No. of painful joints (n = 3,417)			
0	24.3 (21.7–26.8)	26.9 (24.3–29.6)	25.1 (23.1–27.0)
1–4	40.9 (38.1–43.7)	48.7 (45.7–51.7)	43.8 (41.1–45.5)
>4	34.8 (32.1–37.5)	24.4 (21.9–26.9)	31.6 (29.6–33.7)
Depressive symptoms (WHO-5) (n = 3,334)			
No	52.6 (49.6–55.5)	57.3 (54.3–60.3)	54.0 (51.8–56.2)
Mild	23.6 (21.1–26.1)	19.7 (17.3–22.0)	22.4 (20.5–24.3)
Moderate to severe	23.8 (21.4–26.3)	23.0 (20.5–25.5)	23.6 (21.7–25.4)
BMI, mean ± SEM (n = 3,473)	28.4 ± 0.16	28.5 ± 0.13	28.4 ± 0.11
Elixhauser comorbidity index score (ref. 30) (n = 3,564)			
0–1	22.3 (20.1–24.6)	24.5 (22.1–26.9)	23.0 (21.3–24.7)
2–4	49.9 (47.1–52.7)	44.9 (42.0–47.8)	48.4 (46.2–50.5)
5–7	21.8 (19.4–24.2)	22.6 (20.2–25.1)	22.1 (20.2–23.9)
>7	6.0 (4.5–7.4)	8.0 (6.3–9.6)	6.6 (5.5–7.7)
Residential area (n = 3,563)‡			
North	14.7 (12.7–16.7)	13.1 (11.1–15.0)	14.2 (12.7–15.7)
East	30.3 (27.6–32.9)	28.8 (26.1–31.5)	29.8 (27.8–31.8)
South	19.3 (17.1–21.6)	21.7 (19.3–24.1)	20.1 (18.3–21.8)
West	35.7 (33.6–39.2)	36.4 (33.6–39.2)	35.9 (33.9–37.9)
Population of town (n = 3,451)			
<5,000	25.0 (22.5–27.4)	30.6 (28.0–33.3)	26.7 (24.9–28.6)
<20,000	25.4 (22.9–27.8)	26.3 (23.7–28.9)	25.6 (23.8–27.5)
<100,000	25.0 (22.4–27.5)	22.0 (19.5–24.4)	24.0 (22.1–25.9)
<500,000	12.9 (10.9–14.9)	10.5 (8.7–12.3)	12.1 (10.7–13.6)
≥500,000	11.8 (10.0–13.7)	10.6 (8.8–12.4)	11.4 (10.0–12.9)

* Values are the percentage (95% confidence interval) unless indicated otherwise. Numbers are not weighted, percentages are weighted. BMI = body mass index; OA = osteoarthritis; WHO-5 = WHO-5 Well-Being Index; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index (Q1 = score ≤22.8; Q2 = score 22.9–39.5; Q3 = score 39.6–55.0; Q4 = score >55.0).
† Medical consultation: at least 1 consultation in 2016.
‡ North: Bremen, Hamburg, Lower Saxony, Schleswig-Holstein; East: Berlin, Brandenburg, Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt, Thuringia; South: Baden-Wuerttemberg, Bavaria; West: Hesse, North Rhine-Westphalia, Rhineland-Palatinate, Saarland.

Prescription of opioids (15% versus 14%) and number of comorbidities according to the Elixhauser comorbidity index (median 2.5 versus 2.7) were comparable. There were larger differences in the prescription of nonsteroidal antiinflammatory drugs (48% versus 37%) and outpatient orthopedic care (58% versus 45%).

Characteristics of the study population. Table 1 gives an overview of the patients' main sociodemographic and disease-related characteristics stratified by sex. A total of 69% were female, and the mean age was 66.5 years. Overall, 31% had a household income <€1,500. In 2016, 98% of the patients were treated at least once by a general practitioner (96% by other specialists, 64% by orthopedists, and 7% by rheumatologists). The mean WOMAC score was 38.7. In total, 25% had no, 44% had 1–4, and 32% had >4 chronically painful joints in the last 2 years. A total of 22% had mild signs of depression, and 24% had moderate-to-severe signs of depression according to the WHO-5. The proportion of women with a household income <€1,500 was higher than that of men (35% versus 23%). Women were more often diagnosed with polyarticular OA (16% versus 7%), and men were more frequently affected by hip OA (29% versus 22%).

Frequency of prescriptions for PT. Overall, 50% of the study population received PT at least once in 2016. Therapeutic exercise was the type of PT prescribed most commonly (36%), followed by manual therapy (16%) and thermotherapy (13%) (see Supplementary Tables 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24365/abstract). The mean number of prescriptions accounted for was 3.9 (95% CI 3.7–4.1), and the mean number of PT treatments was 24.8 (95% CI 23.2–26.3). Regarding prescribing medical specialist, most PT was prescribed by orthopedists (45%; 32% by general practitioners, 20% by other medical specialists, and 3% by rheumatologists).

Women received PT more frequently than men (54% versus 43%) (Table 2). This pattern was observed in every type of PT (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24365/abstract). The proportion of patients with PT increased by age, from 43% in the age group 30–39 years to 53% in patients 70–79 years, with the exception of those 60–69 years of age (46%). Patients living in the eastern (59%) or southern (52%) states of Germany received PT more frequently than those living in the northern (48%) or western (42%) states. Manual therapy and thermotherapy were prescribed much more frequently in the eastern states of Germany, but therapeutic exercise was prescribed to a lesser extent than in the other states (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24365/abstract). No differences were observed in size of town.

The utilization of PT increased in patients with a higher WOMAC score, high numbers of painful joints, comorbidities, and depressive symptoms (Figure 2 and Supplementary Table 3, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24365/abstract). The proportion of patients who had PT was smallest in those with a WOMAC score ≤22.8 (37%) and highest in those with a WOMAC score >55.0 (63%). PT was also prescribed more frequently if the knee and hip were affected by OA.

Table 2. Utilization of physical therapy (PT) in 2016 depending on sociodemographic factors*

Sociodemographic factors	Proportion with PT
Overall	50.0 (47.9–52.2)
Age, years	
30–39	43.4 (37.6–49.2)
40–49	45.6 (41.2–50.1)
50–59	50.2 (46.2–54.1)
60–69	46.1 (42.5–49.8)
70–79	53.1 (49.5–56.7)
Sex	
Women	53.1 (50.3–55.9)
Men	43.1 (40.2–46.0)
Household income, €	
<1,500	50.0 (45.8–54.2)
1,500–3,200	50.2 (47.2–53.1)
>3,200	49.8 (44.1–55.4)
Population of town	
<5,000	51.5 (47.4–55.6)
<20,000	47.9 (43.7–52.2)
<100,000	45.5 (40.9–50.1)
<500,000	53.5 (47.1–60.0)
≥500,000	57.3 (50.9–63.7)
Residential area†	
North	48.3 (42.6–54.0)
East	58.9 (55.0–62.9)
South	52.4 (47.6–57.1)
West	42.0 (38.6–45.5)

* Values are the percentage (95% confidence interval). Percentages are weighted.
† North: Bremen, Hamburg, Lower Saxony, Schleswig-Holstein; East: Berlin, Brandenburg, Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt, Thuringia; South: Baden-Wuerttemberg, Bavaria; West: Hesse, North Rhine-Westphalia, Rhineland-Palatinate, Saarland.

Univariable and multivariable logistic regression: predictors for utilization of PT. Univariable logistic regression models showed that female sex, hip and knee affection, higher WOMAC score, number of chronically painful joints, depressive symptoms, ≥5 comorbidities, and living in north, east, or south Germany were associated with higher PT utilization, whereas age, household income, BMI, and town size had no influence (Table 3). Multivariable logistic regression analysis revealed that female sex, having hip or knee OA, a higher WOMAC score, an increasing number of painful joints, and residential area were associated with receiving PT to a higher extent. Furthermore, patients with a household income <€1,500 were associated with less frequent utilization of PT. The number of comorbidities and psychological well-being did not remain statistically significant in the multivariable model, and age and BMI still had no influence on the frequency of utilization of PT.

DISCUSSION

The linkage of claims data with patient-reported outcomes enabled us to assess the impact of disease-related and

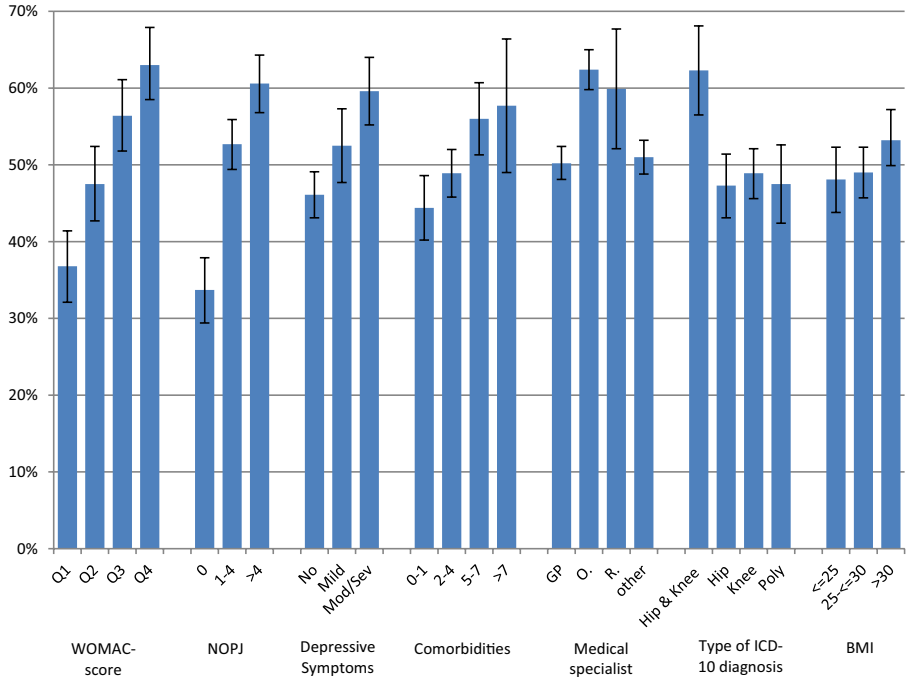

Figure 2. Proportion of physical therapy depending on disease-related factors in percentages with 95% confidence intervals. Percentages are weighted. Comorbidities are classified by the Elixhauser comorbidity index. Depressive symptoms (no, mild, moderate [mod]/severe [sev]) are grouped by the WHO-5 Well-Being Index (WHO-5). BMI = body mass index (kg/m^2); GP = general practitioner; ICD-10 = International Statistical Classification of Diseases and Related Health Problems, Tenth Revision; NOPJ = number of painful joints; O = orthopedist; other = any other medical specialist; poly = polyarticular osteoarthritis; R = rheumatologist; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index (Q1 = score ≤22.8; Q2 = score 22.9–39.5; Q3 = score 39.6–55.0; Q4 = score >55.0). Error bars indicate 95% confidence intervals.

socioeconomic parameters on the utilization of PT in a large and heterogeneous population of individuals with OA. In 2016, 50% of the study population received PT at least once. In addition to higher utilization in women and individuals with high household income, more frequent utilization with higher disease burden was observed both in terms of the extent of pain and functional impairment and of the number of joints affected. Nevertheless, one-third of patients with high disease severity did not receive PT in the last 12 months.

Compared to previous studies, our results on the utilization of PT were high (19–21,35). For example, in a Taiwanese cohort, Yeh et al observed that 25% of incident OA patients received PT 12 months after their diagnosis (20). Less utilization may be explained by the fact that disease burden and disease activity in incident cases is lower. This corresponds to findings from a Canadian cohort of patients with at least moderately severe hip or knee OA, in which only 17% of adults without TJR but 65% of those with TJR received PT in the past year (21). In Germany, claims data from the same statutory health insurance fund as in the current study revealed comparably high PT utilization as the frequency identified here (49%) but in patients with knee or hip OA 1 year before TJR (23). In the US, comparably high utilization of PT as seen in the current study was observed in a population of patients with OA or rheumatoid arthritis by Hagglund et al (including occupational therapy) (39%) (36) and by Iversen et al (52%) (22), although the latter included study participants recruited for a clinical trial of exercise to manage symptomatic knee OA, which potentially led to a greater preference for PT in managing knee symptoms.

Utilization of PT in patients with OA has been associated with a variety of factors. For example, most previous studies have shown that men receive PT significantly less than women (20,22,23,25,35). We also observed a higher utilization of PT in women, confirming these findings. On the one hand, a plausible explanation may be that the time-consuming nature of PT is a critical barrier for patients, especially for men, who tend to prefer fast pain relief interventions such as drug treatment or surgery (37). Furthermore, Yeh et al speculated that men were less flexible in scheduling PT appointments because they worked more hours than women (20). On the other hand, women prefer holistic approaches (38) and are less inclined to receive invasive treatments (39). In addition, Hawker et al estimated that the potential need for arthroplasty was more than twice as great among women than among men (37), which also suggests that women are more receptive of conservative treatments like PT.

Age is another parameter that is discussed as an influencing factor for the utilization of PT. Rommel et al observed a higher utilization with increasing age for all residents in Germany age 18–79 years (35). Other studies showed no association between younger age and PT but less utilization in the elderly (>65 years) (19,20,40,41), which is surprising because function decreases

Table 3. Factors associated with a higher utilization of physical therapy: results from univariable and multivariable logistic regression analyses (n = 3,564)*

Characteristic	Reference	Odds ratio (95% confidence interval)	
		Univariable analysis	Multivariable analysis
Sex			
Women	Men	1.50 (1.29–1.68)†	1.47 (1.23–1.75)†
Age	Per 10 years	1.05 (0.99–1.13)	1.05 (0.96–1.15)
Diagnosis			
Hip and knee OA	Hip OA	1.59 (1.26–2.02)†	1.50 (1.09–2.07)†
Knee OA	Hip OA	1.01 (0.86–1.18)	1.03 (0.82–1.28)
Polyarticular OA	Hip OA	0.88 (0.73–1.06)	0.87 (0.66–1.15)
Household income, €			
1,500–3,200	<€1,500	1.06 (0.87–1.29)	1.21 (0.98–1.50)
>3,200	<€1,500	1.11 (0.85–1.46)	1.53 (1.13– 2.07)†
WOMAC score	Per 10 units	1.17 (1.12–1.23)†	1.11 (1.05–1.18)†
No. of painful joints			
1–4	None	2.08 (1.66–2.62)†	2.85 (1.44–2.37)†
>4	None	2.88 (2.26–3.67)†	2.21 (1.68–2.91)†
Depressive symptoms (WHO-5)			
Mild	No	1.24 (1.99–1.55)†	1.00 (0.78–1.28)
Moderate to severe	No	1.68 (1.35–2.09)†	1.20 (0.92–1.57)
BMI	Per 10 units	1.11 (0.94–1.30)	0.91 (0.76–1.10)
Elixhauser comorbidity index score (ref. 30)			
2–4	0–1	1.15 (0.98–1.34)	1.1 (0.87–1.38)
5–7	0–1	1.32 (1.98–1.60)†	1.24 (0.92–1.67)
>7	0–1	1.78 (1.29–2.44)†	1.24 (0.80–1.91)
Residential area‡			
North	West	1.56 (1.27–1.91)†	1.38 (1.04–1.85)†
East	West	2.10 (1.78–2.47)†	1.95 (1.55–2.46)†
South	West	1.55 (1.30–1.85)†	1.57 (1.22–2.02)†
Population of town			
<20,000	<5,000	0.86 (0.67–1.09)	0.85 (0.66–1.08)
<100,000	<5,000	0.78 (0.91–1.00)	0.84 (0.65–1.09)
<500,000	<5,000	1.07 (0.79–1.45)	1.09 (0.79–1.50)
≥500,000	<5,000	1.26 (0.92–1.71)	1.12 (0.82–1.54)

* Missing values imputed from 3% in body mass index (BMI) to 19% in Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). OA = osteoarthritis; WHO-5 = WHO-5 Well-Being Index.
† Odds ratios of variables significantly associated with the utilization of physical therapy.
‡ North: Bremen, Hamburg, Lower Saxony, Schleswig-Holstein; East: Berlin, Brandenburg, Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt, Thuringia; South: Baden-Wuerttemberg, Bavaria; West: Hesse, North Rhine-Westphalia, Rhineland-Palatinate, Saarland.

with increasing age, which would suggest a greater need for PT. A possible explanation is that low mobility or functional limitations of older adults may represent a hurdle for reaching the PT practice. Consequently, older adults' preference for treatment may shift away from PT. However, there are also several studies showing no association at all between age and the utilization of PT (22,23,42). The results of the current study also showed no clear association between age and the utilization of PT.

Patients affected by both hip and knee OA had a significantly higher rate of PT utilization. It is likely that patients having both diagnoses are affected more with regard to pain and functional outcomes. This corresponds to our finding of a clear and statistically significant association between patient-reported limitation of functional status and the utilization of PT, confirming the results of previous studies of patients with musculoskeletal conditions (19,21,22,43,44). Both higher disease status (WOMAC score) and higher disease activity (number and location of chronically painful joints) were associated with greater utilization of PT. However, this also means that more than one-third of those with highly functional limitations and/or with >4 chronically painful joints did not receive PT in the last 12 months. For those patients, it might not be possible to receive PT continuously for years, but no contact at all with a physical therapist within 1 year might reduce the chances of pain relief, better joint function, and therefore improved quality of life (15,18).

Level of education and income are proxy measures for socioeconomic status (SES). To avoid collinearity, we selected household income to represent SES, and therefore, education was omitted from the model. While some studies provide evidence for an impact of education on the utilization of PT in the general population and in patients with musculoskeletal conditions or OA (19,22,36,40), few data are available on income (22). In

the multivariate analysis of the current study, household income <€1,500 was negatively associated with PT utilization compared to an income of >€3,200 per month, although in the univariate analysis, the association was not significant. In the high-income group, WOMAC scores were lower, and depression was present less often. Both lower WOMAC scores and less depression are associated with low levels of PT utilization, which explains this finding. In a comparable analysis, an association between household income and PT utilization was not found in individuals with rheumatoid arthritis (42). A possible explanation might be the smaller proportion of women and the older study population in the current study.

A few studies have investigated the influence of the size of town on the utilization of PT (20,40). For example, in community-based individuals ≥65 years of age, Freburger et al observed no differences in the utilization of PT in patients living in a rural or metropolitan area, but those living in a metropolitan area had a greater number of PT visits (40). Carter et al reported a pattern with a positive association from residing in an urban area for patients with musculoskeletal conditions (19). We found no variations in the size of town in which patients resided, but we observed regional differences in PT utilization. Especially in the eastern states of Germany, PT seems to play an essential role in the care of patients with musculoskeletal conditions. These regional variations were also found for PT utilization in the general population, both in questionnaire data (35) and claims data (45), as well as for specific diagnoses (23,25). Furthermore, results from the current study show that manual therapy and thermotherapy in particular are more frequently prescribed in East Germany, while exercise therapy is prescribed less often.

The main strength of the current study is the linkage of claims data from a nationwide statutory health insurance fund to questionnaire data from patients with OA. Utilization of PT, the prescribing medical specialist, and residential area are validly coded in claims data. Questionnaire data provided valuable additional information on disease-related, psychological, and lifestyle factors. Consequently, we were able to link the utilization of PT to patient-reported information such as pain and functional status, which, as far as we know, has not been done before.

Some limitations have to be considered as well, mainly relating to the data used. The representativeness of the data needs to be discussed, as there are differences in the sociodemographic characteristics between several funds (46). Compared to the general population, individuals insured with Barmer were slightly older (mean age 53.7 years versus 52.4 years) and less often male (39% versus 47%), but there was no difference regarding higher education (28% versus 29%) (47). This might, to some extent, have led to differences in PT utilization. In addition, we were not able to determine whether a prescription of PT was related to the OA diagnosis coded in claims data or whether it was due to some other indication, which could have led to a higher utilization of PT. However, PT utilization in the German general population is lower (35); thus, a large proportion of PT visits should be due to OA. Moreover, medical specialists could have recommended other types of exercise (e.g., muscle strengthening in a gym), which were not available in claims data. Furthermore, individuals may have received PT in prior years and had already established an individualized program for themselves independently, or they may have worked with another type of health care practitioner. However, when evaluating the utilization of PT in the 2 years before the survey, 53% and 51% of the study population received PT at least once in 2014 and 2015, which is the same as in 2016. Utilization regarding sex, age, and residential region was also comparable to that in 2016, which suggests that this limitation did not hamper our findings. Another limitation is that the patients who responded are not completely representative of the whole sample. It is possible that the responders on average had a higher disease burden or greater dissatisfaction with their health care, which could have led to an overestimation of PT utilization for the general population of individuals with hip, knee, or polyarticular OA. However, we were able to compare characteristics of responders and nonresponders using claims data. Finally, the data did not contain additional information on other potential influencing factors, such as provider characteristics (e.g., sex of the medical specialist and distance to the physical therapist), access to health care (e.g., copayment and transportation costs), or the perceived need for PT (19,20,22,40,48).

In conclusion, compared to the findings of previous studies, the utilization of PT by individuals with OA in the current study was high. Nevertheless, considering current OA management guidelines, which recommend PT as a first-line approach to conservative OA management, a higher utilization of PT is desirable. Clinicians and patients can be encouraged to utilize PT more often, especially male patients, individuals with a low household income, and patients with a high disease burden and functional limitations.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Jacobs, Callhoff, Albrecht, Postler, Lange, Goronzy, Günther, Hoffmann.

Acquisition of data. Jacobs, Callhoff, Saam, Hoffmann.

Analysis and interpretation of data. Jacobs, Callhoff, Hoffmann.

ADDITIONAL DISCLOSURES

Author Saam is an employee of Barmer.

REFERENCES

1. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003;81:646–56.
2. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates

from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014;73:1323–30.

3. World Health Organization. Chronic diseases and health promotion: chronic rheumatic conditions. 2016. URL: http://www.who.int/chp/topics/rheumatic/en/.
4. Bartholdy C, Christensen R, Kristensen LE, Gudbergsen H, Bliddal H, Overgaard A, et al. Association between weight loss and spontaneous changes in physical inactivity in overweight/obese individuals with knee osteoarthritis: an eight-week prospective cohort study. Arthritis Care Res (Hoboken) 2020;72:397–404.
5. Gersing AS, Schwaiger BJ, Nevitt MC, Zarnowski J, Joseph GB, Feuerriegel G, et al. Weight loss regimen in obese and overweight individuals is associated with reduced cartilage degeneration: 96-month data from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2019;27:863–70.
6. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Lancet 2019;393:1745–59.
7. Martel-Pelletier J, Barr AJ, Cicuttini FM, Conaghan PG, Cooper C, Goldring MB, et al. Osteoarthritis. Nat Rev Dis Prim 2016;2:16072.
8. Storheim K, Zwart JA. Musculoskeletal disorders and the Global Burden of Disease study. Ann Rheum Dis 2014;73:949–50.
9. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014;22:363–88.
10. Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis 2013;72:1125–35.
11. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: the Chronic Osteoarthritis Management Initiative of the U.S. Bone and Joint Initiative. Semin Arthritis Rheum 2014;43:701–12.
12. Block JA. OA guidelines: improving care or merely codifying practice? Nat Rev Rheumatol 2014;10:324–6.
13. Bannuru RR, Osani MC, Vaysbrot EE, Arden NK, Bennell K, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthritis Cartilage 2019;27:1578–89.
14. Roos EM, Juhl CB. Osteoarthritis 2012 year in review: rehabilitation and outcomes. Osteoarthritis Cartilage 2012;20:1477–83.
15. Fransen M, Mcconnell S, Harmer AR, van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee: a Cochrane systematic review. Br J Sports Med 2015;49:1554–7.
16. Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. Cochrane Database Syst Rev 2014;22:CD007912.
17. Tanaka R, Ozawa J, Kito N, Moriyama H. Efficacy of strengthening or aerobic exercise on pain relief in people with knee osteoarthritis: a systematic review and meta-analysis of randomized controlled trials. Clin Rehabil 2013;27:1059–71.
18. Uthman OA, van der Windt DA, Jordan JL, Dziedzic KS, Healey EL, Peat GM, et al. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. BMJ 2013;347:f5555.
19. Carter SK, Rizzo JA. Use of outpatient physical therapy services by people with musculoskeletal conditions background and purpose. Phys Ther 2007;87:497–512.
20. Yeh HJ, Chou YJ, Yang NP, Huang N. Receipt of physical therapy among osteoarthritis patients and its influencing factors. Arch Phys Med Rehabil 2015;96:1021–7.
21. Power JD, Cott CA, Badley EM, Hawker GA. Physical therapy services for older adults with at least moderately severe hip or knee arthritis in 2 Ontario counties. J Rheumatol 2005;32:123–9.
22. Iversen MD, Schwartz TA, von Heideken J, Callahan LF, Golightly YM, Goode A, et al. Sociodemographic and clinical correlates of physical therapy utilization in adults with symptomatic knee osteoarthritis. Phys Ther 2018;98:670–8.
23. Lange T, Luque Ramos A, Albrecht K, Günther KP, Jacobs H, Schmitt J, et al. Prescription frequency of physical therapy and analgesics before total hip and knee arthroplasty: an epidemiological analysis of routine health care data from Germany. Orthopade 2018;47:1015–23. In German.
24. Bawa HS, Weick JW, Dirschl DR. Gender disparities in osteoarthritis-related health care utilization before total knee arthroplasty. J Arthroplasty 2016;31:2115–8.e1.
25. Lange U, Berg W, Engel MJ, Fetaj S, Lichti G, Reuss-Borst M, et al. Evidenz zur physikalischen medizin in der rheumatologie. Aktuelle Rheumatol 2009;34:299–302.
26. Brechtel T, Kossack N, Grandt D. BARMER GEK heil- und hilfsmittelreport. URL: https://www.barmer.de/blob/77942/a6a48386ab57c7c566b58166a86d75bf/data/pdf-barmer-gek-heil-und-hilfsmittelreport-2016.pdf.
27. BARMER. Kerndaten der BARMER. 2018. URL: https://www.barmer.de/presse/infothek/daten-und-fakten/kerndaten-42164.
28. GKV-Spitzenverband. Bundeseinheitliches positionsnummernverzeichnis für heilmittel. 2020. URL: https://www.gkv-datenaustausch.de/media/dokumente/leistungserbringer_1/sonstige_leistungserbringer/positionsnummernverzeichnisse_archiv/Heilmittel_20200424.pdf.
29. Kassenärztliche Bundesvereinigung. Richtlinie der kassenärztlichen bundesvereinigung nach § 75 absatz 7 SGB V zur vergabe praxisnetznummern. 2015. URL: http://www.kbv.de/media/sp/Arztnummern_Richtlinie.pdf.
30. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care 1998;36:8–27.
31. Bellamy N. The WOMAC Knee and Hip Osteoarthritis Indices: development, validation, globalization and influence on the development of the AUSCAN Hand Osteoarthritis Indices. Clin Exp Rheumatol 2005;23:148–53.
32. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84:167–76.
33. Inagaki H, Ito K, Sakuma N, Sugiyama M, Okamura T, Awata S. Reliability and validity of the simplified Japanese version of the WHO-Five Well-being Index (S-WHO-5-J). Nihon Koshu Eisei Zasshi 2013;60:294–301. In Japanese.
34. Boonstra AM, Preuper HR, Balk GA, Stewart RE. Cut-off points for mild, moderate, and severe pain on the visual analogue scale for pain in patients with chronic musculoskeletal pain. Pain 2014;155:2545–50.
35. Rommel A, Kroll LE. Individual and regional determinants for physical therapy utilization in Germany: multilevel analysis of national survey data. Phys Ther 2017;97:512–23.
36. Hagglund KJ, Clark MJ, Hilton SA, Hewett JE. Access to healthcare services among persons with osteoarthritis and rheumatoid arthritis. Am J Phys Med Rehabil 2005;84:702–11.
37. Hawker GA, Wright JG, Coyte PC, Williams JI, Harvey B, Glazier R, et al. Differences between men and women in the rate of use of hip and knee arthroplasty. N Engl J Med 2000;342:1016–22.
38. McMurray RJ, Clarke OW, Barrasso JA, Clohan DB, Epps Jr CH, Glasson J, et al. Gender disparities in clinical decision making. JAMA 1991;266:559–62.
39. Jawahar R, Yang S, Eaton CB, McAlindon T, Lapane KL. Gender-specific correlates of complementary and alternative medicine use for knee osteoarthritis. J Womens Health (Larchmt) 2012;21:1091–9.
40. Freburger JK, Holmes GM. Physical therapy use by community-based older people. Phys Ther 2005;85:19–33.

41. Postler A, Luque Ramos A, Goronzy J, Günther KP, Lange T, Schmitt J, et al. Prevalence and treatment of hip and knee osteoarthritis in people aged 60 years or older in Germany: an analysis based on health insurance claims data. Clin Interv Aging 2018;13: 2339–49.
42. Jacobs H, Callhoff J, Hoffmann F, Zink A, Albrecht K. Non-drug treatment of rheumatoid arthritis: an analysis of claims data and a survey of insured persons (project PROCLAIR). Z Rheumatol 2019;78: 119–26.
43. Waltz M. The disease process and utilization of health services in rheumatoid arthritis: the relative contributions of various markers of disease severity in explaining consumption patterns. Arthritis Care Res 2000;13:74–88.
44. Ethgen O, Kahler KH, Kong SX, Reginster JY, Wolfe F. The effect of health related quality of life on reported use of health care resources in patients with osteoarthritis and rheumatoid arthritis: a longitudinal analysis. J Rheumatol 2002;29:1147–55.
45. Kopkow C, Lange T, Schmitt J, Petzold T. Physiotherapeutische versorgungssituation in Deutschland von 2004 bis 2014. Analyse des verordnungsverhaltens bei versicherten der gesetzlichen krankenversicherung. Das Gesundheitswes 2014;2016(79): 153–160.
46. Hoffmann F, Icks A. Unterschiede in der versichertenstruktur von krankenkassen und deren auswirkungen für die versorgungsforschung: ergebnisse des Bertelsmann-gesundheitsmonitors. Gesundheitswesen 2012;74:291–7.
47. Luque Ramos A, Hoffmann F. Differences in chronic back pain and joint disorders among health insurance funds: results of a cross-sectional study based on the data of the socioeconomic panel from 2013. Z Rheumatol 2017;76:238–44.
48. Nelson T, Fernandez JL, Livingston G, Knapp M, Katona C. Does diagnosis determine delivery? The Islington study of older people's needs and health care costs. Psychol Med 2004;34: 147–55.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 1023–1030
DOI 10.1002/acr.24227

American College of Rheumatology
Empowering Rheumatology Professionals

Structural Characteristics Associated With Radiographic Severity of First Metatarsophalangeal Joint Osteoarthritis

Andrew K. Buldt,[1] Shannon E. Munteanu,[1] Jamie J. Allan,[1] Jade M. Tan,[1] Maria Auhl,[1] Karl B. Landorf,[1] Edward Roddy,[2] and Hylton B. Menz[1]

Objective. To determine whether foot structure varies according to the presence and radiographic severity of first metatarsophalangeal (MTP) joint osteoarthritis (OA).

Methods. Weight-bearing dorsiplantar and lateral radiographs were obtained for the symptomatic foot of 185 participants (105 women, ages 22–85 years) with clinically diagnosed first MTP joint OA. A validated atlas was used to classify participants as having radiographic first MTP joint OA and to stratify into 3 categories of severity (none/mild, moderate, severe). Bone length and width and angular measures of the forefoot and medial arch were performed on radiographs, and differences between categories were compared using univariate general linear models, adjusting for confounders.

Results. A total of 150 participants were categorized as having radiographic first MTP joint OA, and participants were further stratified into none/mild (n = 35), moderate (n = 69), or severe (n = 81) OA categories. Participants with radiographically defined first MTP joint OA displayed a greater hallux abductus interphalangeal angle. Greater radiographic severity of first MTP joint OA was associated with a larger hallux abductus interphalangeal angle, a wider first metatarsal and proximal phalanx, and a smaller intermetatarsal angle. No differences in medial arch measurements were observed between the categories.

Conclusion. First ray alignment and morphology differed according to the presence and severity of first MTP joint OA. Prospective studies are required to determine whether the observed differences are a cause or consequence of OA.

INTRODUCTION

Osteoarthritis (OA) of the first metatarsophalangeal (MTP) joint affects 7.8% of the population age ≥50 years and is more prevalent in women and those who work in manual occupations (1). The clinical symptoms of first MTP joint OA include pain and stiffness in and around the joint, leading to significant reduction in quality of life and locomotor function, with 71% of people with the condition reporting disabling symptoms (1,2). Greater radiographic severity of first MTP joint OA is associated with a higher prevalence of pain and deformity and a lower range of joint motion, suggesting that it may be a progressive disorder (3). However, despite many risk factors being suggested, such as age, female sex, and trauma, the mechanisms responsible for the development and progression of first MTP joint OA are not well understood (4).

Variations in skeletal structure have been identified as an intrinsic risk factor for the development and progression of OA in a number of lower-extremity joints, including the knee and hip (5,6). These variations have been attributed to altered joint biomechanics, resulting in changes to the normal distribution of forces acting at the joint (7). For example, in individuals with medial compartment knee OA, varus alignment of the knee increases the knee adduction moment and alters joint compression forces within the medial compartment during gait (8,9). This change in joint biomechanics has been shown to be associated with disease severity and progression (10,11).

Although variations in skeletal structure of the foot are possibly an intrinsic risk factor for first MTP joint OA, the association between first MTP joint OA and the structure of the foot is unclear. Our previous systematic review found evidence that people with

Supported by the National Health and Medical Research Council (Australia) (grants 1105244 and 1049085). Dr. Menz's work was supported by a National Health and Medical Research Council fellowship (ID 1135995).

[1]Andrew K. Buldt, PhD, Shannon E. Munteanu, PhD, Jamie J. Allan, MPodMed, Jade M. Tan, MPodMed, Maria Auhl, BSc, Karl B. Landorf, PhD, Hylton B. Menz, PhD: La Trobe University, Melbourne, Victoria, Australia; [2]Edward Roddy, PhD: Primary Care Centre Versus Arthritis, School of Medicine, Keele University, Keele, Staffordshire, UK.

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

Email correspondence to Andrew K. Buldt, PhD, La Trobe Sport and Exercise Medicine Research Centre, La Trobe University, Melbourne, Victoria 3086, Australia. Email: a.buldt@latrobe.edu.au.

Submitted for publication August 23, 2019; accepted in revised form April 14, 2020.

SIGNIFICANCE & INNOVATIONS

- This is the first study to evaluate foot structure of individuals with first metatarsophalangeal joint osteoarthritis determined using a validated radiographic atlas.
- Weight-bearing dorsiplantar and lateral radiographs were obtained for the symptomatic foot of 185 participants with clinically diagnosed first metatarsophalangeal joint osteoarthritis.
- Participants with radiographically defined first metatarsophalangeal joint osteoarthritis displayed a greater hallux abductus interphalangeal angle. Increasing radiographic severity of first metatarsophalangeal joint osteoarthritis was associated with a larger hallux abductus interphalangeal angle, wider first metatarsal and proximal phalanx, and smaller intermetatarsal angle.
- First ray alignment and morphology differ according to the presence and severity of first metatarsophalangeal joint osteoarthritis.

first MTP joint OA exhibit a wider first metatarsal, wider proximal phalanx of the hallux, longer hallux, and more dorsiflexed first metatarsal compared to people without the condition (12). However, the studies included in the systematic review were limited in that they defined first MTP joint OA using only clinical symptoms or did not use a valid atlas to confirm the presence of radiographically defined first MTP joint OA. These issues make the interpretation of results from previous research difficult because the definition of first MTP joint OA varies between studies (12). The aims of this study were to use a foot-specific radiographic atlas to determine whether skeletal differences exist in people with and without first MTP joint OA, and whether skeletal variations are associated with first MTP joint OA severity.

SUBJECTS AND METHODS

Participants. The study sample consisted of 185 individuals who participated in 2 clinical trials of nonsurgical interventions for first MTP joint OA (13,14). All participants had a clinical diagnosis of first MTP joint OA and met the following inclusion criteria: 1) age ≥18 years, 2) report of having pain in the first MTP joint on most days for at least 12 weeks, 3) report of having pain rated at least 30 mm on a 100-mm visual analog scale, 4) description of pain on palpation of the dorsal aspect of the first MTP joint, 5) restricted dorsiflexion of the first MTP joint (<64° of dorsiflexion range of motion), and 6) ability to walk household distances (>50 meters) without the aid of a walker. Exclusion criteria included: 1) previous first MTP joint surgery, 2) current pregnancy, 3) significant first MTP joint deformity, including hallux valgus (defined as a score of 2 or 3 using the Manchester scale) (15), 4) the presence of any condition within the foot or ankle that could confound pain and functional assessments of the first MTP joint, or 5) the presence of inflammatory conditions such as gout or rheumatoid arthritis. The La Trobe University Human Ethics Committee approved the studies from which participants were drawn (HEC15128 and HEC18375). All radiographic procedures were performed according to the National Health and Medical Research Council of Australia National Statement on Ethical Conduct in Human Research (16).

Radiographic assessment. Weightbearing dorsiplantar and lateral radiographic projections were obtained for all participants while they were standing in a relaxed weightbearing position. If the participant had clinically defined first MTP joint OA in both feet, radiographs were taken on the most symptomatic foot. All radiographs were taken by the same medical imaging group using a Shimadzu UD150LRII 50 kW/30 kHz Generator and 0.6/1.2 P18DE-80S high-speed radiograph tube from a ceiling-suspended tube mount. AGFA MD40 CR digital phosphor plates in a 24-cm × 30-cm cassette were also used. For dorsiplantar projections, the radiograph tube was positioned at an angle of 15° cephalad and centered at the base of the third metatarsal. For the lateral projection, the radiograph tube was positioned at an angle of 90° and centered at the base of the third metatarsal. The film focal distance was 100 cm for both projections.

Radiographs were assessed to confirm the presence and severity of radiographically defined first MTP joint OA. The La Trobe University Radiographic Atlas of Foot Osteoarthritis was used to assess radiographs (17). The atlas has moderate to excellent intrarater reliability and moderate to excellent interrater reliability and is used to determine the severity of osteophytes and joint space narrowing at the first MTP joint (17). The presence of osteophytes was graded as being either absent (score = 0), small (score = 1), moderate (score = 2), or severe (score = 3). The presence of joint space narrowing was graded as being either none (score = 0), definite (score = 1), moderate (score = 2), or severe (score = 3). All assessments were conducted by 2 experienced raters (SEM and HBM), who contributed to the development of the atlas.

Participants were defined as having radiographic first MTP joint OA if they recorded a score of ≥2 for either osteophytes or joint space narrowing in either projection (17). Participants were also assigned to 1 of 3 radiographic severity categories: none/mild OA (defined as 1 score of at least 1 and no score of ≥2 for either osteophytes or joint space narrowing from either the dorsiplantar or lateral radiographs), moderate OA (defined as 1 score of at least 2 and no score of 3), or severe OA (defined as 1 score of 3).

Radiographic measurements of foot structure. The selection of radiographic measurements was based on the need to comprehensively characterize both the structure and architecture of the foot in individuals with first MTP joint OA using measures that had adequate reliability and validity. For all participants, the following variables were measured from dorsiplantar radiographs: first metatarsal length, first metatarsal width, proximal

phalanx length, proximal phalanx width, distal phalanx length, total hallux length, intermetatarsal angle, hallux abductus angle, hallux abductus interphalangeal angle, metatarsus adductus angle (simplified technique), and metatarsal protrusion distance (18–21). For the lateral radiographs, the following variables were measured: calcaneal-first metatarsal angle, first metatarsal declination angle, lateral intermetatarsal angle, dorsal proximal metatarsal angle, dorsal proximal hallux angle, dorsal proximal phalangeal angle, plantar distal metatarsal angle, and plantar distal hallux angle (20,22–24). Measurements are shown in Figures 1 and 2, and detailed explanations are provided in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24227/abstract. All measurements were made on digital radiographs in the same manner for each participant by the same examiner (AKB). Test–retest (intrarater) reliability was evaluated by repeating all radiographic measurements on 2 separate occasions, 2 weeks apart.

Statistical analysis. All analyses were performed using IBM SPSS Statistics, release 24 for Windows. Reliability was calculated using intraclass correlation coefficients (ICCs) (model 3,1) with 95% confidence intervals (95% CIs) (25). To determine whether there were any significant differences in radiographic measurements between participants with and without radiographically defined first MTP joint OA, or within participants with different first MTP joint OA radiographic severity categories, univariate general linear models were calculated for all variables. To determine appropriate covariates for the models, a series of independent samples *t*-tests or chi-square tests were conducted for the comparison of cases and non-cases. Covariates were identified where there were significant differences between severity categories (*P* values less than 0.05 were considered significant). General linear models with the entry of covariates and least significant difference adjustment were conducted to determine differences in structural variables between cases and non-cases, and between severity categories of first MTP joint OA. For all analyses, adjusted mean differences were calculated, with *P* values less than 0.05 considered significant. Effect sizes (Cohen's *d*) were calculated for all significant structural differences to allow comparison of magnitude of differences across measures.

RESULTS

Reliability. Mean ± SDs for tests and re-tests for all radiographic measures, along with ICCs and 95% CIs, are shown in Supplementary Tables 2–4. All measures displayed good-to-excellent intrarater reliability, with ICCs ranging between 0.82 and 0.98.

Participant characteristics. Participant characteristics for participants with and without radiographic first MTP joint OA are shown in Table 1. Of the 185 participants, 150 (81.1%) had radiographic first MTP joint OA. Participants with

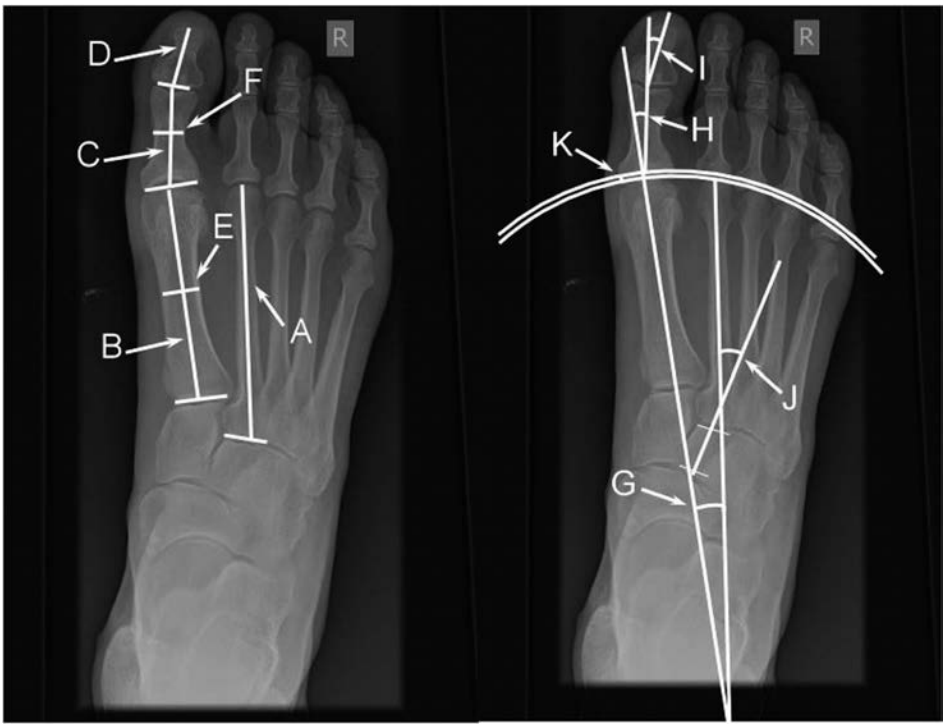

Figure 1. Dorsiplantar radiographic measurement techniques: bone length and width measurements. **A**, Second metatarsal length; **B**, First metatarsal length; **C**, Proximal phalanx length; **D**, Distal phalanx length; **E**, First metatarsal width; **F**, Proximal phalanx width. Dorsiplantar radiographic measurement techniques: angle measures; **G**, Intermetatarsal angle; **H**, Hallux abductus angle; **I**, Hallux abductus interphalangeal angle; **J**, Simplified metatarsus adductus angle; **K**, Metatarsal protrusion distance.

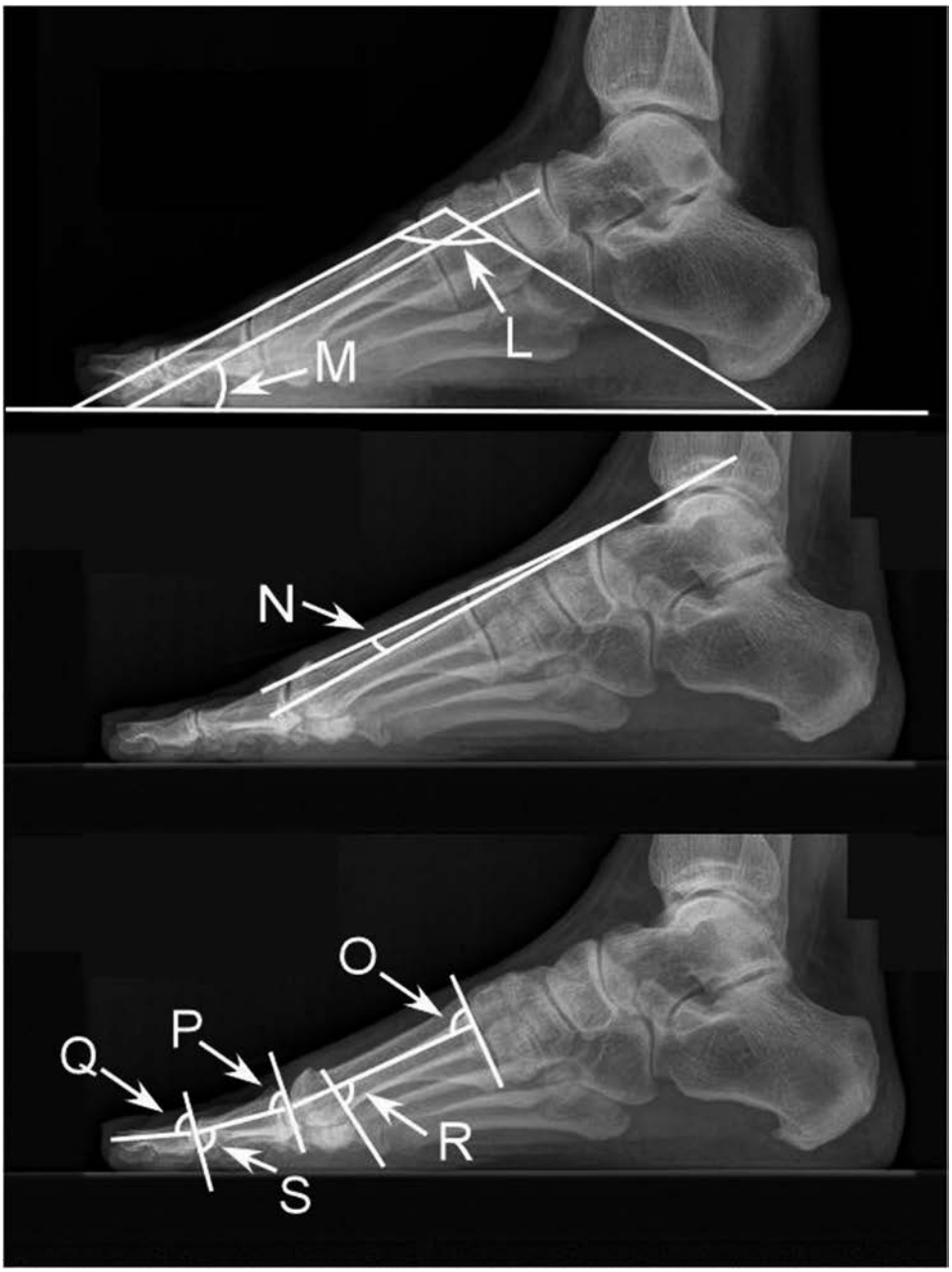

Figure 2. Lateral radiographic measurement techniques. **L**, Calcaneal-first metatarsal angle; **M**, First metatarsal declination angle; **N**, Lateral intermetatarsal angle; **O**, Dorsal proximal metatarsal angle; **P**, Dorsal proximal hallux angle; **Q**, Dorsal proximal phalangeal angle; **R**, Plantar distal metatarsal angle; **S**, Plantar distal hallux angle.

radiographic first MTP joint OA exhibited significantly greater weight (mean difference 5.8 kg [95% CI 1.0, 10.6]), body mass index (BMI) (mean difference 1.9 kg/m^2 [95% CI 0.2, 3.5]), and duration of symptoms (mean difference 35.9 months [95% CI 7.3, 64.5]) compared to those without radiographic OA. For the general linear models, BMI and duration of symptoms were considered to be confounders and were entered as covariates. Although weight was significantly different between cases and non-cases, it was not included in addition to BMI as a covariate to avoid possible overadjustment because the 2 variables were strongly correlated (r = 0.799, $P < 0.001$).

Participant characteristics for the comparison between radiographic severity categories are shown in Table 2. There were 35 participants (18.9%) in the none/mild category, 69 (37.2%) in the moderate category, and 81 (43.8%) in the severe category. Participants in the severe category were significantly older than in categories for none/mild (mean difference 4.9 years [95% CI 0.7, 9.1]) and moderate (mean difference 3.8 years [95% CI 0.5, 7.2])

Table 1. Characteristics of participants with (case) and without (non-case) radiographically defined first MTP joint osteoarthritis*

Characteristic	Non-case (n = 35)	Case (n = 150)	*P*
Age, years	55.0 ± 13.4	58.1 ± 9.9	0.12
Female, no. (%)	22 (62.9)	85 (56.7)	0.45
Height, cm	166.7 ± 8.5	166.9 ± 8.6	0.68
Weight, kg	74.2 ± 13.2	80.1 ± 13.1	0.02
Body mass index, kg/m^2	26.8 ± 4.6	28.7 ± 4.5	0.03
Self-reported duration of symptoms, months	35.8 ± 42.3	72.9 ± 83.2	0.01

* Values are the mean ± SD unless indicated otherwise. MTP = metatarsophalangeal.

and exhibited significantly greater weight compared to both none/mild (mean difference 9.1 kg [95% CI 4.1, 14.2]) and moderate (mean difference 7.0 kg [95% CI 2.9, 11.1]). The severe category also exhibited significantly greater BMI compared to the none/mild category (mean difference 2.5 kg/m^2 [95% CI 0.7, 4.3]). Finally, the severe category exhibited significantly greater self-reported duration of symptoms compared to the categories for none/mild (mean difference 51.9 months [95% CI 21.5, 82.3]) and moderate (mean difference 33.2 months [95% CI 8.7, 57.8]).

Structural differences between participants with and without radiographic first MTP joint OA. Structural characteristics in case and non-case categories are shown in Table 3. The case category exhibited greater hallux abductus interphalangeal angle compared to the non-case category (mean difference 4.1° [95% CI 2.0, 6.3], d = 0.77). There were no other statistically significant differences in measures of structure between those with and without radiographic first MTP joint OA.

Structural differences according to radiographic severity in those with radiographic first MTP joint OA. Structural characteristics according to radiographic severity are shown in Table 4. The severe radiographic OA category exhibited a significantly wider first metatarsal compared to the moderate category (mean difference –1.0% [95% CI –1.9, –0.1], d = 0.54), a wider proximal phalanx compared to the moderate category (mean difference –2.2% [95% CI –3.7, –0.6], d = 0.51), a smaller intermetatarsal angle compared to both the none/mild (mean difference 1.1° [95% CI 0.2, 2.1], d = 0.37) and moderate (mean difference 0.8° [95% CI 0.1, 1.6], d = 0.36) categories, and a significantly greater hallux abductus interphalangeal angle compared to the none/mild category (mean difference 4.3° [95% CI 1.9, 6.7], d = 0.79). The moderate category displayed a greater hallux abductus interphalangeal angle compared to the none/mild category (mean difference 3.9° [95% CI 1.6, 6.2], d = 0.77).

DISCUSSION

The objective of this study was to determine whether skeletal foot structure varies according to the presence and radiographic severity of first MTP joint OA. To the best of our knowledge, this is the first study to evaluate foot structure of individuals with first MTP joint OA determined using a validated radiographic atlas. A comprehensive suite of radiographic measurements was used, and we found that some radiographic measurements related to first ray alignment and morphology differ according to the radiographic severity of first MTP joint OA.

Among all structural variables, the hallux abductus interphalangeal angle was the only measure that was significantly different between those with and without first MTP joint OA and between severity categories. The magnitude of the differences, determined by effect sizes, was also largest for this measurement. These findings indicate that there is greater lateral deviation of the distal phalanx relative to the proximal phalanx in individuals with radiographically defined first MTP joint OA. Furthermore, there was evidence of a dose-response relationship as the degree of distal phalanx deviation increased with increasing severity of radiographic first MTP joint OA.

Two previous studies found no difference in hallux abductus interphalangeal angle between cases and controls (20,26). However, in those studies, the inclusion criteria used to recruit participants were either clinical symptoms or first MTP joint range-of-motion testing. In addition, 1 study only recruited participants with early signs of first MTP joint OA (26). Therefore, because our study recruited participants who exhibited a range of radiographic severities, the findings suggest that a temporal relationship may exist between longer duration of first MTP joint OA and lateral deviation of the distal phalanx.

The mechanism that leads to greater hallux abductus interphalangeal angle may involve alterations in forces acting on the

Table 2. Participant characteristics according to radiographic severity of first MTP joint osteoarthritis*

Characteristic	None/mild (n = 35)	Moderate (n = 69)	Severe (n = 81)	*P*
Age, years	55.0 ± 13.4	56.1 ± 10.8	59.9 ± 8.9	0.02†
Female, no. (%)	22 (62.9)	44 (63.8)	39 (48.1)	0.10
Height, cm	166.7 ± 8.5	165.2 ± 8.7	168.4 ± 8.3	0.57
Weight, kg	74.2 ± 13.2	76.3 ± 11.3	83.4 ± 13.7	<0.01†
Body mass index, kg/m^2	26.8 ± 4.5	28.0 ± 4.6	29.3 ± 4.6	0.02‡
Duration, months	35.8 ± 42.3	56.6 ± 66.6	86.5 ± 93.8	<0.01†

* Values are the mean ± SD unless indicated otherwise. MTP = metatarsophalangeal.
† Significant difference between none/mild and severe, and between moderate and severe.
‡ Significant difference between none/mild and severe.

Table 3. Comparison of structural characteristics between participants with (case) and without (non-case) radiographically defined first MTP joint osteoarthritis*

Characteristic	Non-case (n = 35)	Case (n = 150)	*P*	Adjusted mean difference (95% CI)
Dorsiplantar projection				
First metatarsal length†	85.4 ± 3.6	84.6 ± 3.1	0.36	−0.5 (−1.7, 0.7)
Proximal phalanx length†	43.6 ± 3.3	43.6 ± 4.4	0.58	0.3 (−1.2, 1.9)
Distal phalanx length†	30.9 ± 2.9	31.4 ± 3.2	0.40	0.3 (−0.8, 1.5)
Hallux length†	74.6 ± 5.2	74.9 ± 6.5	0.90	0.1 (−2.4, 2.2)
First metatarsal width‡	19.8 ± 2.5	20.1 ± 2.6	0.72	0.3 (−0.9, 0.9)
Proximal phalanx width‡	38.5 ± 4.9	38.8 ± 4.7	0.96	−0.1 (−1.8, 1.6)
Intermetatarsal angle, °	10.1 ± 2.7	9.6 ± 2.2	0.90	−0.6 (−1.5, 0.2)
Hallux abductus angle, °	12.4 ± 5.6	10.9 ± 4.7	0.11	−1.4 (−3.3, 0.4)
Hallux abductus interphalangeal angle, °	11.2 ± 5.5	15.6 ± 5.8	<0.01	4.1 (2.0, 6.3)§
Metatarsal protrusion distance, mm	2.2 ± 3.6	1.7 ± 3.6	0.48	−0.6 (−1.9, 0.8)
Lateral projection				
Metatarsus adductus angle, °	22.5 ± 5.4	22.9 ± 5.5	0.74	0.2 (−1.8, 2.3)
Calcaneal, first metatarsal angle, °	131.1 ± 8.9	132.3 ± 7.2	0.75	0.3 (−2.6, 3.2)
First metatarsal declination angle, °	24.2 ± 3.6	23.1 ± 3.2	0.19	−0.7 (−2.0, 0.5)
Lateral intermetatarsal angle, °	1.9 ± 1.4	1.9 ± 1.2	0.82	−0.1 (−0.5, 0.4)
Dorsal proximal metatarsal angle, °	89.9 ± 2.3	89.1 ± 2.6	0.19	0.6 (−0.3, 1.6)
Dorsal proximal hallux angle, °	84.8 ± 4.0	83.8 ± 5.8	0.29	1.1 (−0.9, 3.2)
Dorsal proximal phalangeal angle, °	76.3 ± 4.4	77.6 ± 5.6	0.15	−1.5 (−3.6, 0.6)
Plantar distal metatarsal angle, °	82.1 ± 4.9	81.4 ± 5.6	0.89	1.0 (−1.9, 2.2)
Plantar distal hallux angle, °	89.6 ± 9.8	90.5 ± 7.9	0.35	−1.7 (−5.2, 1.9)

* Values are the mean ± SD unless indicated otherwise. Body mass index and duration of symptoms entered as covariates in general linear model. 95% CI = 95% confidence interval; MTP = metatarsophalangeal.
† Expressed as a percentage of length of the second metatarsal.
‡ Expressed as a percentage of length of the corresponding bone.
§ Effect size (*d*) = 0.77 showing significant difference.

interphalangeal joint of the hallux when walking. This angular difference is supported by biomechanical research conducted on individuals with and without first MTP joint OA that found both greater force on the hallux when walking and greater lateral deviation of the center of pressure in those with OA (27,28). These findings imply that a greater deviating force is placed on the hallux in people with first MTP joint OA compared to people without the condition. However, more research is needed to understand the long-term effects that biomechanical variations related to first MTP joint OA have on adjacent joints of the foot.

We found that the first metatarsal and proximal phalanx were significantly wider in individuals with severe first MTP joint OA compared to individuals with moderate first MTP joint OA. This finding is consistent with a previous study that found a significantly wider first metatarsal and proximal phalanx in cases of first MTP joint OA compared to asymptomatic controls (21). However, our findings are novel in that no difference was found between individuals with and without radiographic first MTP joint OA. There are 2 possible explanations for these findings. First, a wider first metatarsal and proximal phalanx may provide a relatively square (as opposed to round) joint surface that causes uneven and increased joint compression, leading to the initial development and progression of the condition over time (21). Second, bony remodeling may occur in response to altered loading in individuals with first MTP joint OA, resulting in increased width of the first metatarsal and proximal phalanx.

Individuals with severe first MTP joint OA also exhibited a smaller angle between the first and second metatarsals compared to both the none/mild and moderate categories, indicating a less medially deviated first metatarsal relative to the lateral forefoot. Studies of normal foot mechanics indicate that the first metatarsal moves in a direction of adduction relative to bones of the midfoot, allowing for abduction of the hallux during the propulsive phase of gait (29). Our findings suggest that in people with first MTP joint OA, the first metatarsal does not move into adduction to adequately facilitate normal function of the first MTP joint. Such a mechanism may lead to increased joint compression in the first MTP joint. However, this observation is also possibly a consequence of first MTP joint pathology.

No significant differences were found for any angular measurements from lateral projections that characterize foot posture. This result suggests that sagittal plane measures of the medial longitudinal arch are not associated with first MTP joint OA. In terms of previous research that investigated foot medial arch shape characteristics in people with first MTP joint OA, our findings differ from Mahiquez et al, who found that individuals with a rearfoot valgus angle of 5°, indicative of a flatter foot, were 23% more likely to develop first MTP joint OA (30). However, the findings of this prospective study used frontal plane heel position as an indicator of medial arch shape characteristics, whereas our study used a suite of lateral radiographic angular measurements. Further prospective work should use both sagittal plane and

Table 4. Comparison for radiographic measures according to radiographic severity of first MTP joint osteoarthritis*

	None/mild (n = 35)	Moderate (n = 69)	Severe (n = 81)	*P*	Between-group adjusted mean difference (95% CI)		
					None/mild vs. moderate	None/mild vs. severe	Moderate vs. severe
Dorsiplantar projection							
First metatarsal length†	85.4 ± 3.6	84.8 ± 3.1	84.4 ± 3.1	0.74	0.4 (–0.8, 1.7)	0.5 (–0.8, 1.8)	0.1 (–1.0, 1.1)
Proximal phalanx length†	43.6 ± 3.3	43.6 ± 5.7	43.7 ± 2.9	0.79	–0.2 (–1.9, 1.5)	–0.6 (–2.2, 1.2)	–0.3 (–1.7, 1.0)
Distal phalanx length†	30.9 ± 2.9	31.3 ± 3.3	31.6 ± 3.1	0.82	–0.3 (1.6, 0.9)	–0.5 (–1.8, 0.8)	–0.1 (–1.1, 0.9)
Hallux length†	74.6 ± 5.2	74.5 ± 7.7	75.3 ± 5.2	0.89	0.2 (–2.3, 2.8)	–0.3 (–2.9, 2.3)	–0.5 (–2.6, 1.6)
First metatarsal width‡	19.8 ± 2.6	19.5 ± 2.3	20.6 ± 2.8	0.07	0.3 (–0.7, 1.4)	–0.7 (–1.8, 0.4)	–1.0 (–1.9, –0.1)§
Proximal phalanx width‡	38.5 ± 4.9	37.6 ± 3.6	39.9 ± 5.3	0.03	0.9 (–1.0, 2.8)	–1.3 (–3.2, 0.7)	–2.2 (–3.7, –0.6)§
Intermetatarsal angle, °	10.1 ± 2.7	10.0 ± 2.3	9.2 ± 2.1	0.03	0.3 (–0.6, 1.2)	1.1 (0.2, 2.1)§	0.8 (0.1, 1.6)§
Hallux abductus angle, °	12.4 ± 5.6	10.7 ± 4.7	11.2 ± 4.8	0.27	1.6 (–0.4, 3.7)	1.4 (–0.7, 3.4)	–0.3 (–1.9, 1.4)
Hallux abductus interphalangeal angle, °	11.2 ± 5.5	15.3 ± 5.1	15.9 ± 6.3	<0.01	–3.9 (–6.2, –1.6)§	–4.3 (–6.7, –1.9)§	–0.3 (–2.2, 1.6)
Metatarsus adductus angle, °	22.5 ± 5.4	22.4 ± 4.8	23.3 ± 6.0	0.63	0.1 (–2.1, 2.4)	–0.7 (–3.0, 1.6)	–0.9 (–2.7, 1.0)
Metatarsal protrusion distance, mm	2.2 ± 3.6	1.7 ± 3.4	1.8 ± 3.7	0.78	0.5 (–0.9, 2.1)	0.3 (–1.2, 1.8)	–0.2 (–1.4, 0.9)
Lateral projection							
Calcaneal, first metatarsal angle, °	131.1 ± 8.9	131.2 ± 7.7	133.3 ± 6.7	0.43	0.3 (–2.8, 3.4)	–1.1 (–4.3, 2.1)	–1.4 (–3.9, 1.1)
First metatarsal declination angle, °	24.2 ± 3.6	23.7 ± 3.0	22.6 ± 3.3	0.12	0.3 (–1.0, 1.6)	1.3 (–0.1, 2.6)	0.9 (–0.1, 2.0)
Lateral intermetatarsal angle, °	1.8 ± 1.4	1.7 ± 1.1	2.1 ± 1.2	0.33	0.2 (–0.3, 0.7)	–0.1 (–0.6, 0.4)	–0.3 (–0.7, 0.1)
Dorsal proximal metatarsal angle, °	89.9 ± 2.3	89.1 ± 2.7	89.1 ± 2.6	0.40	0.7 (–0.4, 1.8)	0.6 (–0.5, 1.7)	–0.1 (–1.0, 0.7)
Dorsal proximal hallux angle, °	84.8 ± 4.0	83.8 ± 5.4	84.0 ± 6.1	0.59	1.1 (–1.1, 3.4)	1.1 (–1.2, 3.4)	–0.1 (–1.9, 1.8)
Dorsal proximal phalangeal angle, °	76.3 ± 4.4	77.9 ± 5.3	77.1 ± 5.8	0.32	–1.6 (–3.8, 0.5)	–0.7 (–2.9, 1.5)	0.9 (–0.9, 2.7)
Plantar distal metatarsal angle, °	82.1 ± 4.9	82.1 ± 4.9	80.8 ± 5.5	0.79	–0.4 (–2.6, 1.8)	0.2 (–2.0, 2.5)	0.6 (–1.2, 2.4)
Plantar distal hallux angle, °	90.5 ± 7.9	90.6 ± 9.5	88.7 ± 10.2	0.19	0.5 (–3.4, 4.4)	3.0 (–0.9, 7.0)	2.5 (–0.6, 5.7)

* Values are the mean ± SD unless indicated otherwise. Age, body mass index, and self-reported duration of symptoms entered as covariates in general linear model. 95% CI = 95% confidence interval; MTP = metatarsophalangeal.
† Expressed as a percentage of length of the second metatarsal.
‡ Expressed as a percentage of length of the corresponding bone.
§ Significant adjusted mean difference and effect size (*d*); for none/mild vs. moderate: *d* = 0.77 for hallux abductus interphalangeal angle; for none/mild vs. severe: *d* = 0.37 intermetatarsal angle and *d* = 0.79 for hallux abductus interphalangeal angle; for moderate vs. severe: *d* = 0.54 for first metatarsal width, *d* = 0.51 for proximal phalanx width, and *d* = 0.36 for intermetatarsal angle.

frontal plane measures of foot posture to provide further insights into the association between medial arch shape characteristics and the development of first MTP joint OA.

We expected the plantar distal hallux angle, indicative of distal phalanx dorsiflexion, to be greater in severe first MTP joint OA as a compensatory response to the lack of dorsiflexion range of motion available in the first MTP joint. Such a finding was reported in a study whereby individuals with limited first MTP joint range of motion (<55°) displayed significantly greater dorsiflexion of the hallux interphalangeal joint compared to controls with normal first MTP joint range of motion (31). However, no such significant difference was found. Rather, a comparatively greater variance in plantar distal hallux angle was found compared to other angular measures, particularly among individuals with severe first MTP joint OA. The greater variance in plantar distal hallux angle suggests that while some individuals displayed a dorsiflexed distal phalanx of the hallux, others displayed a plantarflexed hallux, similar to that observed in a hammertoe deformity.

Strengths of this study include the use of a validated atlas for first MTP joint OA and the analysis of a comprehensive suite of radiographic measurements of foot structure. However, several limitations also need to be considered. First, because this was a cross-sectional study, temporal relationships cannot be inferred. Long-term prospective studies are needed to determine whether structural differences identified in this study influence the progression of first MTP joint OA. Second, although structural factors have been identified as important factors that may contribute to the progression of first MTP joint OA, there are likely other factors, such as previous trauma, work/occupation, biomechanics, or genetics that could contribute to its development and progression (32). Third, the investigator taking radiographic measurements was not blinded to the pathology, a factor that is inherently difficult to achieve, introducing the risk of measurement bias. Fourth, we were limited to radiographic measures obtained from dorsiplantar and lateral radiographic views. Further studies could include measures from other views, such as frontal plane calcaneal measures from frontal plane views. Finally, all participants were symptomatic as they were recruited for clinical trials. Therefore, further study is required to understand the relationships between structure, radiographic severity, and the development of symptoms.

The presence and severity of radiographic first MTP joint OA is associated with larger hallux abductus interphalangeal angle, a wider first metatarsal and proximal phalanx, and a smaller intermetatarsal angle. These findings suggest that foot structure may be involved in the development and progression of first MTP joint OA. However, long-term prospective studies are required to further understand the role of these factors in the development of this condition.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Buldt, Munteanu, Landorf, Roddy, Menz.
Acquisition of data. Buldt, Munteanu, Allan, Tan, Auhl, Landorf, Menz.
Analysis and interpretation of data. Buldt, Munteanu, Landorf, Roddy, Menz.

REFERENCES

1. Roddy E, Thomas MJ, Marshall M, Rathod T, Myers HB, Menz HB, et al. The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: cross-sectional findings from the Clinical Assessment Study of the Foot. Ann Rheum Dis 2015;74:156–63.
2. Bergin SM, Munteanu SE, Zammit GV, Nikolopoulos N, Menz HB. Impact of first metatarsophalangeal joint osteoarthritis on health-related quality of life. Arthritis Care Res (Hoboken) 2012;64:1691–8.
3. Menz HB, Roddy E, Marshall M, Thomas M, Rathod T, Myers H, et al. Demographic and clinical factors associated with radiographic severity of first metatarsophalangeal joint osteoarthritis: cross-sectional findings from the Clinical Assessment Study of the Foot. Osteoarthritis Cartilage 2015;23:77–82.
4. Zammit GV, Menz HB, Munteanu SE, Landorf KB, Gilheany MF. Interventions for treating osteoarthritis of the big toe joint. Cochrane Database Syst Rev 2010;9:CD007809.
5. Nicholls AS, Kiran A, Pollard TC, Hart DJ, Arden CP, Spector T, et al. The association between hip morphology parameters and nineteen-year risk of end-stage osteoarthritis of the hip: a nested case–control study. Arthritis Rheum 2011;63:3392–400.
6. Haverkamp DJ, Schiphof D, Bierma-Zeinstra SM, Weinans H, Waarsing JH. Variation in joint shape of osteoarthritic knees. Arthritis Rheum 2011;63:3401–7.
7. Martel-Pelletier J, Di Battista J, Lajeunesse D. Biochemical factors in joint articular tissue degradation in osteoarthritis. In: Osteoarthritis. NY: Springer; 1999: p. 156–87.
8. Cho Y, Ko Y, Lee W. Relationships among foot position, lower limb alignment, and knee adduction moment in patients with degenerative knee osteoarthritis. J Phys Ther Sci 2015;27:265–8.
9. Sharma L, Hurwitz DE, Thonar EJ, Sum JA, Lenz ME, Dunlop DD, et al. Knee adduction moment, serum hyaluronan level, and disease severity in medial tibiofemoral osteoarthritis. Arthritis Rheum 1998;41:1233–40.
10. Sharma L, Chmiel JS, Almagor O, Felson D, Guermazi A, Roemer F, et al. The role of varus and valgus alignment in the initial development of knee cartilage damage by MRI: the MOST study. Ann Rheum Dis 2013;72:235–40.
11. Miyazaki T, Wada M, Kawahara H, Sato M, Baba H, Shimada S. Dynamic load at baseline can predict radiographic disease progression in medial compartment knee osteoarthritis. Ann Rheum Dis 2002;61:617–22.
12. Zammit GV, Menz HB, Munteanu SE. Structural factors associated with hallux limitus/rigidus: a systematic review of case control studies. J Orthop Sports Phys Ther 2009;39:733–42.
13. Menz HB, Levinger P, Tan JM, Auhl M, Roddy E, Munteanu SE. Rocker-sole footwear versus prefabricated foot orthoses for the treatment of pain associated with first metatarsophalangeal joint osteoarthritis: study protocol for a randomised trial. BMC Musculoskeletal Disord 2014;15:86.
14. Munteanu SE, Landorf KB, McClelland JA, Roddy E, Cicuttini FM, Shiell A, et al. Shoe-stiffening inserts for first metatarsophalangeal joint osteoarthritis (the SIMPLE trial): study protocol for a randomised controlled trial. Trials 2017;18:198.
15. Garrow AP, Papageorgiou A, Silman AJ, Thomas E, Jayson MI, Macfarlane GJ. The grading of hallux valgus: the Manchester Scale. J Am Podiatr Med Assoc 2001;91:74–8.
16. National Health and Medical Research Council. National statement on ethical conduct in human research. Canberra 2007 (updated 2018). URL: https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018.
17. Menz HB, Munteanu SE, Landorf KB, Zammit GV, Cicuttini FM. Radiographic classification of osteoarthritis in commonly affected joints of the foot. Osteoarthritis Cartilage 2007;15:1333–8.
18. Hardy R, Clapham J. Observations on hallux valgus. J Bone Joint Surg Br 1951;33:376–91.
19. Kaschak T, Laine W. Surgical radiology. Clin Podiatr Med Surg 1988;5:797–829.
20. Bryant A, Tinley P, Singer K. A comparison of radiographic measurements in normal, hallux valgus, and hallux limitus feet. J Foot Ankle Surg 2000;39:39–43.
21. Munuera PV, Dominguez G, Castillo JM. Radiographic study of the size of the first metatarso-digital segment in feet with incipient hallux limitus. J Am Podiatr Med Assoc 2007;97:460–8.
22. Bryant A, Mahoney B, Tinley P. Lateral intermetatarsal angle: a useful measurement of metatarsus primus elevatus? J Am Podiatr Med Assoc 2001;91:251–4.
23. Menz HB, Munteanu SE. Validity of 3 clinical techniques for the measurement of static foot posture in older people. J Orthop Sports Phys Ther 2006;36:479–86.
24. Lamm BM, Stasko PA, Gesheff MG, Bhave A. Normal foot and ankle radiographic angles, measurements, and reference points. J Foot Ankle Surg 2016;55:991–8.
25. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327:307–10.
26. Castillo-Lopez JM, Ramos-Ortega J, Reina-Bueno M, Domínguez-Maldonado G, Palomo-Toucedo IC, Munuera PV. Hallux abductus interphalangeus in normal feet, early-stage hallux limitus, and hallux valgus. J Am Podiatr Med Assoc 2014;104:169–73.
27. Zammit GV, Menz HB, Munteanu SE, Landorf KB. Plantar pressure distribution in older people with osteoarthritis of the first metatarsophalangeal joint (hallux limitus/rigidus). J Orthop Res 2008;26:1665–9.
28. Menz HB, Auhl M, Tan JM, Buldt AK, Munteanu SE. Centre of pressure characteristics during walking in individuals with and without first metatarsophalangeal joint osteoarthritis. Gait Posture 2018;63:91–6.
29. Glasoe WM, Yack HJ, Saltzman CL. Anatomy and biomechanics of the first ray. Phys Ther 1999;79:854–9.
30. Mahiquez MY, Wilder FV, Stephens HM. Positive hindfoot valgus and osteoarthritis of the first metatarsophalangeal joint. Foot Ankle Int 2006;27:1055–9.
31. Munuera PV, Trujillo P, Güiza I. Hallux interphalangeal joint range of motion in feet with and without limited first metatarsophalangeal joint dorsiflexion. J Am Podiatr Med Assoc 2012;102:47–53.
32. Roddy E, Menz HB. Foot osteoarthritis: latest evidence and developments. Ther Adv Musculoskelet Dis 2018;10:91–103.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 1031–1037
DOI 10.1002/acr.24193

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Changes in Medial Meniscal Three-Dimensional Position and Morphology As Predictors of Knee Replacement in Rapidly Progressing Knee Osteoarthritis: Data From the Osteoarthritis Initiative

Melanie Roth,[1] Katja Emmanuel,[1] Wolfgang Wirth,[2] C. Kent Kwoh,[3] David J. Hunter,[4] Michael J. Hannon,[5] and Felix Eckstein[2]

Objective. To assess whether quantitative changes in the meniscus predict progression from early knee osteoarthritis (OA) to knee replacement (KR).

Methods. A nested case–control study was conducted among Osteoarthritis Initiative participants: all 35 case knees with baseline Kellgren/Lawrence (K/L) grade ≤2 that had KR between 36 and 60 months were matched 1:1 by age, sex, and baseline K/L grade to 35 control knees without subsequent KR. Quantitative 3-dimensional medial meniscus position and morphologic measures were determined from magnetic resonance imaging at the visit just before KR and 2 years before. Paired *t*-tests and case–control odds ratios (ORs, standardized per SD of change in controls) were used to compare changes between groups.

Results. Cases (52% women, age 65 ± 7 years, body mass index [BMI] 30 ± 4 kg/m^2, K/L grades 0/1/2: 5/8/22 participants, respectively) and controls (52% women, age 64 ± 7 years, BMI 30 ± 5 kg/m^2, K/L grades 0/1/2: 9/4/22 participants, respectively) were similar. Compared to control knees, KR case knees displayed longitudinal changes, specifically, a decrease in tibial plateau coverage, an increase in meniscal extrusion, and a decrease in meniscal width. The odds for KR increased with greater reduction in the percentage of tibial plateau coverage (OR 2.28 [95% CI confidence interval (95% CI) 1.43, 3.64]), a greater increase in maximal extrusion (OR 1.40 [95% CI 1.12, 1.75]), and a greater reduction of mean meniscal width (OR 2.01 [95% CI 1.23, 3.26]). The odds for KR increased with medial compartment cartilage thickness loss (OR 2.86 [95% CI 1.51, 5.41]) for comparison.

Conclusion. Quantitative measures of meniscal position and morphology are associated with subsequent KR in knees with rapidly progressing knee OA. These findings show that structural changes of the meniscus are related to an important clinical and economic outcome of knee OA.

ClinicalTrials.gov identifier: NCT00080171.

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

[1]Melanie Roth, PhD, Katja Emmanuel, MD: Paracelsus Medical University Salzburg and Nuremberg, Salzburg, Austria; [2]Wolfgang Wirth, PhD, Felix Eckstein, MD: Paracelsus Medical University Salzburg and Nuremberg, Salzburg, Austria, and Chondrometrics GmbH, Ainring, Germany; [3]C. Kent Kwoh, MD: University of Arizona College of Medicine, Tucson; [4]David J. Hunter, MD: Royal North Shore Hospital and University of Sydney, Sydney, New South Wales, Australia; [5]Michael J. Hannon, MA: Pinney Associates, Pittsburgh, Pennsylvania.

Dr. Wirth has received consulting fees from Galapagos NV Belgium (less than $10,000) and is a co-owner of Chondrometrics GmbH. Dr. Kwoh has received consulting fees from Fidia, Taiwan Liposome Company, Regeneron, Astellas, and Thusane (less than $10,000 each), and from Regulus, EMD Serono, and Express Scripts (more than $10,000 each), and has received grants from AbbVie, Pfizer, and EMD Serono. Dr. Hunter has received consulting fees from Merck Serono, Pfizer, TLC Bio, and Tissuegene (less than $10,000 each). Dr. Eckstein has received consulting fees from Merck Serono, Bioclinica/Synarc, Kolon Tissuegene, Samumed, AbbVie, Servier, Galapagos, Medtronic, and Roche (less than $10,000 each) and is CEO of Chondrometrics GmbH and has received grants from Pfizer, Eli Lilly and Company, Merck Serono, Novartis, Stryker, AbbVie, Kolon/Tissuegene, Synarc, Ampio, BICL, Orthotrophix, Medivir, Servier, and Galapagos. No other disclosures relevant to this article were reported.

Address correspondence to Melanie Roth, PhD, Department of Imaging and Functional Musculoskeletal Research, Paracelsus Medical University Salzburg and Nuremberg, Strubergasse 21, 5020 Salzburg, Austria. Email: melanie.roth@pmu.ac.at.

Submitted for publication June 26, 2019; accepted in revised form March 17, 2020.

SIGNIFICANCE & INNOVATIONS

- Longitudinal change in 3-dimensional quantitative measures of meniscal position and morphology are related to subsequent risk of knee replacement.
- Greater progression in loss of tibial plateau coverage and reduction of meniscal width showed a >2-fold odds of subsequent knee replacement compared to matched controls.
- The findings highlight the fact that quantitative structural measures associated with progression of knee osteoarthritis are related to an important clinical outcome.

INTRODUCTION

Knee replacement (KR) represents an important clinical and economic end point of knee osteoarthritis (OA) and is a major contributor to treatment costs associated with knee OA, the main clinical indications for KR being knee pain and functional impairment (1). Reduction of radiographic joint space width (JSW) has been associated with KR (2) and suggested to result from a combined progression of cartilaginous and meniscal pathologies (3). JSW represents the currently recommended outcome for disease-modifying knee OA drug trials by the Food and Drug Administration and the European Medicines Agency.

Quantitative cartilage thickness loss predicts KR, particularly in knees with early knee OA (i.e., Kellgren/Lawrence [K/L] grade ≤2) that exhibit rapid clinical and structural progression toward KR (4). Additionally, previous studies have analyzed the association between semiquantitative meniscal lesions and extrusion and KR, with varying results (5,6). Quantitative 3-dimensional (3-D) measures of meniscal morphology and position from magnetic resonance imaging (MRI) (7) are sensitive to longitudinal change in knee OA progression (8). However, no study has explored the relationship of quantitative meniscal measures with future KR.

The primary aim of this study was to cross-sectionally and longitudinally compare quantitative measures of meniscal morphology and position between KR case knees with rapid clinical and structural progression (4) and matched controls without subsequent KR. The second aim was to relate the observed case–control differences to those observed for cartilage thickness change (4).

PATIENTS AND METHODS

Study design. A nested, matched case–control study was conducted among participants of the Osteoarthritis Initiative (OAI), a multicenter, prospective observational cohort study (9) approved by the institutional review board for the University of California, San Francisco, and by each OAI study site (approval H5254-20499-09). All participants provided informed consent. Clinical data and 3 Tesla MRIs of the knees were acquired at annual visits (9).

To be eligible as a case for the current study, total or unicompartmental medial KR had to be recorded at the 36- through

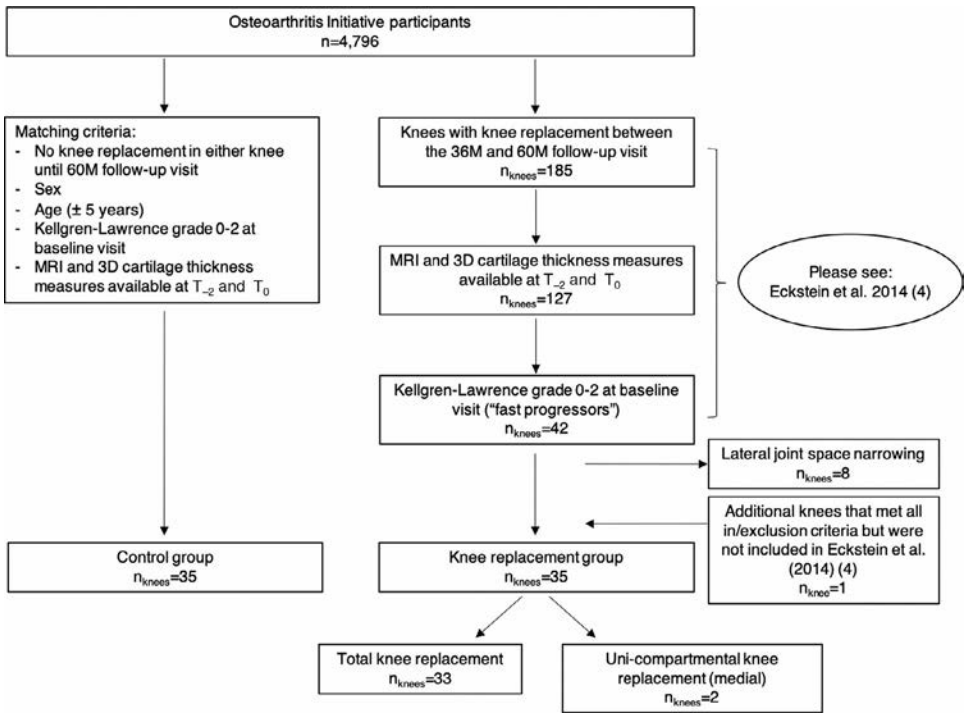

Figure 1. Flow chart depicting participant selection from the Osteoarthritis Initiative database. M = months; MRI = magnetic resonance imaging; T_{-2} = visit 2 years prior to knee replacement, T_0 = visit just before knee replacement.

60-month follow-up visit and confirmed by medical record and/or the radiographic follow-up visits. MRI acquisitions had to be available for the annual visit immediately before KR (T_0) and from the annual visit 2 years before T_0 (T_{-2}). KR cases detected at the 12-month or 24-month follow-up visit were not included, lacking the 2-year study window before KR. The analyses were restricted to knees that progressed to KR with baseline K/L grade ≤2 (OAI release 0.4 from central readings) because these knees exhibited very rapid clinical and structural progression (i.e., high rates of radiographic JSW and articular cartilage loss) (4,10). Knees with baseline lateral Osteoarthritis Research Society International joint space narrowing were excluded from the analysis because they were not likely to show medial compartment progression, and this study focused on changes in the medial meniscus. All 35 OAI knees from 33 participants fulfilling the above selection criteria were included (Figure 1).

To be eligible as a control, both knees had to be free from KR through the 60-month follow-up. Cases and controls were matched 1:1 (via algorithm) by sex, age (± 5 years), and baseline radiographic disease stage (K/L grade 0/1 or 2) (9) and were subsequently analyzed at T_{-2} and T_0 time points, dictated by the cases.

3-D quantitative meniscal measurements. In all, 3 Tesla MRIs were acquired using a sagittal double-echo steady-state sequence (DESS) with water excitation (8,9). Coronal multiplanar reconstructions were used for the segmentation of the medial meniscus at T_{-2} and T_0 (8,9) (Figure 2A). The medial tibial plateau area (i.e., the area of cartilage surface including denuded areas of subchondral bone) and the surfaces of the medial meniscus (tibial, femoral, and external area) were segmented manually by the first author (MR), with blinding to acquisition order, using

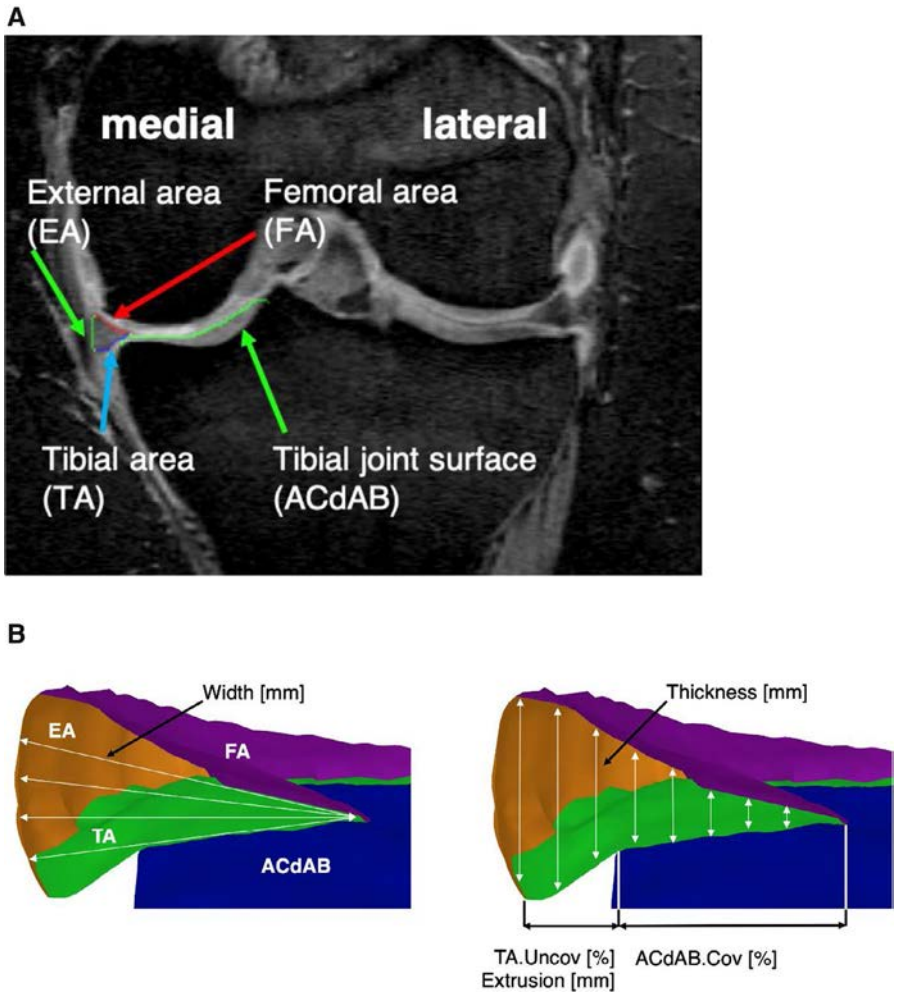

Figure 2. Three-dimensional (3-D) analysis of the menisci. Magnetic resonance imaging (MRI)–based, quantitative 3-D analysis of the menisci. **A**, Coronal MRI from a double-echo steady-state sequence with water excitation, with segmentation; **B**, 3-D reconstruction of the medial meniscus depicting joint surface of the tibia, consisting of 1) the area of the cartilage surface and denuded areas of subchondral bone (ACdAB), if applicable; 2) the femoral, tibial, and external surface of the medial meniscus; 3) meniscal thickness and width; 4) overlap between the tibial area (TA) and ACdAB (ACdAB.Cov [%]), uncovered area of the TA (TA.Uncov [%]), and 5) extrusion distance. The mean parameters represent the average, the maximum parameters the greatest value across all slices.

proprietary software (Chondrometrics GmbH) (Figure 2A), as described previously (7). All segmentations were quality controlled by 1 expert reader (KE). Meniscal morphology was characterized by the mean thickness, mean width, and volume (Figure 2B) (7). Parameters of meniscal position included the percentage of the tibial plateau area covered by the meniscus (ACdAB.Cov%), the percentage of the tibial meniscal surface area not covering the tibial plateau, and the mean and maximal extrusion distance (7) (Figure 2B).

Quantitative cartilage thickness measurements and JSW. Quantitative cartilage thickness in the total medial femorotibial compartment (MFTC) and central MFTC were available from previously performed manual cartilage segmentations from sagittal DESS with water imaging at T_{-2} and T_0 (4). Minimal medial JSW (mJSW) from fixed-flexion radiography was added as reference.

Statistical analysis. To analyze the differences between cases and controls, paired sample *t*-tests applying bootstrap (1,000 replications, simple sampling, BCa method), or Wilcoxon's signed-rank tests were used, as appropriate for the distribution. Comparisons were performed at T_{-2} and for the longitudinal changes observed between T_{-2} and T_0. Conditional logistic regression was applied to obtain case–control odds ratios (ORs). This analysis was conducted using a survival analysis procedure because the resulting hazard ratios are interpretable as ORs, given constant time to censoring. This procedure permitted us to account for potential correlation between 2 knees of the same individual (11). Case–control ORs were standardized per SD of the change from T_{-2} to T_0 in the control group. Robustness of these odds was assessed by performing adjustment for the effects of body mass index (BMI) and pain at T_{-2} (case–control OR_{bp}). This procedure was done due to reported associations between meniscal pathology and pain (12) and between cartilage loss and BMI (13). These adjustments were made for standard categories of BMI (normal/overweight/obese) and for pain frequency (no pain/infrequent pain/frequent pain [i.e., pain on most days of the past 30 days]), commonly used to classify symptomatic knee OA (2,4,9). To enable comparison of ORs across meniscal and cartilage measures, selected ORs were inverted so that a greater clinical worsening in the case group resulted in ORs >1.

Case–control areas under the receiver operating curve, adjusted for BMI and pain (AUC_{bp}) were calculated to evaluate the discrimination ability between case and control knees. The case–control AUC_{bp} was adjusted for BMI and pain at T_{-2}, as well as for age, sex, and K/L grade (matching values). For exploratory purposes, OR and AUC calculations were also performed for T_{-2} and T_0, and case–control OR_{bp} and case–control AUC_{bp} were additionally adjusted for concurrent change in central MFTC.

Case–control OR, case–control OR_{bp}, and case–control AUC_{bp} were calculated using SAS Studio software, version 3.6. The remaining analyses were done using SPSS software, version 25. Given multiple comparisons, Bonferroni adjustments were made, and *P* values less than 0.005 were considered significant.

RESULTS

Sample description. Of the 33 OAI case participants (T_{-2}: age 65 ± 7 years, BMI 30.0 ± 4.0 kg/m^2), 23 were women; of the 35 case knees, 5, 8, and 22 were K/L grade 0/1/2, respectively, at baseline, and 6, 10, and 19 had no pain/infrequent pain/frequent pain, respectively, at T_{-2}. In all, 13 received KR at the 36-month follow-up, 7 at 48 months, and 15 at 60 months; 2 were medial unicompartmental knee replacements. The median time between T_0 and time of KR was 193 days. Of the 35 control subjects (T_{-2}: age 64 ± 7 years, BMI 30.0 ± 4.9 kg/m^2), 23 were women; of the 35 knees, 9, 4, and 22 were K/L grade 0/1/2 at baseline, respectively, and 11, 4, and 20 had no pain/infrequent pain/frequent pain, respectively, at T_{-2}. Until T_{-2}, 2 cases and 1 control progressed to K/L grade 3 or 4; until T_0, 18 cases and 1 control progressed to K/L grade 3 or 4.

Meniscal measures in case and control knees. At T_{-2}, no statistically significant differences were observed between case and control knees in meniscal or cartilage parameters (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24193/abstract). Within-group analysis of the case knees between T_{-2} and T_0 showed that the loss in ACdAB.Cov% and change in meniscal extrusion measures (percentage of the tibial meniscal surface area not covering the tibial plateau, mean and maximal extrusion distance) were statistically significant. Among morphologic measures, the decrease in mean thickness was statistically significant ($P < 0.001$), but not mean width or volume (Table 1). In controls, no statistically significant changes were observed (Table 1). Between-group analysis showed statistically significant greater loss in ACdAB.Cov% and cartilage thickness in cases versus controls (Table 1).

Risk of subsequent KR. The odds of subsequent KR (defined as case–control OR_{bp}) increased with greater longitudinal changes in parameters associated with meniscal position. The case–control OR_{bp} values ranged from 1.40 for maximal extrusion distance (95% confidence interval [95% CI] 1.12, 1.75; $P < 0.005$) to 2.28 for ACdAB.Cov% (95% CI 1.43, 3.64; $P < 0.001$). Among morphologic measures, the odds of subsequent KR increased with a greater longitudinal reduction of mean width (case–control OR_{bp} 2.01 [95% CI 1.23, 3.26]; $P < 0.005$) (Table 1). Among all parameters measured, ACdAB.Cov% displayed the strongest discrimination between case–control pairs (case–control AUC_{bp} 0.78 [95% CI 0.67, 0.89]; $P < 0.001$). Case–control OR_{bp} for cartilage thickness change was 2.86 (95% CI 1.51, 5.41; $P < 0.001$) for MFTC, and 3.52 (95% CI 1.15, 10.72; $P = 0.027$) for central MFTC (Table 1). Case–control OR_{bp} for mJSW change was 3.56 (95% CI 1.95, 6.51; $P < 0.001$).

Table 1. Two-year within-group and between-group changes in medial meniscal, cartilage, and joint space width parameters ($T_{-2} > T_0$), with case–control odds ratios (ORs) for knee replacement and case–control areas under the receiver operating curve (AUC)*

	Change $T_{-2} > T_0$		Case–control OR			Case–control AUC	
	Knee replacement (n = 35)	Control (n = 35)	OR	Adjusted for BMI, pain	Adjusted for BMI, pain, central MFTC	Adjusted for BMI, pain	Adjusted for BMI, pain, central MFTC
Meniscal position							
ACdAB.Cov%	−6.3 (−9.1, −3.6)†	−0.1 (−1.2, 1.0)	2.13 (1.20, 3.78)	2.28 (1.43, 3.64)‡	2.26 (1.04, 4.90)	0.77 (0.66, 0.88)‡	0.78 (0.67, 0.89)‡
TA Uncov%§	6.8 (3.4, 10.2)‡	0.9 (−0.4, 2.1)	1.43 (1.05, 1.94)	1.51 (1.18, 1.94)¶	1.07 (0.76, 1.50)	0.74 (0.62, 0.86)‡	0.73 (0.61, 0.85)‡
Ex. mean, mm§	0.8 (0.4, 1.2)¶	0.1 (−0.1, 0.3)	1.36 (1.00, 1.85)	1.42 (1.11, 1.82)	1.10 (0.74, 1.62)	0.74 (0.62, 0.86)‡	0.73 (0.61, 0.85)‡
Ex. max, mm§	0.6 (0.2, 1.0)¶	0.2 (0.0, 0.3)	1.31 (1.03, 1.67)	1.40 (1.12, 1.75)¶	0.91 (0.67, 1.25)	0.61 (0.48, 0.75)	0.61 (0.47, 0.75)
Meniscal morphology							
Width mean, mm	−0.5 (−0.8, −0.3)‡	−0.1 (−0.3, 0.1)	1.84 (1.08, 3.14)	2.01 (1.23, 3.26)¶	1.98 (1.21, 3.23)	0.72 (0.60, 0.84)‡	0.72 (0.60, 0.84)‡
Thickness mean, mm§	0.1 (0.0, 0.1)	0.0 (0.0, 0.1)	1.48 (1.07, 2.05)	1.41 (1.00, 2.00)	0.84 (0.46, 1.54)	0.56 (0.42, 0.70)	0.56 (0.42, 0.70)
Volume, mm^3§	12 (−59, 84)	0 (−54, 55)	0.94 (0.69, 1.27)	1.00 (0.73, 1.38)	1.17 (0.69, 1.97)	0.49 (0.36, 0.63)	0.50 (0.36, 0.64)
Cartilage thickness							
MFTC, mm	−0.4 (−0.6, −0.2)†	0.0 (−0.1, 0.0)	2.48 (1.51, 4.07)‡	2.86 (1.51, 5.41)¶	NA	0.71 (0.59, 0.83)‡	NA
Central MFTC, mm	−0.7 (−1.0, −0.4)†	0.0 (−0.1, 0.1)	3.04 (1.26, 7.33)	3.52 (1.15, 10.72)	NA	0.76 (0.64, 0.87)‡	NA
Joint space width							
mJSW, mm	−1.3 (−1.9, −0.8)†	−0.2 (−0.5, 0.0)	3.03 (1.96, 5.57)‡	3.56 (1.95, 6.51)‡	1.64 (0.66, 4.03)	0.74 (0.60, 0.87)‡	0.74 (0.61, 0.87)‡

* Values are the mean (95% confidence interval) unless indicated otherwise. Adjustment for body mass index (BMI) and pain was performed at T_{-2}. Ex. max = maximal meniscal extrusion; Ex. mean = mean meniscal extrusion; MFTC = medial femorotibial compartment; mJSW = minimal medial joint space width; NA = not applicable; T_{-2}/T_0 = visit 2 years prior to knee replacement/visit just before knee replacement; TA Uncov% = percentage of tibial meniscal surface area not covering the tibial plateau; ACdAB.Cov% = percentage of tibial plateau area covered by the meniscus.

† Within group and between group, $P < 0.001$.

‡ Within group, $P < 0.001$.

§ Case–control OR adjusted for BMI and pain is presented as 1/case–control OR:BMI, pain adjusted.

¶ Within group, $P < 0.005$.

Exploratory analyses at T_{-2} showed that no cross-sectional parameter was associated (case–control OR_{bp}) with KR 2 years later (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24193/abstract). At T_0, however, cross-sectional meniscal position displayed greater odds for KR than cartilage thickness and mJSW (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24193/abstract). Exploratory analyses with central MFTC as an additional covariate mostly showed reduced effect sizes (Table 1 and Supplementary Tables 1 and 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24193/abstract).

DISCUSSION

The objective of this study was to assess whether change in quantitative 3-D measures of meniscal position and morphology is associated with clinical progression from early knee OA to KR and to compare these associations to those previously reported for cartilage thickness loss. Our results show that odds were greater for KR knees to display a longitudinal decrease in tibial coverage by the meniscus, an increase in extrusion, and a reduction in meniscal width than for control knees. Although these odds did not appear to be as strong as those for cartilage thickness and mJSW change, receiver operating curve analysis suggested that meniscal positional parameters display discrimination ability between cases and controls similar to cartilage thickness and mJSW.

A limitation of this study is the limited sample size due to the selection of baseline knees with K/L grades 0–2 with rapid progression toward KR (4,10). This choice was made because considerable JSW and cartilage thickness changes were previously observed in this sample (4,10); thus, the question of interest could be pursued efficiently. Since cartilage thickness change was previously shown to differ substantially between case and control knees in this sample (4), and since comparison of meniscal with cartilaginous parameters was one of the main objectives of this analysis, the study was adequately powered. Further, precise matching of cases versus controls with the same sex and similar radiographic stage and age, as well as the statistical adjustments for pain and BMI, minimized the potential impact of confounders. The results clearly encourage further exploration of the relationship of quantitative meniscal measures with relevant clinical outcomes of knee OA in larger samples and knees with advanced radiographic knee OA (i.e., K/L grades 3–4) to better understand the previous differences in the rate of radiographic JSW loss between KR case and non-KR control knees (2,4). This may reflect differences in the rate of change in cartilage thickness but also in differences in changes of meniscal width and extrusion.

A strength of the study was the temporal synchronization to a clinically important end point (KR surgery); nonetheless, being aware of its dependency on, among others, comorbidities, willingness to undergo surgery, and socioeconomic status. Although the T_{-2} to T_0 window is a relatively long period for clinical decision-making, it was chosen because strong differences in cartilage loss were previously observed between case and control knees over this period in a similar sample (4), and because 2-year changes are more robust. This study focused on the medial femorotibial compartment because it is more commonly affected by knee OA than the lateral compartment and has greater mechanical load transmitted through it (14).

Another limitation of this study is that it focused on quantitative meniscal measures, not taking into account the number or type of meniscal tears during the study period. However, our finding that a longitudinal reduction in meniscal tibial plateau coverage and a longitudinal increase in meniscal extrusion predict KR increases the knowledge on this topic. Previous studies employed semiquantitative scoring to assess whether the presence of meniscal extrusion predicts KR (5,6). While Roemer et al found no significant association between cross-sectional meniscal extrusion 1 year before KR and KR itself, using the MRI Osteoarthritis Knee Score on ~200 knee OA KR knees (5), Hafezi-Nejad et al reported baseline meniscal extrusion to be associated with subsequent KR when using the Boston Leeds Osteoarthritis Knee Score on 25 knee OA KR knees, but not when using the Whole-Organ Magnetic Resonance Imaging Score (6). In addition, 2-year change in meniscal extrusion assessed by either scoring system did not predict subsequent KR (6).

The fully quantitative 3-D method applied in our study, however, allowed for determination of 3-D meniscal extrusion using continuous measures. Further, the 3-D method applied here reports the percent coverage of the tibial plateau by the meniscus, a measure that no semiquantitative scoring system evaluates. From a functional and biomechanical perspective, measures of tibial plateau coverage are potentially superior to measures of extrusion because they may provide more relevant information on the mechanical protection of the cartilage by the meniscus. Indeed, parameters of tibial plateau coverage displayed a strong relationship with subsequent KR in this study. This finding is in line with responsiveness to 1-year change reported applying shape models in a progression but non-KR cohort (15).

This study not only evaluated quantitative measures of meniscal position in the context of predicting KR but also quantitative measures of meniscal morphology. Meniscal thickness tended to increase before KR, and the volume appeared to remain constant, whereas meniscal width displayed a significant reduction. This reduction in meniscal width may reflect substance loss at the internal edge of the meniscus, which in conjunction with increased meniscal extrusion may contribute to a mechanically unfavorable reduction in tibial plateau coverage and increase mechanical stress on the tibiofemoral cartilage before KR. This mechanism is reflected by the high discrimination ability of ACdAB.Cov% between KR case and non-KR control knees that was observed to be of similar strength to that of cartilage thickness loss. These observations support current

surgical strategies aiming at maintaining as much of the degenerated menisci as is functionally useful to protect femorotibial cartilage from undue mechanical loading.

The ORs for cartilage thickness observed differ slightly from those previously reported (4) because the current analysis focused on the medial meniscus, and thus, few subjects with lateral joint space narrowing were excluded. Overall, the odds of subsequent KR were observed to be somewhat weaker for meniscal morphology and position than for cartilage thickness and mJSW loss. Nevertheless, since the 95% CIs of the ORs overlap, one cannot be regarded as significantly better or worse than the other. The observation that change in meniscal parameters, and not only the change in cartilage thickness parameters, is associated with KR reinforces the conception of knee OA to be a whole joint disease. OA is a disease affected by changes in most, if not all, articular structures, with changes in 1 structure often happening concurrently with changes in another structure.

In conclusion, this study suggests that quantitative measures of meniscal morphology and position, namely tibial plateau coverage, meniscal extrusion, and meniscal width, predict subsequent KR in rapidly progressing knee OA. The findings show that not only cartilage thickness loss, but also longitudinal change in quantitative meniscal measures, are related to an economically important clinical outcome of knee OA.

ACKNOWLEDGMENTS

The authors thank the OAI participants and investigators, as well as the OAI Clinical Center's staff for generating this publicly available image data set.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Roth, Wirth, Eckstein.

Acquisition of data. Roth, Emmanuel, Wirth, Kwoh, Hannon.

Analysis and interpretation of data. Roth, Wirth, Hunter, Hannon, Eckstein.

REFERENCES

1. Gossec L, Paternotte S, Maillefert JF, Combescure C, Conaghan PG, Davis AM, et al. The role of pain and functional impairment in the decision to recommend total joint replacement in hip and knee osteoarthritis: an international cross-sectional study of 1909 patients. Report of the OARSI-OMERACT Task Force on total joint replacement. Osteoarthritis Cartilage 2011;19:147–54.
2. Eckstein F, Boudreau R, Wang Z, Hannon MJ, Duryea J, Wirth W, et al. Comparison of radiographic joint space width and magnetic resonance imaging for prediction of knee replacement: a longitudinal case-control study from the Osteoarthritis Initiative. Eur Radiol 2016;26:1942–51.
3. Bloecker K, Wirth W, Hunter DJ, Duryea J, Guermazi A, Kwoh CK, et al. Contribution of regional 3D meniscus and cartilage morphometry by MRI to joint space width in fixed flexion knee radiography: a between-knee comparison in subjects with unilateral joint space narrowing. Eur J Radiol 2013;82:e832–9.
4. Eckstein F, Boudreau RM, Wang Z, Hannon MJ, Wirth W, Cotofana S, et al. Trajectory of cartilage loss within 4 years of knee replacement: a nested case-control study from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2014;22:1542–9.
5. Roemer FW, Kwoh CK, Hannon MJ, Hunter DJ, Eckstein F, Wang Z, et al. Can structural joint damage measured with MR imaging be used to predict knee replacement in the following year? Radiology 2015;274:810–20.
6. Hafezi-Nejad N, Zikria B, Eng J, Carrino JA, Demehri S. Predictive value of semi-quantitative MRI-based scoring systems for future knee replacement: data from the Osteoarthritis Initiative. Skeletal Radiol 2015;44:1655–62.
7. Wirth W, Frobell RB, Souza RB, Li X, Wyman BT, Le Graverand MP, et al. A three-dimensional quantitative method to measure meniscus shape, position, and signal intensity using MR images: a pilot study and preliminary results in knee osteoarthritis. Magn Reson Med 2010;63:1162–71.
8. Roth M, Emmanuel K, Wirth W, Kwoh CK, Hunter DJ, Eckstein F. Sensitivity to change and association of three-dimensional meniscal measures with radiographic joint space width loss in rapid clinical progression of knee osteoarthritis. Eur Radiol 2018;28:1844–53.
9. Eckstein F, Kwoh CK, Link TM. Imaging research results from the Osteoarthritis Initiative (OAI): a review and lessons learned 10 years after start of enrolment. Ann Rheum Dis 2014;73:1289–300.
10. Eckstein F, Kwoh CK, Boudreau R, Wang Z, Hannon M, Cotofana S, et al. Quantitative magnetic resonance imaging measures of cartilage predict knee replacement: a case-control study from the Osteoarthritis Initiative. Ann Rheum Dis 2013;72:707–14.
11. Gaccione P. More on proc phreg: conditional logistic regression (or should we do another study..?). SAS Conf Proc Northeast SAS Users Gr 1995:791–800.
12. Wenger A, Englund M, Wirth W, Hudelmaier M, Kwoh K, Eckstein F. Relationship of 3D meniscal morphology and position with knee pain in subjects with knee osteoarthritis: a pilot study. Eur Radiol 2012;22:211–20.
13. Pelletier JP, Raynauld JP, Berthiaume MJ, Abram F, Choquette D, Haraoui B, et al. Risk factors associated with the loss of cartilage volume on weight-bearing areas in knee osteoarthritis patients assessed by quantitative magnetic resonance imaging: a longitudinal study. Arthritis Res Ther 2007;9:R74.
14. Eckstein F, Wirth W, Hudelmaier M, Stein V, Lengfelder V, Cahue S, et al. Patterns of femorotibial cartilage loss in knees with neutral, varus, and valgus alignment. Arthritis Rheum 2008;59:1563–70.
15. Dube B, Bowes MA, Kingsbury SR, Hensor EM, Muzumdar S, Conaghan PG. Where does meniscal damage progress most rapidly? An analysis using three-dimensional shape models on data from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2018;26:62–71.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 1038–1040
DOI 10.1002/acr.24213

BRIEF REPORT

Achievement of the 2019 European Alliance of Associations for Rheumatology/American College of Rheumatology Criteria for Systemic Lupus Erythematosus and Amount of Damage Accrual: Results From a Multiethnic Multicenter Cohort

Manuel F. Ugarte-Gil,[1] Guillermo J. Pons-Estel,[2] Russell Griffin,[3] Luis M. Vilá,[4] John D. Reveille,[5] and Graciela S. Alarcón[6]

Objective. To determine the difference in outcomes in patients who achieved or did not achieve the 2019 European Alliance of Associations for Rheumatology (EULAR)/American College of Rheumatology (ACR) criteria.

Methods. Patients from the LUpus in MInorities, NAture versus nurture (LUMINA) cohort were included. For these analyses, we compared those patients who achieved the 2019 EULAR/ACR criteria any time during follow-up to those who did not. The predefined outcomes were the last Systemic Lupus International Collaborating Clinics/ACR Damage Index (SDI) scores and survival. Univariable and multivariable negative binomial regression models were performed; adjustment models were based on a forward selection process.

Results. In total, 98 of 640 patients never achieved the 2019 EULAR/ACR criteria. There was no difference in mean baseline SDI score among the patients who did not achieve the criteria compared to those who did. Conversely, the mean ± SD SDI score at last visit was lower for those who never achieved the criteria (1.2 ± 1.7 versus 2.0 ± 2.3, $P = 0.0004$). In the final adjusted model, the SDI score at last visit was 31% lower for those who never achieved the criteria ($P = 0.0077$). These patients were also more likely to survive, but this was not statistically significant.

Conclusion. In our cohort, patients who did not achieve the 2019 EULAR/ACR criteria accrued less damage, suggesting that these criteria could allow us to identify a subset of patients with more severe disease than allowed by previous criteria.

INTRODUCTION

The 2019 European Alliance of Associations for Rheumatology (EULAR)/American College of Rheumatology (ACR) criteria for the classification of patients with systemic lupus erythematosus (SLE) were proposed in order to improve the sensitivity and specificity of the previously published SLE criteria (1,2).

These criteria, however, need to be validated in several ethnic groups; for example, in 2 multiethnic, multicenter cohorts, one from the US (LUpus in MInorities, NAture versus nurture [LUMINA]) and the other from Latin America (Grupo Latinoamericano de Estudio del Lupus Eritematoso [GLADEL]), we found a sensitivity of 84.8% and 91.3%, respectively, using as gold-standard the 1982/1997 ACR revised classification criteria for SLE as updated in 1997 (3–7). Similarly, in Brazil, in a cohort of childhood-onset SLE patients, the sensitivity of the EULAR/ACR criteria was 87.7%, and the specificity was 67.4% using as gold-standard the clinical criteria (8), while in the Netherlands, the sensitivity was 87.% and the specificity was 74% in a cohort of SLE patients with neuropsychiatric symptoms (9).

[1]Manuel F. Ugarte-Gil, MD, MSc: Hospital Nacional Guillermo Almenara Irigoyen, EsSalud and Universidad Científica del Sur, Lima, Peru; [2]Guillermo J. Pons-Estel, MD, PhD: Hospital Provincial de Rosario, Rosario, and Grupo Oroño, Sanatorio Parque Rosario, Santa Fe, Argentina; [3]Russell Griffin, PhD: University of Alabama at Birmingham; [4]Luis M. Vilá, MD: University of Puerto Rico Medical Sciences Campus, San Juan, Puerto Rico; [5]John D. Reveille, MD: The University of Texas Health McGovern Medical School, Houston; [6]Graciela S. Alarcón, MD, MPH: University of Alabama at Birmingham, and Universidad Peruana Cayetano Heredia, Lima, Peru.

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

Address correspondence to Manuel F. Ugarte-Gil, MD, MSc, School of Medicine, Universidad Cientifica del Sur, Lima, Peru, Avenida Panamericana Sur KM19, Villa El Salvador, Lima 15067, Peru. Email: mugarte@cientifica.edu.pe.

Submitted for publication October 31, 2019; accepted in revised form April 7, 2020.

SIGNIFICANCE & INNOVATIONS

- The European Alliance of Associations for Rheumatology/American College of Rheumatology classification criteria for systemic lupus erythematosus have been developed in order to improve the sensitivity and specificity of previous criteria.
- However, the impact of not achieving these criteria on the course of the disease has not been properly evaluated.
- In this multiethnic lupus cohort, not achieving these new criteria is associated with a better outcome.

In addition to considering sensitivity and specificity, it is important to determine if patients who would not be classified as having SLE according to these criteria have different short- and long-term outcomes than those who could classified as having SLE. Thus, the aim of this study was to determine the difference in damage accrual and survival in patients who achieved or who did not achieve these new criteria.

PATIENTS AND METHODS

The LUMINA cohort has been previously described (10,11). SLE was defined using the 1982/1997 ACR revised classification criteria for SLE as updated in 1997 (6,7), and disease duration could have been up to 5 years at cohort entry.

We have to point out that some manifestations were not recorded, and that precluded us from including them in the analyses (fever, alopecia, delirium, acute pericarditis, and complement levels). Clinical and laboratory variables were measured at cohort entry, at 6 and 12 months, and yearly thereafter.

For these analyses, we compared those patients who achieved the 2019 EULAR/ACR criteria any time during the follow-up to those who did not. The predefined outcomes were the Systemic Lupus International Collaborating Clinics/ACR Damage Index (SDI) scores and survival at last visit. Possible confounders included were age, sex, race/ethnicity, poverty, disease duration at baseline, baseline Systemic Lupus Activity Measure (SLAM) score (for disease activity), baseline SDI score, and treatment with antimalarials, glucocorticoids, and immunosuppressive drugs at baseline. For the last SDI score and survival, univariable and multivariable negative binomial regression models were performed; adjustment models were based on a forward selection process. For the damage analysis, 2 alternative analyses were performed, the first one in order to evaluate the impact of follow-up on damage accrual, dividing the patients in analyses with <1 year, 1–3 years, 3–6 years, and ≥6 years, and the second one in order to evaluate the impact of being classified earlier, at the same time, or later using the 2019 EULAR/ACR criteria. In order to determine the impact of achieving these criteria in each domain, binary logistic regression models were performed using as outcome the increase in damage per domain, and as possible confounders age, sex, race/ethnicity, poverty, disease duration at baseline, baseline SLAM score (for disease activity), baseline SDI score (same domain), and treatment with antimalarials, glucocorticoids, and immunosuppressive drugs at baseline. For the survival analysis, a single model was performed (survival among those achieving and those not achieving the criteria). Adjustment variables were similar to those used in the damage analysis. The statistical analyses were performed using SAS software, version 9.1.3.

RESULTS

In total, 98 of 640 patients never achieved the 2019 EULAR/ACR criteria. The 98 patients were older and less likely to be Hispanic or African American than those who did achieve the criteria. There was no difference in mean ± SD baseline SDI score among the patients who did not achieve the criteria (0.6 ± 1.2) compared to those who achieved the criteria (0.8 ± 1.2, $P = 0.3580$). Conversely, the mean ± SDI score at last visit was lower for those who never achieved the criteria (1.2 ± 1.7 versus 2.0 ± 2.3, $P = 0.0004$).

In the final adjusted model (Table 1), the SDI score at last visit was 31% lower for those not achieving the criteria ($P = 0.0077$). When the analyses were performed dividing the patients according to the length of follow-up, in the group with at least 6 years of follow-up, those not achieving the criteria accrued less damage (estimate –0.96, SEM 0.41; $P = 0.0205$); the association in the other subgroups did not remain significant. There were no differences in damage accrual if the patients were classified earlier, at the same time, or later using the 2019 EULAR/ACR criteria (data not shown).

Table 1. Impact of achieving or not achieving the 2019 European Alliance of Associations for Rheumatology (EULAR)/American College of Rheumatology (ACR) criteria on damage accrual*

Parameter	Estimate	SEM	*P*
Intercept	–1.6924	0.2224	<0.0001
Not achieving the 2019 EULAR/ACR criteria	–0.3773†	0.1416	0.0077
Age at baseline, years	0.0174	0.0037	<0.0001
Ethnicity			
Hispanic (Texas)	0.9339	0.1740	<0.0001
African American	0.9513	0.1592	<0.0001
White	0.5969	0.1661	0.0003
Puerto Rican	Ref.		
SLAM score at baseline	0.0562	0.0084	<0.0001
Poverty	0.1785	0.0946	0.0593
Disease duration at baseline, years	0.1439	0.0327	<0.0001
Receiving immunosuppressive drugs at baseline	0.3626	0.1125	0.0013

* In addition to the variable of interest (not achieving the 2019 EULAR/ACR criteria), possible confounders associated with damage in the univariable analyses were included. Ref. = reference; SLAM = Systemic Lupus Activity Measure.
† The ratio of the average damage was 0.69 (not achieving versus achieving the 2019 EULAR/ACR criteria), which translates into a 31% decreased probability of not accruing damage for those not achieving the criteria.

In the per domain analyses, the same trend was shown for all domains, but significance was only achieved for the ocular domain (data not shown). In total, 87 patients (88.8%) who never achieved the 2019 EULAR/ACR criteria were still alive at the end of follow-up, and 456 (84.1%) among those who achieved the criteria. This difference, however, was not statistically significant in both univariable and multivariable models (data not shown).

DISCUSSION

Our study shows that patients who did not achieve the 2019 EULAR/ACR criteria accrued less damage than those who did, in particular those who had a longer follow-up. Our findings are consistent with the data reported by Carneiro et al, according to whom a higher EULAR/ACR criteria score was associated with a higher SDI score adjusted for age and sex (12); however, they did not adjust the model for other possible confounders. Additionally, Carneiro et al found that the higher the score with the new criteria, the more likely the presence of renal damage after adjusting for age and sex.

In the LUMINA and GLADEL cohorts, those who were classified with the 2019 EULAR/ACR criteria earlier than with the 1982/1997 ACR criteria had a lower frequency of milder manifestations at baseline, suggesting that these criteria could be useful for patients with a more severe SLE subset (3,13), which is consistent with our findings about the impact of achieving the criteria on damage.

Although there was lower survival among patients who achieved the new criteria versus those who did not, this difference was not significant, which probably relates to the sample size; these data are consistent with the data reported by Carneiro et al (12).

Our study has some limitations. First, as we did not collect some variables included in the new criteria in our cohort, this could have some impact on our results. Second, as all patients should have satisfied the 1982/1997 ACR criteria, patients who could achieve the EULAR/ACR criteria and not the 1982/1997 ACR criteria were not included in the cohort. Third, the lack of association between survival and not achieving the EULAR/ACR criteria could be due to lack of power; so this should be examined in larger cohorts. In conclusion, we found that not achieving the EULAR/ACR criteria was associated with a better prognosis even after adjustment for possible confounders.

ACKNOWLEDGMENTS

The authors thank the LUMINA patients and staff, who made this cohort possible.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Ugarte-Gil, Pons-Estel, Griffin, Alarcón.
Acquisition of data. Pons-Estel, Alarcón.
Analysis and interpretation of data. Ugarte-Gil, Pons-Estel, Griffin, Vilá, Reveille, Alarcón.

REFERENCES

1. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71:1400–12.
2. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Ann Rheum Dis 2019;78:1151–9.
3. Pons-Estel GJ, Ugarte-Gil MF, Harvey GB, Wojdyla D, Quintana R, Saurit V, et al. Applying the 2019 EULAR/ACR lupus criteria to patients from an established cohort: a Latin American perspective. RMD Open 2020;6:e001097.
4. Ugarte-Gil MF, Pons-Estel GJ, Griffin R, Harvey GB, Wojdyla D, Vila LM, et al. 83 classification of systemic lupus erythematosus patients: the 2018 EULAR/ACR criteria vs the 1982/1997 ACR criteria: a comparative study in two multicenter multiethnic cohorts. Lupus Sci Med 2019;6 Suppl 1:A59–60.
5. Pons-Estel GJ, Ugarte-Gil MF, Harvey GB, Wojdyla D, Quintana R, Saurit V, et al. Applying the 2019 EULAR/ACR lupus criteria to patients from an established cohort: a Latin American perspective. RMD Open 2020;6:e001097.
6. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271–7.
7. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.
8. Rodrigues Fonseca A, Felix Rodrigues MC, Sztajnbok FR, Gerardin Poirot Land M, Knupp Feitosa de Oliveira S. Comparison among ACR1997, SLICC and the new EULAR/ACR classification criteria in childhood-onset systemic lupus erythematosus. Adv Rheumatol 2019;59:20.
9. Gegenava M, Beaart HJ, Monahan RC, Brilman E, Beaart-van de Voorde LJ, Magro-Checa C, et al. Performance of the proposed ACR-EULAR classification criteria for systemic lupus erythematosus (SLE) in a cohort of patients with SLE with neuropsychiatric symptoms. RMD Open 2019;5:e000895.
10. Alarcón GS, Friedman AW, Straaton KV, Moulds JM, Lisse J, Bastian HM, et al. Systemic lupus erythematosus in three ethnic groups. Part III. A comparison of characteristics early in the natural history of the LUMINA cohort. LUpus in MInorities, NAture versus nurture. Lupus 1999;8:197–209.
11. Alarcón GS, Roseman J, Bartolucci AA, Friedman AW, Moulds JM, Goel N, et al, for the LUMINA Study Group. Systemic lupus erythematosus in three ethnic groups. Part II. Features predictive of disease activity early in its course. Arthritis Rheum 1998;41:1173–80.
12. Carneiro AC, Ruiz MM, Freitas S, Isenberg D. A comparison of three classification criteria sets for systemic lupus erythematosus: a study looking at links to outcome and mortality. Arthritis Care Res (Hoboken) 2020;72:1611–4.
13. Ugarte-Gil MF, Pons-Estel GJ, Harvey GB, Vilá LM, Griffin R, Alarcón GS. Applying the 2019 EULAR/ACR lupus criteria to patients from the LUMINA cohort: results from LUMINA (LXXXI), a multi-ethnic, multi-center US cohort. Arthritis Care Res (Hoboken) doi: http://onlinelibrary.wiley.com/doi/10.1002/acr.24367/abstract. E-pub ahead of print.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 1041–1048
DOI 10.1002/acr.24220

American College of Rheumatology
Empowering Rheumatology Professionals

Does Muscle Strength Change Over Time in Patients With Hypermobile Ehlers-Danlos Syndrome/Hypermobility Spectrum Disorder? An Eight-Year Follow-Up Study

Marie Coussens,[1] Patrick Calders,[1] Bruno Lapauw,[2] Bert Celie,[1] Thiberiu Banica,[2] Inge De Wandele,[2] Verity Pacey,[3] Fransiska Malfait,[2] and Lies Rombaut[2]

Objective. Reduced maximal muscle strength and strength endurance have been found in patients with hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorder (hEDS/HSD) and are recognized as common associated features of the disorder. However, the extent to which these parameters change over time is currently not documented. Therefore, the purpose of this 8-year follow-up study was to investigate this evolution.

Methods. Thirty female patients (mean age 41 years) with hEDS/HSD and 17 controls participated at baseline and 8 years later. Maximal muscle strength and strength endurance tests of the knee flexors and extensors, and 2 lower-extremity posture maintenance tests were performed to evaluate static strength endurance. In addition, muscle mass and density were evaluated by dual-energy x-ray absorptiometry and peripheral quantitative computed tomography.

Results. Maximal muscle strength and strength endurance were significantly lower at both baseline and follow-up in the hEDS/HSD group compared to the control group ($P \leq 0.007$). Maximal muscle strength of the knee flexors (decreased in the control group: $p\eta^2 = 0.139$), strength endurance of the knee extensors (decreased in the hEDS/HSD group and increased in the control group: $p\eta^2 = 0.244$), and muscle density (decreased in the hEDS/HSD group: $p\eta^2 = 0.263$) showed a significantly different evolution over 8 years. No other significant differences in evolution were found.

Conclusion. Decreased muscle strength was identified at both time points in patients with hEDS/HSD. The evolution of most muscle strength parameters over time did not significantly differ between groups. Future studies should focus on the effectiveness of different types of muscle training strategies in hEDS/HSD patients.

INTRODUCTION

Ehlers-Danlos syndrome (EDS) is a heterogeneous group of hereditary connective tissue disorders caused by mutations in the genes encoding for collagen or enzymes involved in the processing or modification of collagen. Hence, the most important consequences are joint hypermobility, tissue fragility, and skin hyperextensibility (1). The current EDS classification distinguishes 13 subtypes, caused by defects in 19 different genes (2). However, the genetic basis of the hypermobile type of EDS, which is the most common subtype, remains largely unknown and diagnosis is therefore solely based on clinical criteria.

Over time, these clinical criteria have been revised to describe hypermobile EDS (hEDS) in detail and to delineate it from related conditions. Initially, "the hypermobility type of Ehlers-Danlos syndrome" (EDS-HT) was described based on its major and minor clinical characteristics, which include generalized joint hypermobility and a hyperextensible or soft velvety skin (major criteria), a positive family history, recurrent joint dislocations, and chronic pain (minor criteria) (1). In 2017, the description was refined and now also emphasizes associated soft tissue fragility (e.g., multiple abdominal hernias, prolapse of organs at the level of the pelvic floor, etc.) (2). By consensus, the hypermobile type of EDS is now referred to as hEDS. Patients with a previous diagnosis of EDS-HT who no longer fully meet the stricter 2017 criteria for hEDS, are now described as having "hypermobility spectrum disorder" (HSD). Consequently, a group of patients diagnosed with EDS-HT in the past now consists of patients with hEDS and HSD.

[1]Marie Coussens, MSc, Patrick Calders, PhD, Bert Celie, PhD: Ghent University, Ghent, Belgium; [2]Bruno Lapauw, PhD, MD, Thiberiu Banica, MSc, MD, Inge De Wandele, PhD, Fransiska Malfait, PhD, MD, Lies Rombaut, PhD: Ghent University Hospital, Ghent, Belgium; [3]Verity Pacey, PhD: Macquarie University, Sydney, New South Wales, Australia.

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

Address correspondence to Marie Coussens, MSc, Ghent University, Department of Rehabilitation Sciences and Physiotherapy, Corneel Heymanslaan 10, 9000 Ghent, Belgium. Email: Marie.Coussens@Ugent.be.

Submitted for publication August 23, 2019; accepted in revised form April 7, 2020.

SIGNIFICANCE & INNOVATIONS

- Patients with hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorder (hEDS/HSD) demonstrate substantially higher pain scores, greater functional impairment, and reduced lower-extremity muscle strength continuing over an 8-year period in comparison with controls.
- Patients with hEDS/HSD show no muscle atrophy or higher muscle lipid content in comparison with controls.
- Muscle strength parameters remain relatively stable in patients with hEDS/HSD after an 8-year period, in which several factors may play a role, such as physical therapy and exercises.

In addition to generalized joint hypermobility and recurrent joint dislocations, patients with hEDS/HSD report multiple other musculoskeletal symptoms and problems. In 2012, Rombaut et al identified reduced maximal muscle strength and muscle strength endurance in 43 patients with hEDS/HSD, compared to healthy controls (3). This decrease may result from musculoskeletal pain and exercise avoidance and is likely to be related to abnormalities of the extracellular matrix of the muscle (2,4,5). By illustration, a study by Rombaut et al found an increased tendon extensibility in patients with hEDS/HSD compared to controls, which may lead to a decreased efficiency in force transmission (6). Another argument that supports the link between decreased muscle strength and connective tissue involvement is that mild-to-moderate neuromuscular involvement has also been found in several other types of EDS (4).

Unfortunately, decreased muscle strength further compromises joint stability and contributes to altered movement patterns and overload injuries in this patient population. Moreover, reduced maximal muscle strength and muscle strength endurance, muscle cramps, ruptures, and pain are related to activity limitations in hEDS/HSD (3,5). Although Castori et al (7,8) mentioned muscle weakness as part of the disease progression in hEDS/HSD, longitudinal studies about the evolution of muscle weakness over time are lacking (7–9). Because muscle weakness is a major contributor to functional impairment, knowledge of how muscle strength changes over time may provide a crucial understanding of the quality of life, prognosis, and follow-up of patients with hEDS/HSD (3). Therefore, this longitudinal study aimed to investigate the evolution of maximal muscle strength, muscle strength endurance, muscle mass, and density in patients with hEDS/HSD over a period of 8 years.

PATIENTS AND METHODS

Participants. This study protocol was reviewed and approved by the Ethical Committee of Ghent University Hospital (EC number 2017/1278), and written informed consent was obtained from all participants. Female patients diagnosed according to the Villefranche criteria and controls matched for sex and age were selected in 2009–2010 (at baseline or T1), as described by Rombaut et al (3). In 2017 (at follow-up or T2), patients and healthy controls were contacted a second time. Thirty patients with hEDS/HSD (follow-up rate of 70% or n = 30 of 43) and 17 controls (follow-up rate of 40% or n = 17 of 43) participated again at T2. The main reasons for dropout were no up-to-date contact details, work commitments, or no interest. Of the 30 patients previously diagnosed with EDS-HT, 10 patients had an hEDS diagnosis according to the 2017 EDS nosology, while 20 were reclassified as having HSD (2,9). No differences in muscle characteristics between participants with hEDS and HSD were found. Due to the small group hEDS patients (n = 10), further analyses were performed on the total patient group, referred to as hEDS/HSD in this article.

Procedure. Participants were invited by email or phone to participate in this follow-up study at Ghent University Hospital. A few weeks before the measurements were obtained, each participant was asked to fill in a self-developed follow-up questionnaire evaluating physical therapy, sports, physical activities, and medical history (injuries, surgeries, and pregnancies) over the past 8 years.

Subject characteristics, including age, height, weight, and body mass index (BMI) were collected. Lean mass of the dominant leg (LMDL, kg) and subtotal lean mass (SLM [whole body without the head], kg) were evaluated by total body dual x-ray absorptiometry (DXA) with a Hologic QDR-Discovery device (Hologic software, version 2.3.1) (3). Furthermore, muscle density of the dominant leg (mg/cm^3), which reflects the lipid content of the muscle (the lower the muscle density, the higher the lipid content), was measured by peripheral quantitative computed tomography with an XCT-2000 device (Stratec, Medizintechnik) (10). Subsequently, participants were evaluated according to the protocol described below.

Measurements. Prior to obtaining measurements, general average pain severity on the day of the tests was measured using a visual analog scale (VAS, mm) (3). Maximal muscle strength of the knee flexors (hamstrings) and extensors (quadriceps) was evaluated by isokinetic tests (Biodex) at an angular velocity of 60°/second and 5 repetitions, following the protocol described by Rombaut et al (3). If test results showed a coefficient of variation higher than 15%, the test was repeated (11). Absolute peak torque (Nm), i.e., the highest force output accomplished by the muscle at any moment during a repetition, was assessed, and peak torque/SLM (Nm/kg) and peak torque/LMDL (Nm/kg) were calculated for both knee flexion and extension.

Muscle strength endurance of the flexors and extensors was evaluated by isokinetic tests at an angular velocity of 240°/second and 30 repetitions, and of the lower-extremity muscles by 2 posture maintenance tests in which participants had to hold a posture as long as possible, as explained by Rombaut et al (3). For the isokinetic tests, the amount of work performed during all 30

repetitions (total work; J), the first 10 repetitions (work first third; J) and last 10 repetitions (work last third; J), and the ratio of difference between those first and last 10 repetitions, or work fatigue (%), were assessed for the knee flexors and extensors. For the posture maintenance tests, the length of time (seconds) during which a patient could maintain the correct position was recorded. Relative values (normalized for SLM and LMDL) were calculated for total work (J/kg) and work fatigue (%/kg), and for SLM for posture maintenance (seconds/kg).

Additionally, pain severity (VAS) was evaluated before and immediately after each muscle strength test and 1 minute after each muscle endurance test. Finally, physical activity and functional impairment were respectively evaluated by the Baecke questionnaire and the Arthritis Impact Measurement Scales (AIMS) (3). The mobility, walking and bending, hand, finger, and arm function subscales of the AIMS as well as the total Baecke score were used for analyses.

Statistical analysis. Data analysis was performed using the statistical package SPSS, version 24. Normality was evaluated using the Shapiro-Wilk test and Q-Q plots. Data (normally distributed) are shown as mean ± SDs, except for the AIMS questionnaire (not normally distributed), which is shown as medians and interquartile ranges. Pain scores are shown as clustered box plots with medians and interquartile ranges. Because all statistical assumptions were fulfilled, repeated-measures analysis of variance (ANOVA) was performed to identify significant differences in evolution between both groups (hEDS/HSD and controls). From a clinical point of view, age, pain, and BMI are important factors impacting muscle strength. However, due to a small sample size, only the variable with the biggest impact (BMI) was included as a covariate. Post hoc paired-sample *t*-tests with Bonferroni correction were executed when a significant interaction (time × group) effect was observed, in order to identify significant time effects within either the hEDS/HSD group or the control group.

For the AIMS questionnaire, a nonparametric Wilcoxon's test was performed to identify significant time effects, and the Mann-Whitney U test for group differences on T1 and T2 and for the difference scores of the 2 time points between hEDS/HSD and the control group, all with Bonferroni correction. *P* values less than 0.05 were considered statistically significant. Additionally, size effect estimates are shown by partial eta squared ($p\eta^2$) for repeated-measures ANOVA and by eta squared (η^2) for the Mann-Whitney U test, of which values of 0.01, 0.06, and 0.14 represent small, medium, and large size effects, respectively (12). Finally, results of the follow-up questionnaires were analyzed by descriptive statistics (frequency tables) and pain severity scores before and after the muscle strength tests by an independent samples *t*-test.

RESULTS

Characteristics. Subject characteristics at both T1 and T2 are shown in Table 1. There were no significant differences between the patient and control groups at baseline nor at follow-up, except for a significantly higher Beighton score in the patient group in comparison with the control group at T1 ($P < 0.001$). The evolution in muscle density was significantly different between the hEDS/HSD and control group (P [time × group] = 0.001, $p\eta^2 = 0.263$), with a mild decrease in muscle density in the hEDS/HSD group over time ($P < 0.001$), but not in the control group. Over 8 years, SLM increased significantly (main effect for time $P = 0.012$), with a similar evolution for both groups.

Maximal muscle strength. Maximal muscle strength results are shown in Table 2. The repeated-measures ANOVA showed a significant group effect for all maximal muscle strength variables, indicating that the hEDS/HSD group was significantly weaker than the control group at baseline and follow-up (main effect for group $P \leq 0.011$).

The evolution of most maximal muscle strength variables did not significantly differ between patients and controls, except for the peak torque of the flexors, and these normalized for LMDL (patients: P [time × group] = 0.012, $p\eta^2 = 0.139$; controls: P [time × group] = 0.045, $p\eta^2 = 0.099$) due to a significant decline over 8 years in the control group (peak torque of the flexors: $P = 0.052$;

Table 1. Characteristics of patients and controls*

Characteristic	hEDS/HSD group		Control group				
	T1	T2	T1	T2	*P* time	*P* group	*P* (time × group)
Age, years	41.3 ± 11.39	49.2 ± 11.36	40.65 ± 11.66	48.65 ± 11.78	<0.001†	0.892	0.398
Body mass index, kg/m^2	27.3 ± 6.14	28.7 ± 5.79	24.5 ± 3.99	25.9 ± 4.93	0.011†	0.087	0.981
Beighton (/9)	6.7 ± 1.65‡	4.2 ± 2.17	1.4 ± 1.62	NA	NA	NA	NA
SLM, kg	40.9 ± 6.31	43.2 ± 7.71	43.1 ± 5.71	44.8 ± 6.12	0.012†	0.340	0.679
LMDL, kg	7.3 ± 1.33	7.5 ± 1.46	7.9 ± 1.22	8.0 ± 1.15	0.131	0.211	0.446
Muscle density, mg/cm^3	76.6 ± 2.34	74.3 ± 3.25	76.6 ± 1.65	76.7 ± 1.83	0.002†	0.101	0.001†

* Values are the mean ± SD unless indicated otherwise. T1 was baseline (2009); T2 was at follow-up (2017). hEDS/HSD = hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorder patient group; LMDL = lean mass dominant leg; NA = not applicable (measurements were not performed at T2); *P* group = *P* value of comparison hEDS/HSD with control group; *P* time = *P* value for comparison of T2 with T1; *P* (time × group) = *P* value of comparison of evolution between hEDS/HSD and control group; SLM = subtotal lean mass.
† $P < 0.05$.
‡ $P < 0.05$ for hEDS/HSD group compared with control group at T1 (analyzed by an independent samples *t*-test).

Table 2. Maximal muscle strength: absolute and relative values*

Peak torque†	hEDS/HSD group		Control group					
	T1	T2	T1	T2	*P* time	*P* group	*P* (time × group)	$p\eta^2$
Extensors, Nm	85.3 ± 36.97	84.7 ± 31.12	128 ± 23.09	119.3 ± 23.39	0.986	<0.001‡	0.271	0.028
Flexors, Nm	44.2 ± 21.43	45.4 ± 15.35	65.7 ± 14.18	58.7 ± 15.97	0.068	<0.001‡	0.012‡	0.139
Extensors/SLM, Nm/kg	2.1 ± 0.83	1.9 ± 0.62	2.9 ± 0.52	2.7 ± 0.52	0.613	0.001‡	0.681	0.004
Flexors/SLM, Nm/kg	1.1 ± 0.45	1.1 ± 0.28	1.5 ± 0.26	1.3 ± 0.33	0.750	0.007‡	0.066	0.084
Extensors/LMDL, Nm/kg	11.7 ± 4.39	11.1 ± 3.54	16.0 ± 2.69	14.9 ± 2.73	0.556	0.001‡	0.761	0.002
Flexors/LMDL, Nm/kg	6.0 ± 2.25	6.1 ± 1.67	8.2 ± 1.39	7.1 ± 1.80	0.707	0.011‡	0.045‡	0.099

* Values are the mean ± SD unless indicated otherwise. T1 was baseline (2009); T2 was at follow-up (2017). hEDS/HSD = hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorder patient group; LMDL = lean mass dominant leg; *P* group = *P* value of comparison hEDS/HSD with control group; *P* time = *P* value of comparison T2 with T1; *P* (time × group) = *P* value of comparison of evolution between hEDS/HSD and control group with body mass index as covariate; $p\eta^2$ = partial eta squared (relative size effect) for comparison of evolution between hEDS/HSD and control group; SLM = subtotal lean mass.
† Angular velocity 60°/second, isokinetic test.
‡ $P < 0.05$.

peak torque of the flexors normalized for LMDL: $P = 0.028$), whereas the patient group appeared to remain stable ($P = 1.000$). Furthermore, no changes over time were observed.

Muscle strength endurance. Muscle strength endurance results are shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24220/abstract. Absolute and relative values of total work and work performed in the first and last 10 repetitions were significantly lower in the hEDS/HSD group compared to the controls at baseline and follow-up (main effect for group $P \leq 0.004$). Work fatigue, expressed as a ratio of difference between the work first third and last third (absolute and relative values), did not significantly differ between the 2 groups at T1 and T2, except for work fatigue of the flexors normalized for SLM and LMDL ($P = 0.045$ and $P = 0.028$, respectively).

A significantly different evolution of total work (absolute and relative values) and work in the first and last third performed by the extensors was identified (P [time × group] ≤ 0.022, $p\eta^2$ varying between 0.121 and 0.363) (Figure 1). Post hoc tests showed a decrease of total work and work first third performed by the extensors in the patient group ($P = 0.006$ and $P = 0.002$, respectively), in contrast to an increase of total work and work last third in the control group over 8 years ($P = 0.010$ and $P = 0.032$, respectively). No differences in evolution for the flexors were observed. Work fatigue (absolute and relative values) of the extensors and flexors showed a similar evolution and no significant changes over time except for work fatigue of the flexors, which significantly improved over time (main effect for time $P = 0.025$).

Posture maintenance tests (absolute and relative values) showed significantly lower values in the hEDS/HSD group compared to the control group over both time points (main effect for group $P < 0.001$). No significant differences in evolution or over time were found between both groups.

Pain associated with maximal muscle strength, muscle strength endurance tests. Pain severity was significantly higher at baseline and follow-up in the patient group in comparison with the control group, both before and after the strength tests ($P < 0.001$) (Figure 2).

Questionnaires. Results of the questionnaires are shown in Tables 3 and 4. Functional impairment ($P < 0.001$) and pain were significantly higher and physical activity significantly lower in the hEDS/HSD group compared to the control group at baseline and follow-up (main effect for group $P \leq 0.001$). All

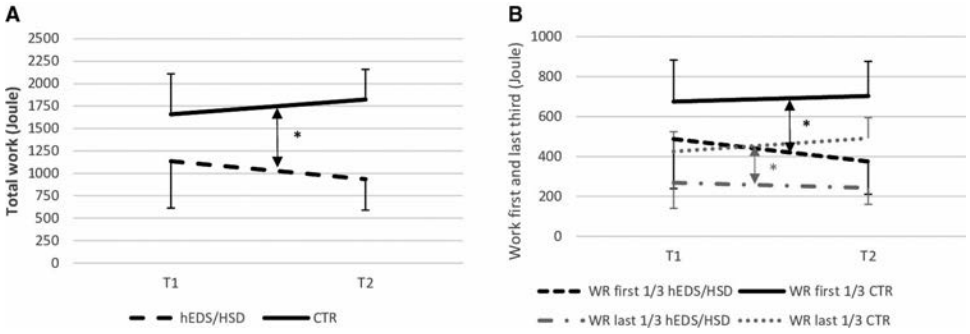

Figure 1. **A**, Evolution of total work (extensors). **B**, Evolution of work, first and last third (extensors). CTR = control group; hEDS/HSD = hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorder patient group; T1 = baseline (2009); T2 = follow-up (2017); WR first 1/3 = work performed in the first third; WR last 1/3 = work performed in the last third; * = $P < 0.05$.

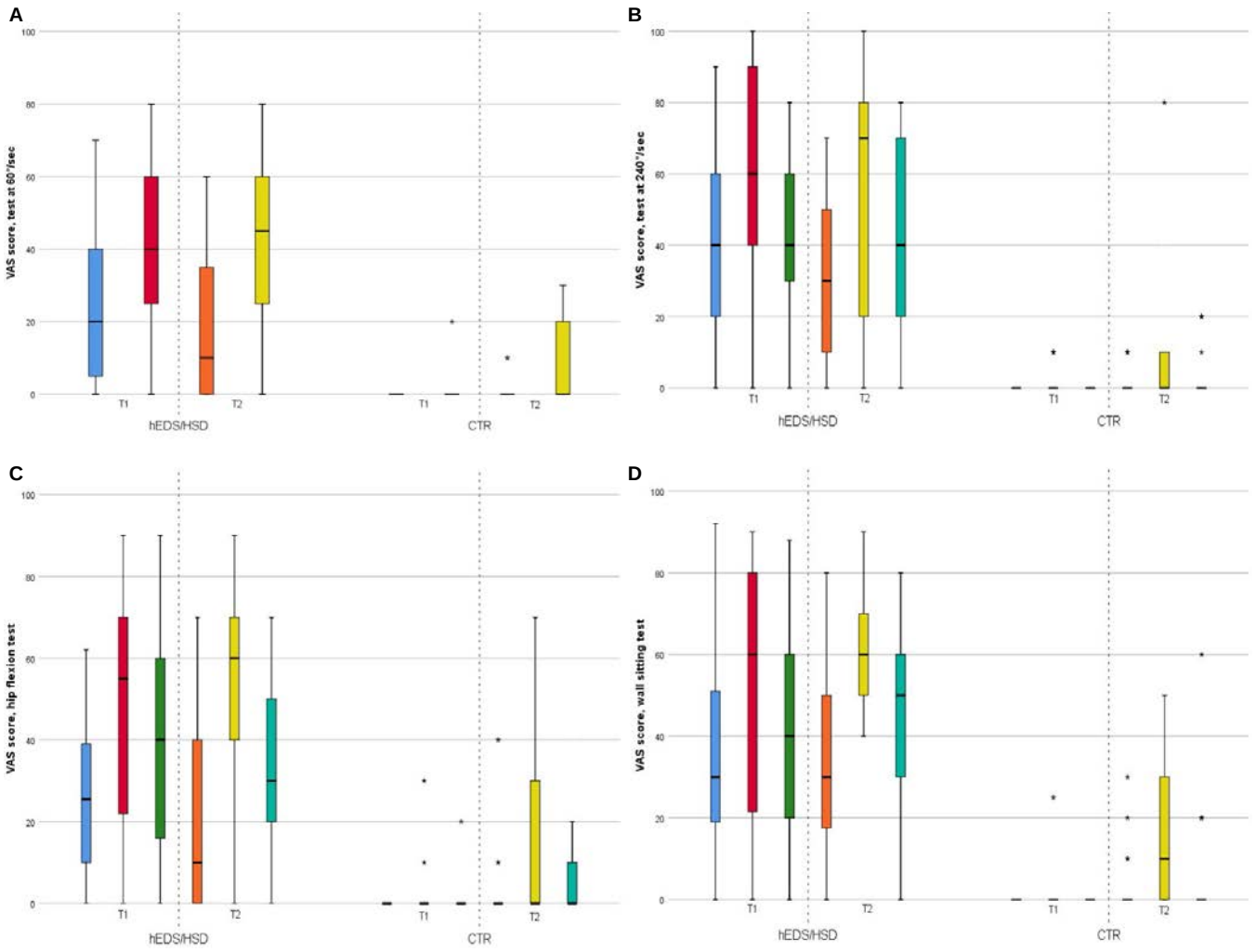

Figure 2. Pain associated with maximal muscle strength (**A**) and muscle strength endurance (**B**, **C**, and **D**) tests. Clustered box plots are shown with medians and interquartile ranges. CTR = control group; hEDS/HSD = hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorder patient group; T1 = baseline (2009); T2 = follow-up (2017); VAS = visual analog scale for pain; blue bars = just before the test at baseline; red bars = immediately after the test at baseline; dark green bars = 1 minute after the test at baseline; orange bars = just before the test at follow-up; yellow bars = immediately after the test at follow-up; light green bars = 1 minute after the test at follow-up; * = outlier.

variables showed no significantly different evolution between the 2 groups and no significant changes over time.

Results on the follow-up questionnaire showed that over 8 years, 40% of the patients reported they received physical therapy multiple times a week, 20% weekly, 13.3% monthly, and 16.7% incidentally. This finding is in contrast to the controls, of which 0% received physical therapy weekly or monthly and 64.7% incidentally. Physical therapy sessions were reported to consist mainly of stabilization exercises (hEDS/HSD 60%, controls 23.5%), manual therapy (hEDS/HSD 56.7%, controls 58.8%), muscle strength training (hEDS/HSD 53.3%, controls 17.6%), and massage (hEDS/HSD 50%, controls 0%). Additionally, 33.3% of the patients performed exercises multiple times a week, 20% weekly, and 10% monthly, whereas 5.9% of the controls exercised multiple times a week or weekly and 0% monthly. The most commonly reported sports and physical activities undertaken by the participants included walking (hEDS/HSD 60%, controls 70.6%), cycling (hEDS/HSD 46.6%, controls 76.5%), swimming (hEDS/HSD 43.3%, controls 41.2%), and aquagym/hydrotherapy (hEDS/HSD 26.7%, controls 0%). Regarding their medical history over the 8 years, compared to the control group, more of the patient group reported experiencing injuries (hEDS/HSD 56.7%, controls 47.1%) and ≥1 surgery (any type) (hEDS/HSD 56.1%, controls 35.3%). Similar proportions of each group also reported having given birth to 1 or more children over the 8 years (hEDS/HSD 13.4%, controls 11.8%).

DISCUSSION

This study has provided new insight into the evolution of muscle strength over a period of 8 years in patients with hEDS/HSD in

Table 3. Baecke and VAS questionnaire results*

	hEDS/HSD group		Control group					
	T1	T2	T1	T2	*P* time	*P* group	*P* (time × group)	$p\eta^2$
Total Baecke	6.9 ± 2.13	7.2 ± 1.24	8.2 ± 1.44	8.6 ± 1.17	0.194	0.001†	0.785	0.001
VAS (/100)	40.8 ± 20.35	45.2 ± 21.65	7.1 ± 13.12	11.2 ± 13.64	0.115	<0.001	0.836	0.000

* Values are the mean ± SD unless indicated otherwise. T1 was baseline (2009); T2 was at follow-up (2017). Higher scores on the visual analog scale (VAS) and Baecke questionnaire indicate higher pain levels and higher physical activity, respectively. hEDS/HSD = hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorder patient group; *P* group = *P* value of comparison hEDS/HSD with control group; *P* time = *P* value of comparison of T2 with T1; *P* (time × group) = *P* value of comparison of evolution between hEDS/HSD and control group; $p\eta^2$ = partial eta squared (relative size effect of Mann-Whitney U test) for comparison of evolution between hEDS/HSD and control group.
† $P < 0.05$.

comparison with controls. In general, at baseline and follow-up, maximal muscle strength and muscle strength endurance were significantly lower in patients than in controls. The main finding of this study is that the strength parameters remained relatively stable in the patient group over a period of 8 years.

Similar to the baseline results, maximal muscle strength and muscle strength endurance generally remained significantly lower at follow-up in the hEDS/HSD group in comparison with the control group. Several factors may be responsible for reduced muscle strength in hEDS/HSD (3).

As suggested by Rombaut et al, musculoskeletal pain related to joint hypermobility, subluxations, and central sensitization processes may contribute to lower muscle strength by inhibiting maximal voluntary contraction force (3,4,13,14). This finding is in accordance with our results, showing significantly higher VAS scores before and after the muscle strength tests in comparison with the control group. Future research focusing on strength generation in (asymptomatic) hypermobile individuals could further explore this impact.

Furthermore, this study identified lower habitual physical activity levels in the hEDS/HSD group in comparison with the controls, which may lead to deconditioning and decreased muscle strength (15). When results on DXA scans were compared to the controls, this study did not identify any signs of muscle atrophy in the hEDS/HSD group, which is in accordance with previous studies (1,3,4,16–18). Therefore, muscle atrophy is less likely to provide an explanation for the observed lower muscle strength in hEDS/HSD.

In addition, alterations in the structural integrity of the connective tissue in the tendons and surrounding the muscle cells could contribute to a reduced force transmission from the muscle fibers onto the skeleton, eventually leading to an altered muscle function and reduced muscle strength (3,4,19,20). Poor proprioception, associated with generalized joint hypermobility, may be related to reduced muscle strength as well (19).

The current study did not demonstrate any significant changes in maximal muscle strength over time in hEDS/HSD patients. However, a decrease of maximal muscle strength generated by the knee flexors was identified in the control group. Though this decrease could be explained by an age-related deterioration of muscle function, this decline appears to be absent in the hEDS/HSD group (21–23). The high engagement with physical therapy and exercise in the hEDS/HSD group could give a likely explanation, because these are major contributors to the maintenance of muscle strength and mass over time and are prescribed as an important aspect of the multidisciplinary treatment of the pathology (24,25). Furthermore, more than half of the physical therapy consultations consisted of muscle strengthening exercises in the patient group, in contrast to 18% in the control group. Although no information is available about exact methods of this strength training and accomplished strength enhancements, these findings only suggest that physical therapy and exercise could have a positive impact in preventing further deterioration of maximal muscle strength in hEDS/HSD.

In general, this study identified no differences over time in static muscle strength endurance and muscle strength endurance of the knee flexors in both groups. However, decreased muscle strength endurance of the knee extensors in the hEDS/HSD group was observed, which is in contrast to the increase in the control group over a period of 8 years.

The evolution of muscle strength endurance over time in the control group could be attributed to age-related changes

Table 4. Arthritis Impact Measurement Scales (AIMS) questionnaire results*

	hEDS/HSD group			Control group				
AIMS	T1	T2	*P* time	T1	T2	*P* time	*P* diff	η^2
Movement abilities	3.0 (1.5–5.0)	3.0 (1.5–4.1)	1.000	0.0 (0.0–1.3)	0.0 (0.0–0.0)	0.176	1.000	0.009
Walking and bending	6.0 (5.0–8.0)	6.0 (5.0–7.8)	0.778	0.0 (0.0–1.0)	0.0 (0.0–2.0)	0.678	0.308	0.045
Hand and finger function	4.5 (2.0–6.0)	3.2 (2.0–4.5)	0.644	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000	0.746	0.017
Arm function	2.5 (0.5–3.6)	1.5 (0.9–2.5)	0.064	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.634	0.332	0.041

* Values are the median (Q1–Q3) unless indicated otherwise. T1 was baseline (2009); T2 was at follow-up (2017). Higher scores on the AIMS indicate higher functional impairment. hEDS/HSD = hypermobile Ehlers-Danlos syndrome/hypermobility spectrum disorder patient group; *P* diff = *P* value of comparison of difference scores of T1 and T2 between hEDS/HSD and control group; *P* time = *P* value of comparison of T2 with T1; η^2 = eta squared (effect size of Mann-Whitney U test).

in muscle fiber type. Findings about changes of type I muscle fibers, mainly used during daily living and aerobic endurance activities, are inconclusive but range from a higher type I to type II fiber ratio to nonaffected type I fibers during the aging process (25–27). These changes in muscle fiber type eventually could result in an increase or stabilization of muscle strength endurance, as shown in the control group of this study.

By contrast, in the patient group, muscle strength endurance of the knee extensors significantly decreased over the period of 8 years (total work and work in the first one-third of the isokinetic test). This decrease might be explained by the content of the physical therapy sessions in hEDS/HSD patients participating in this study, which mainly focused on joint stability exercises (60%) and manual therapy (56.7%) over the past 8 years rather than improving muscle strength endurance, which is in accordance with research identifying smaller distances performed in the 6-minute walking test in comparison with healthy controls (19). Furthermore, the implementation of cognitive-behavioral treatment, including coping strategies in the multidisciplinary treatment program for hEDS/HSD could provide a likely explanation of why results on the strength endurance tests were lower (28–30). It could be hypothesized that the decrease of total work over a period of 8 years in the hEDS/HSD group can be attributed to coping strategies learned in physical therapy sessions to avoid maximal load on the muscles and unstable joints. Decreased work in the first third might reflects the fact that patients with hEDS/HSD possibly try to spread the load over time to be able to perform the entire test, consisting of 30 repetitions. However, cognitive-behavioral treatment was not evaluated in this study, and therefore this hypothesis remains purely speculative.

In addition to the fact that pain and fatigue frequently occur in hEDS/HSD, reduced muscle strength is a major contributor to functional impairment and has a considerable impact on the daily living activities of these patients (3,14,19,31). For instance, because hamstring muscles play a major role in power transfer in the lower extremity, muscle weakness could lead to altered motor control and propulsion in these patients (19,32). Therefore, treatment focusing on pain relief, joint stabilization exercises, coping strategies, and muscle strengthening exercises could be recommended to improve quality of life and reduce disability (9,19,33).

Physical therapy plays a key role within the multidisciplinary team in the management of this patient population (33). A pilot study performed by Bathen et al (29) showed positive effects on daily functioning, kinesiophobia, and both muscle strength and endurance after an intensive training program with learning coping strategies. Bathen el al suggest that improving muscle function in hEDS/HSD is possible (29). Furthermore, endurance training to improve cardiovascular, physical, and musculoskeletal fitness should be included in the training program, as we observed a longitudinally decreased muscle strength endurance in hEDS/HSD (33). However, evidence-based clinical trials evaluating intervention programs are scarce. Therefore, further research should determine whether or not exercise is effective in this patient group and which types of exercises should be recommended, specifically to improve maximal muscle strength and muscle strength endurance in hEDS/HSD.

This was the first longitudinal study evaluating muscle strength parameters in patients with hEDS/HSD over 8 years, with a high follow-up rate of the patients with hEDS/HSD. Along with objective measurements of muscle strength, this study also retrospectively questioned several muscle strength–related parameters. However, the results of the study should be viewed within its limitations. First, the use of this self-developed follow-up questionnaire could create a bias due to the dependence on the patient's ability to recall the information. Second, patients were only measured twice over a period of 8 years. Future studies should systematically evaluate medical parameters, muscle strength, and muscle mass, preferably each year to better evaluate the evolution in this patient population. Third, a low follow-up rate of controls was achieved. However, no significant differences in outcomes (measured at T1) were observed between the dropouts and the follow-up participants, either in the patient group or the control group. Though the power of this study is decreased by these dropouts, especially in the control group, the impact of this low follow-up rate is therefore probably limited. Finally, results cannot be generalized to the upper extremity, because only the lower extremity was evaluated.

In conclusion, this follow-up study showed at baseline and follow-up a significant reduced muscle strength and muscle strength endurance in hEDS/HSD patients compared to the controls, of which the underlying causes are possibly multifactorial. With regard to the evolution, the majority of the strength parameters remained relatively stable in the patient group over a period of 8 years. Future studies should focus on both effectiveness and efficiency of different types of muscle training strategies and their effect on pain and functioning in hEDS/HSD patients.

ACKNOWLEDGMENTS

The authors thank the patients with hEDS/HSD who participated in this study. They also thank the Department of Endocrinology for their cooperation in this study.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Coussens, Calders, De Wandele, Pacey, Malfait, Rombaut.

Acquisition of data. Coussens, Banica, Rombaut.

Analysis and interpretation of data. Coussens, Calders, Lapauw, Celie, Banica, De Wandele, Pacey, Malfait, Rombaut.

REFERENCES

1. Beighton P, De Paepe A, Steinmann B, Tsipouras P, Wenstrup RJ. Ehlers–Danlos syndromes: revised nosology, Villefranche, 1997.

Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). Am J Med Genet 1998;77:31–7.
2. Malfait F, Francomano C, Byers P, Belmont J, Berglund B, Black J, et al. The 2017 international classification of the Ehlers-Danlos syndromes. Am J Med Genet C Semin Med Genet 2017;175:8–26.
3. Rombaut L, Malfait F, De Wandele I, Taes Y, Thijs Y, De Paepe A, et al. Muscle mass, muscle strength, functional performance, and physical impairment in women with the hypermobility type of Ehlers-Danlos syndrome. Arthritis Care Res (Hoboken) 2012;64:1584–92.
4. Voermans NC, van Alfen N, Pillen S, Lammens M, Schalkwijk J, Zwarts MJ, et al. Neuromuscular involvement in various types of Ehlers-Danlos syndrome. Ann Neurol 2009;65:687–97.
5. Rombaut L, Malfait F, Cools A, De Paepe A, Calders P. Musculoskeletal complaints, physical activity and health-related quality of life among patients with the Ehlers-Danlos syndrome hypermobility type. Disabil Rehabil 2010;32:1339–45.
6. Rombaut L, Malfait F, De Wandele I, Mahieu N, Thijs Y, Segers P, et al. Muscle–tendon tissue properties in the hypermobility type of Ehlers-Danlos syndrome. Arthritis Care Res (Hoboken) 2012;64:766–72.
7. Castori M, Camerota F, Celletti C, Danese C, Santilli V, Saraceni VM, et al. Natural history and manifestations of the hypermobility type Ehlers-Danlos syndrome: a pilot study on 21 patients. Am J Med Genet A 2010;152a:556–64.
8. Castori M, Morlino S, Celletti C, Ghibellini G, Bruschini M, Grammatico P, et al. Re-writing the natural history of pain and related symptoms in the joint hypermobility syndrome/Ehlers-Danlos syndrome, hypermobility type. Am J Med Genet A 2013;161a:2989–3004.
9. Tinkle B, Castori M, Berglund B, Cohen H, Grahame R, Kazkaz H, 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;175:48–69.
10. Goodpaster BH, Thaete FL, Kelley DE. Thigh adipose tissue distribution is associated with insulin resistance in obesity and in type 2 diabetes mellitus. Am J Clin Nutr 2000;71:885–92.
11. Biodex Medical Systems. Isokinetic testing and data interpretation: data analysis. URL: http://www.biodex.com/sites/default/files/manual-clinical-resources-data.pdf.
12. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 2012;141:2–18.
13. Castori M, Voermans NC. Neurological manifestations of Ehlers-Danlos syndrome(s): a review. Iran J Neurol 2014;13:190–208.
14. Voermans NC, Knoop H, Bleijenberg G, van Engelen BG. Pain in Ehlers-Danlos syndrome is common, severe, and associated with functional impairment. J Pain Symptom Manage 2010;40:370–8.
15. Chacon-Cabrera A, Lund-Palau H, Gea J, Barreiro E. Time-course of muscle mass loss, damage, and proteolysis in gastrocnemius following unloading and reloading: implications in chronic diseases. PloS One 2016;11:e0164951.
16. Pretorius ME, Butler IJ. Neurologic manifestations of Ehlers-Danlos syndrome. Neurology 1983;33:1087–9.
17. Banerjee G, Agarwal RK, Shembesh NM, el Mauhoub M. Ehlers Danlos syndrome: masquerading as primary muscle disease. Postgrad Med J 1988;64:126–7.
18. Gerrits KH, Voermans NC, de Haan A, van Engelen BG. Neuromuscular properties of the thigh muscles in patients with Ehlers-Danlos syndrome. Muscle Nerve 2013;47:96–104.
19. Scheper M, Rombaut L, de Vries J, De Wandele I, van der Esch M, Visser B, et al. The association between muscle strength and activity limitations in patients with the hypermobility type of Ehlers-Danlos syndrome: the impact of proprioception. Disabil Rehabil 2017;39:1391–7.
20. Bilkey WJ, Baxter TL, Kottke FJ, Mundale MO. Muscle formation in Ehlers-Danlos syndrome. Arch Phys Med Rehabil 1981; 62:444–8.
21. Aoyagi Y, Shephard RJ. Aging and muscle function. Sports Med 1992;14:376–96.
22. Deschenes MR. Effects of aging on muscle fibre type and size. Sports Med 2004;34:809–24.
23. Carter HN, Chen CC, Hood DA. Mitochondria, muscle health, and exercise with advancing age. Physiology (Bethesda) 2015;30: 208–23.
24. Curtis E, Litwic A, Cooper C, Dennison E. Determinants of muscle and bone aging. J Cell Physiol 2015;230:2618–25.
25. Gomes MJ, Martinez PF, Pagan LU, Damatto RL, Cezar MD, Lima AR, et al. Skeletal muscle aging: influence of oxidative stress and physical exercise. Oncotarget 2017;8:20428–40.
26. Fielding RA, Vellas B, Evans WJ, Bhasin S, Morley JE, Newman AB, et al. Sarcopenia: an undiagnosed condition in older adults. Current consensus definition: prevalence, etiology, and consequences. International working group on sarcopenia. J Am Med Dir Assoc 2011;12:249–56.
27. Talbot J, Maves L. Skeletal muscle fiber type: using insights from muscle developmental biology to dissect targets for susceptibility and resistance to muscle disease. Wiley Interdiscip Rev Dev Biol 2016;5:518–34.
28. Talotta R, Bazzichi L, Di Franco M, Casale R, Batticciotto A, Gerardi MC, et al. One year in review 2017: fibromyalgia. Clin Exp Rheumatol 2017;35 Suppl 105:6–12.
29. Bathen T, Hangmann AB, Hoff M, Andersen LO, Rand-Hendriksen S. Multidisciplinary treatment of disability in Ehlers-Danlos syndrome hypermobility type/hypermobility syndrome: a pilot study using a combination of physical and cognitive-behavioral therapy on 12 women. Am J Med Genet A 2013;161a:3005–11.
30. Zhou Z, Rewari A, Shanthanna H. Management of chronic pain in Ehlers-Danlos syndrome: two case reports and a review of literature. Medicine (Baltimore) 2018;97:e13115.
31. Voermans NC, Knoop H, Bleijenberg G, van Engelen BG. Fatigue is associated with muscle weakness in Ehlers-Danlos syndrome: an explorative study. Physiotherapy 2011;97:170–4.
32. Morin JB, Gimenez P, Edouard P, Arnal P, Jimenez-Reyes P, Samozino P, et al. Sprint acceleration mechanics: the major role of hamstrings in horizontal force production. Front Physiol 2015 ;6:404.
33. Engelbert RH, Juul-Kristensen B, Pacey V, de Wandele I, Smeenk S, Woinarosky N, et al. The evidence-based rationale for physical therapy treatment of children, adolescents, and adults diagnosed with joint hypermobility syndrome/hypermobile Ehlers Danlos syndrome. Am J Med Genet C Semin Med Genet 2017;175:158–67.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 1049–1054
DOI 10.1002/acr.24205

American College of Rheumatology
Empowering Rheumatology Professionals

Mortality in Patients With Gout Treated With Allopurinol: A Systematic Review and Meta-Analysis

Charles A. Hay, James A. Prior, John Belcher, Christian D. Mallen, and Edward Roddy

Objective. Urate-lowering therapy (predominantly allopurinol) is highly effective as a treatment for gout, but its wider long-term effects remain unclear. This systematic review and meta-analysis aimed to ascertain the association between mortality and the use of allopurinol in patients with gout.

Method. Medline, Embase, CINAHL, and the Cochrane Library were searched from inception to August 2018. Articles eligible for inclusion used a cohort design and examined cardiovascular or all-cause mortality in patients diagnosed with gout and prescribed allopurinol. Information on study characteristics, design, sample size, and mortality risk estimates were extracted. Article quality was assessed using the Newcastle-Ottawa Scale. Included articles were described in a narrative synthesis and, where possible, risk estimate data were pooled.

Results. Four articles reported a hazard ratio (HR) risk estimate for all-cause mortality in patients with gout using allopurinol, and 2 of these also reported cardiovascular mortality. Two articles found allopurinol to be protective in patients with gout, 1 found no statistically significant association, and 1 found no statistically significant effect of escalation of allopurinol dosage on all-cause or cardiovascular-related mortality. Data pooling was possible for all-cause mortality and found no association between allopurinol use in patients with gout and all-cause mortality compared to patients with gout not using allopurinol (adjusted HR 0.80 [95% confidence interval 0.60–1.05]).

Conclusion. There was no significant association between all-cause mortality and allopurinol use in people with gout. However, the number of included studies was small, suggesting that further studies are needed.

INTRODUCTION

Gout is the most common inflammatory arthritis, affecting 2.5% of UK adults (1). Its pathogenesis is well understood: elevation of serum urate levels above 360 μmol/liter (6 mg/dl) can lead to formation and deposition of monosodium urate crystals in joints and soft tissues that can result in painful acute flares of joint inflammation (2). Without treatment, flare frequency increases, chronic joint damage occurs, and mobility/function decrease, resulting in impaired health-related quality of life (3). There is also an increased risk of serious comorbidities (e.g., cardiovascular disease) and premature mortality (4,5).

Treating gout should be straightforward due to the availability of safe, effective, long-term treatment to lower urate levels (urate-lowering therapy [ULT]), allowing dissolution of existing urate crystals and prevention of new crystal formation, leading to the cessation of gout flares (6,7). International guidelines recommend that ULT be offered to all patients with gout and initiated upon confirmation of diagnosis, once any current flare has abated (8,9). Allopurinol is the first-line ULT and should be initiated at a low dose (≤100 mg daily), followed by uptitration in 100-mg increments until urate levels are suppressed below 360 μmol/liter (6 mg/dl). Despite clear guidelines and benefits, only 30% of patients are prescribed allopurinol, and of those, only 40% have treatment escalated to achieve the target serum urate level of <360 μmol/liter, suggesting that many patients with gout could receive better ULT (10).

In addition to its success in treating gout, other benefits of allopurinol have been suggested in patients with kidney and cardiovascular diseases. The drug has been shown to be associated with a decreased likelihood of renal events (initiation of dialysis, doubling serum creatinine, ≥50% decrease in estimated glomerular filtration rate) in two-thirds of patients with chronic kidney disease (11). Improvements in cardiovascular function include increased peripheral blood-flow due to improved endothelial

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

Supported by the Collaboration for Leadership in Applied Health Research and Care Oxford West Midlands. Mr. Hay's work was supported by an NIHR School for Primary Care Research PhD Studentship. Dr. Mallen's work was supported by the NIHR Collaboration for Leadership in Applied Health Research and Care West Midlands, the NIHR School for Primary Care Research, and an NIHR Research Professorship in General Practice (NIHR-RP-2014-04-026).

Charles A. Hay, MA, James A. Prior, PhD, John Belcher, PhD, Christian D. Mallen, PhD, Edward Roddy, PhD: Keele University, Newcastle, UK.

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

Address correspondence to Charles A. Hay, MA, School of Medicine, Keele University, Staffordshire, ST5 5BG, UK. Email: c.hay@keele.ac.uk.

Submitted for publication September 6, 2019; accepted in revised form March 31, 2020.

SIGNIFICANCE & INNOVATIONS

- We found no significant association between all-cause mortality and allopurinol use in people with gout.
- Further studies taking into account allopurinol dose and achievement of target serum urate levels are required.

function in patients with chronic heart failure (12). However, despite these improvements in morbidity, whether the benefits of allopurinol extend to reducing premature mortality in patients with gout remains unclear. In patients with hyperuricemia (the precursor to gout), the use of allopurinol has been estimated to be associated with a 25% lower risk of mortality during follow-up compared with untreated patients (13,14).

Despite guidelines recommending earlier prescription of ULT (8,9) and the reported benefits on comorbidities, the use of allopurinol to treat gout remains suboptimal. Though the reasons behind this hesitancy are multifaceted, 1 contributing factor relates to the apprehension of patients and clinicians to initiate life-long treatment without a clear understanding of the long-term effects (1). Because the overall balance of potential benefit and risk in the role of allopurinol on mortality in patients with gout remains unclear, this systematic review and meta-analysis examined the association between the use of allopurinol in patients with gout and cardiovascular or all-cause mortality.

MATERIALS AND METHODS

A systematic review of research literature was conducted. Medical literature databases were searched to identify articles that included patients with gout treated using allopurinol and that reported the risk of cardiovascular or all-cause mortality in their sample. Meta-analysis was used to determine pooled risk estimates of mortality. The protocol for this systematic review and meta-analysis was registered on PROSPERO (ID CRD42017056011) and the systematic review was undertaken following Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

Data sources, searches, and study selection. Four electronic bibliometric databases were searched for articles (Embase, Medline, CINAHL, and Cochrane Studies) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24205/abstract). These were required to fulfil the following eligibility criteria: 1) the study sample was formed from adults with a diagnosis of gout, 2) allopurinol was used to treat gout, 3) risk estimates of all-cause or cardiovascular mortality were reported, and 4) the study used a cohort design. Cohort studies were targeted specifically because their populations are more likely than randomized controlled trials (RCTs) to be representative of the general population and of normal courses of treatment, therefore increasing the likelihood that this systematic review and meta-analysis produces a generalizable result. Case–control and cross-sectional studies were excluded because they would not describe outcomes over time. No restrictions were imposed on the time periods for publication, with medical literature databases searched from inception to August 2018. There were no language restrictions, but if translational facilities were not available for an article, it was excluded.

Data extraction. Data were extracted by 2 authors (CAH and JAP) with the main data including demographic information (age, sex, country of origin, etc.), study sample size, numbers of patients with gout, study setting (e.g., primary care), exposures (e.g., allopurinol), mortality outcome (e.g., all-cause, cardiovascular), definition of gout, and method of adjusted risk estimates regarding the association between gout treated with allopurinol and cardiovascular and all-cause mortality risk estimates.

Quality assessment. All articles finally included in the systematic review were quality appraised independently by 2 assessors (CAH and JAP). Any disagreement on initial scoring was discussed, and if the difference could not be agreed on, the decision was arbitrated by a third reviewer (ER). Methodologic quality was assessed using the Newcastle-Ottawa Scale for cohort studies (15).

Meta-analysis. Where a sufficient number of articles (≥3) were identified, a random-effects meta-analysis was used to pool reported mortality risk estimates along with their 95% confidence intervals (95% CIs). Heterogeneity was assessed by I^2. The meta-analysis was undertaken in Stata software, version 14.

RESULTS

Literature search. From 362 articles identified by the initial literature search, 90 duplicates were removed. The titles of the remaining 272 articles were screened, after which 37 articles remained. After an abstract review of these, 32 articles were excluded. The full text of the remaining 5 articles was reviewed in full, and a final 4 articles were deemed to fulfil the inclusion criteria (Figure 1).

Sample characteristics. Four articles examined all-cause mortality (16–19) and 2 of these (16,19) also examined cardiovascular mortality in the same population. Of the articles included in the review, 1 study population was from Taiwan, 1 was from the US, and the other 2 were from the UK. All 4 articles estimated the risk of mortality using HRs (16–19) (Table 1).

The Taiwanese study by Chen et al (16) sourced its cohort from the medical insurance data from MJ Health Screening

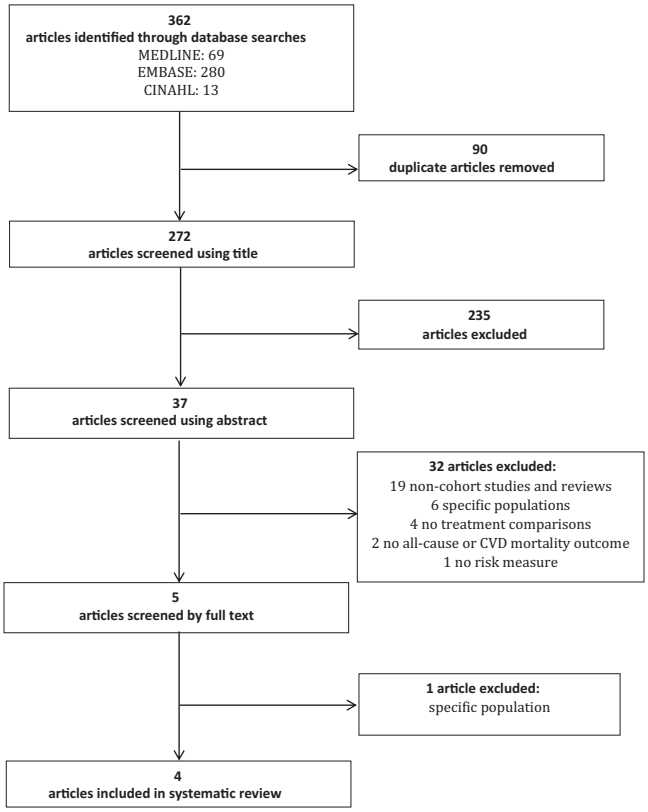

Figure 1. Flow diagram of the number of articles at each stage of the search and screening process. CVD = cardiovascular disease.

Center, which contained 49,460 individuals age >17 years who had consultations since 1996. Gout was defined using International Classification of Diseases, Ninth Revision (ICD-9) codes for cases identified between 1997 and 2002. Kuo et al (18) used a UK primary care data source, the Clinical Practice Research Datalink, and defined incident gout by Read codes between 1995 and 1999. Dubreuil et al (17) used a different UK primary care data source (The Health Improvement Network), defining gout by Read codes between January 2000 to May 2010. The study by Coburn et al (19) sourced its cohort from the US Department of Veterans Affairs Health Administration between 2001 and 2008 and defined gout by its ICD-9 definition. Unlike the previous 3 studies, Coburn et al focused specifically on the effect on risk of all-cause and cardiovascular mortality of increasing allopurinol dosage in patients.

All articles were cohort studies and used matching based on propensity scores. Chen et al and Dubreuil et al followed up their patients from exposure onward (date of diagnosis for Chen et al and initiation of allopurinol for Dubreuil et al). Kuo et al and Coburn et al both used landmark analysis to avoid immortal time bias. Kuo et al only included patients who were alive by the landmark time points (1 year and 3 years); this method excludes the initial time period immediately after gout diagnosis, reducing the possibility of conferring an unfair survival advantage on the allopurinol-treated group. Coburn et al used 2 models; in model 1, they followed up patients from exposure, and in model 2, they followed up patients after a 2-year landmark.

Quality assessment results. All 4 articles included representative patients with gout and assessed exposures and outcomes using secure methods (medical records), and employed appropriate methods to compare subjects, with and without gout, to avoid confounding by indication affecting the veracity of results. All 4 studies also employed propensity score matching. In addition, 3 of the 4 articles used methods to attempt to negate immortal time bias; Chen et al used time-index matching between patients and controls and Coburn et al used an analytical method that involved only following up patients who were alive 2 years after allopurinol initiation. Kuo et al employed a landmark analysis method that only followed up patients who were alive after 1 year post-allopurinol initiation

Table 1. The characteristics, demographics, and risk values of the study sample used in each included article*

Author, year (ref.)	Population	Study period	Sample size	Incident gout, no.	Male, %	Mean ± SD age, years	Adjusted HR (95% CI)
Chen et al, 2015 (16)	MJ Health Screening Center database, Taiwan	1997–2002	1,457	286	89	52.7 ± 15.4	0.39 (0.22–0.70)
Coburn et al, 2018 (19)	US Department of Veterans Affairs Health Administration	1999–2010	111,694	6,428†	99.7	64.4 ± 10.5	1.05 (0.96–1.15)‡
Dubreuil et al, 2015 (17)	The Health Improvement Network, UK	2000–2010	9,590	483	69	67§	0.81 (0.70–0.92)
Kuo et al, 2015 (18)	Clinical Practice Research Datalink, UK	1995–2013	19,549	3519	72	64 (52–73)¶	0.99 (0.87–1.12)# 1.01 (0.92–1.09)**

* 95% CI = 95% confidence interval; HR = hazard ratio; ref. = reference.
† Gout patients receiving dose escalation.
‡ This HR represents the risk of all-cause mortality in patients with gout treated with escalating doses of allopurinol compared to patients with gout on a constant dosage of allopurinol.
§ No SD was reported.
¶ Median (interquartile range).
# 1-year landmark analysis.
** 3-year landmark analysis.

and then 3 years after initiation. Loss to follow-up was minimal and was accounted for in analyses.

Risk of all-cause mortality. Chen et al and Dubreuil et al both found allopurinol to have a protective effect on all-cause mortality in patients with gout. Chen et al reported an adjusted HR of 0.39 (95% CI 0.22–0.70) (allopurinol was slightly more protective against all-cause mortality than the use of any ULT medication, with HR 0.47 [95% CI 0.29–0.79]), and Dubreuil et al reported an adjusted HR of 0.81 (95% CI 0.70–0.92). However, Kuo et al found no association between the use of allopurinol in patients with gout and all-cause mortality, with HR 0.99 (95% CI 0.87–1.12) for 1-year landmark analysis and HR 1.01 (95% CI 0.92–1.09) for the 3-year landmark analysis, the latter of which was included in the pooled analysis. Finally, Coburn et al reported an HR for all-cause mortality in patients with gout for whom allopurinol dosage was increased, compared with patients with gout using a constant dose. The researchers reported a significant increase in all-cause mortality for model 1 (propensity score matching HR 1.08 [95% CI 1.01–1.17]) and a nonsignificant HR for model 2 (inclusion of 2-year landmark analysis HR 1.05 [95% CI 0.96-1.15]). However, because these HRs were based on stratification by dose, their inclusion in the pooled analysis was not possible. The pooled adjusted HR for all-cause mortality calculated from the 3 applicable cohorts was 0.80 (95% CI 0.60–1.05), and heterogeneity was statistically significant (87.6%; $P < 0.001$) (Figure 2).

Risk of cardiovascular mortality. Chen et al reported a protective effect of allopurinol on cardiovascular mortality, finding an HR in patients with gout treated with allopurinol of 0.37 (95% CI 0.01–0.48) compared to non-allopurinol users. Coburn et al initially reported an association between increased cardiovascular-related mortality in those with escalated allopurinol dose compared to those with a stable dose for model 1 (HR 1.08 [95% CI 0.97–1.21]), but no association remained in model 2 (HR 1.05 [95% CI 0.92–1.20]). Due to the sparsity of data related to cardiovascular mortality, we were unable to conduct pooled analysis for this outcome (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24205/abstract).

DISCUSSION

Our systematic review and meta-analysis of 3 studies showed no significant association between the use of allopurinol and all-cause mortality in patients with gout. The results of studies into cardiovascular mortality were contradictory and limited (preventing data pooling).

Our findings are not consistent with reports of statistically significantly decreased mortality associated with allopurinol use in hyperuricemic patients and a protective effect against cardiovascular and chronic kidney disease in patients with gout (11–14,20). Though not directly comparable, such findings supported our initial hypothesis that a reduction in mortality for patients with gout using allopurinol would be observed. There are, however, some important differences between the studies in our meta-analysis and those that have previously shown protective effects of allopurinol. Notably, the studies included in our review used observational data from clinical practice, where allopurinol dosage is commonly insufficient to lower urate significantly (only 40% having treatment escalated to achieve the target serum urate level [10]).

Studies that have shown a protective effect of allopurinol dosage on the risk of cardiovascular events often involves dosage of >600 mg/day, compared to the more common 100–300 mg/day

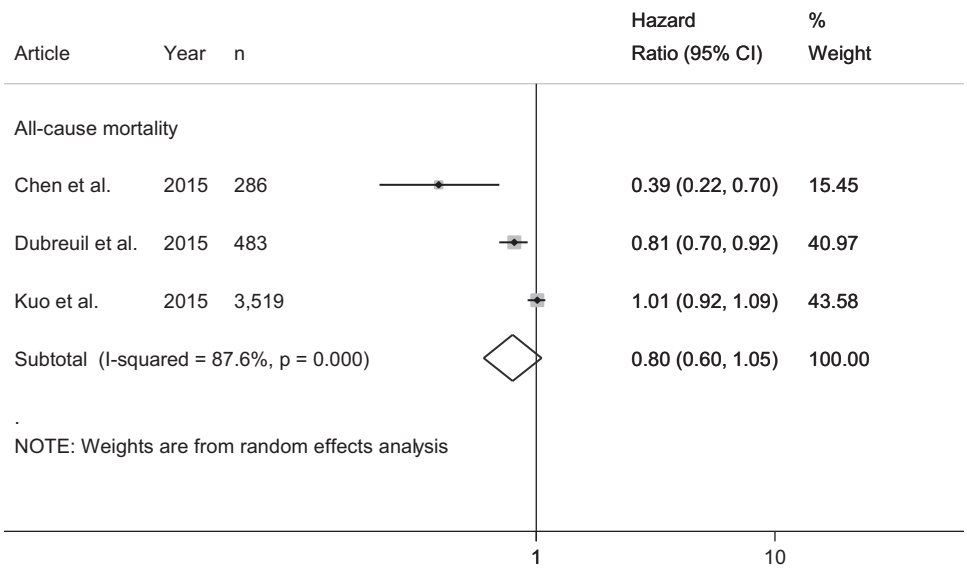

Figure 2. Random effects meta-analysis of the hazard ratio.

found throughout normal primary care gout management. Also in the case of RCTs (11), dosage of ULT was managed, observed, and escalated in a more systematic way than in the cohort studies included in our review. Possibly, therefore, the nonsignificant protective effect reported by our meta-analysis is related to the fact that lower dosages infrequently facilitate the achievement of target serum urate levels in patients, and there are frequently lower levels of compliance and treatment observation in the general population compared to RCT populations.

Our findings support the existing body of evidence on the short-term safety of allopurinol (21,22), because our included articles used large, nationally representative data sets and provided a combined sample of >10,000 patients with gout in which to examine all-cause mortality. In particular in the UK, where the majority of patients with gout are managed in primary care, Kuo et al and Dubreuil et al formed the principal weighting within the meta-analysis, with data from 2 different primary care data sets. A key methodologic difference between the studies is the use of landmark analysis by Kuo et al to address the potential for immortal time bias, and this methodologic difference may well be the cause of the disagreement between the 2 studies regarding risk. Though Chen et al demonstrated a protective effect of allopurinol use, their sample was small, and they did not include landmark analysis. However, they attempted to avoid immortal time bias by matching for the index date of ULT prescription using a propensity score (16). Possibly the difference in reported effects between the study of Chen et al and the other 3 studies in this systematic review is due to the difference in populations.

The pooled HR of 0.80 with its 95% CI of 0.60 to 1.05 could suggest a possible small protective effect of allopurinol; however, statistical significance was not reached, and the 2 largest of the 3 included studies contributed the greatest weighting in the meta-analysis and had HRs closest to 1. Further large studies into the effect of allopurinol on both all-cause and cause-specific mortality in patients with gout are needed. Our findings are complicated by the results of Coburn et al, which showed an increase in the risk of all-cause mortality in patients with gout whose dosage was escalated, although these associations became nonsignificant upon closer matching of patients with dose escalation to patients without dose escalation.

Given the protective effects of allopurinol found in RCTs and several cohort studies, further research in this area to produce a more cohesive and conclusive view of the association between treatment of gout with allopurinol and mortality is essential. Consideration should be given to the effect of allopurinol on mortality in specific subgroups, such as men and women and those with different comorbidities or tophaceous gout. Also of high importance in this research would be the effects of treatment adherence, because this adherence is so low in patients with gout that it may be undermining not just the primary aims for allopurinol, but also possible secondary positive outcomes, such as a lower risk of early all-cause mortality.

We are unable to draw any conclusions on any potential role of allopurinol use in cardiovascular mortality in patients with gout. Only 2 articles were identified and results were varied, so further research is required. However, from 1 of these articles (19), the consideration of allopurinol dose arises as an important issue in the matter of the role of allopurinol on mortality in patients with gout. Coburn et al found no significant difference between either all-cause or cardiovascular mortality in those patients who had their dose of allopurinol escalated over 2 years and those whose dose remained stable. To address the fact that the majority of patients with gout using allopurinol never reach target serum urate levels, the researchers performed a sensitivity analysis using only those patients who did reach the guideline target levels. Within this subsample, they found that for all-cause mortality there remained a similar HR (not reported); however, for cardiovascular mortality, though not significant, they now found a reduction in risk of 7% (HR 0.93 [95% CI 0.76–1.14]). The role of allopurinol and its dose on the risk of premature mortality (particularly cardiovascular) in patients with gout using allopurinol requires much further study.

This is the first systematic review to examine the association between patients with gout treated with allopurinol and cardiovascular or all-cause mortality. Our search criteria were extensive (not limited by language) and included cohort studies from large, nationally representative samples using data over similar time periods to provide a more generalizable picture of the role of allopurinol on mortality in patients with gout. Risk estimates for all-cause mortality from different studies were pooled. The principal limitations of our review are the small number of articles available and statistical heterogeneity in the pooled analysis. However, despite the low number of studies, those included in this systematic review are of high methodologic quality, having factored in methods for avoiding immortal time bias and confounding by indication.

Our systematic review and meta-analysis did not find a significant association between allopurinol use and cardiovascular or all-cause mortality. However, the small number of studies suitable for inclusion and the evidence from the wider literature that allopurinol may have cardiovascular and renal benefits suggest that further studies into the effect of allopurinol use on mortality in people with gout are required, particularly regarding the role of allopurinol dose and the importance of reaching target serum urate levels.

ACKNOWLEDGMENT

The authors thank the School of Primary, Community and Social Care, Keele University for its support.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Hay, Prior, Roddy.
Acquisition of data. Hay, Prior, Roddy.
Analysis and interpretation of data. Hay, Prior, Belcher, Mallen, Roddy.

REFERENCES

1. Kuo CF, Grainge MJ, Mallen C, Zhang W, Doherty M. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. Ann Rheum Dis 2015;74:661–7.
2. Kuo CF, Grainge MJ, Zhang W, Doherty M. Global epidemiology of gout: prevalence, incidence and risk factors. Nat Rev Rheumatol 2015;11:649–2.
3. Chandratre P, Roddy E, Clarson L, Richardson J, Hider SL, Mallen CD. Health-related quality of life in gout: a systematic review. Rheumatology (Oxford) 2013;52:2031–40.
4. Clarson LE, Hider SL, Belcher J, Heneghan C, Roddy E, Mallen CD. Increased risk of vascular disease associated with gout: a retrospective, matched cohort study in the UK clinical practice research datalink. Ann Rheum Dis 2015;74:642–7.
5. Kuo CF, Luo SF. Gout: risk of premature death in gout unchanged for years. Nat Rev Rheumatol 2017;13:200–1.
6. Li-Yu J, Clayburne G, Sieck M, Beutler A, Rull M, Eisner E, et al. Treatment of chronic gout: can we determine when urate stores are depleted enough to prevent attacks of gout? J Rheumatol 2001;28:577–80.
7. Perez-Ruiz F, Herrero-Beites AM. Evaluation and treatment of gout as a chronic disease. Adv Ther 2012;29:935–46.
8. Hui M, Carr A, Cameron S, Davenport G, Doherty M, Forrester H, et al. The British Society for Rheumatology guideline for the management of gout. Rheumatology (Oxford) 2017;56:e1–20.
9. Richette P, Doherty M, Pascual E, Barskova V, Becce F, Castaneda-Sanabria J, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis 2017;76:29–42.
10. Cottrell E, Crabtree V, Edwards JJ, Roddy E. Improvement in the management of gout is vital and overdue: an audit from a UK primary care medical practice. BMC Fam Pract 2013;14:170.
11. Goicoechea M, de Vinuesa SG, Verdalles U, Verde E, Macias N, Santos A, et al. Allopurinol and progression of CKD and cardiovascular events: long-term follow-up of a randomized clinical trial. Am J Kidney Dis 2015;65:543–9.
12. Doehner W, Schoene N, Rauchhaus M, Leyva-Leon F, Pavitt DV, Reaveley DA, et al. Effects of xanthine oxidase inhibition with allopurinol on endothelial function and peripheral blood flow in hyperuricemic patients with chronic heart failure: results from 2 placebo-controlled studies. Circulation 2002;105:2619–24.
13. Luk AJ, Levin GP, Moore EE, Zhou XH, Kestenbaum BR, Choi HK. Allopurinol and mortality in hyperuricaemic patients. Rheumatology (Oxford) 2009;48:804–6.
14. Wei L, Mackenzie IS, Chen Y, Struthers AD, MacDonald TM. Impact of allopurinol use on urate concentration and cardiovascular outcome. Br J Clin Pharmacol 2011;71:600–7.
15. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
16. Chen JH, Lan JL, Cheng CF, Liang WM, Lin HY, Tsay GJ, et al. Effect of urate-lowering therapy on the risk of cardiovascular disease and all-cause mortality in patients with gout: a case-matched cohort study. J Rheumatol 2015;42:1694–701.
17. Dubreuil M, Zhu Y, Zhang Y, Seeger JD, Lu N, Rho YH, et al. Allopurinol initiation and all-cause mortality in the general population. Ann Rheum Dis 2015;74:1368–72.
18. Kuo CF, Grainge MJ, Mallen C, Zhang W, Doherty M. Effect of allopurinol on all-cause mortality in adults with incident gout: propensity score–matched landmark analysis. Rheumatology (Oxford) 2015;54:2145–50.
19. Coburn BW, Michaud K, Bergman DA, Mikuls TR. Allopurinol dose escalation and mortality among patients with gout: a national propensity-matched cohort study. Arthritis Rheumatol 2018;70:1298–307.
20. White WB, Saag KG, Becker MA, Borer JS, Gorelick PB, Whelton A, et al. Cardiovascular safety of febuxostat or allopurinol in patients with gout. N Engl J Med 2018;378:1200–10.
21. Li S, Yang H, Guo Y, Wei F, Yang X, Li D, et al. Comparative efficacy and safety of urate-lowering therapy for the treatment of hyperuricemia: a systematic review and network meta-analysis. Sci Rep 2016;6:33082.
22. Ramasamy SN, Korb-Wells CS, Kannangara DR, Smith MW, Wang N, Roberts DM, et al. Allopurinol hypersensitivity: a systematic review of all published cases, 1950–2012. Drug Saf 2013;36:953–80.

Arthritis Care & Research
Vol. 73, No. 7, July 2021, pp 1055–1060

American College of Rheumatology
Empowering Rheumatology Professionals

LETTERS

DOI 10.1002/acr.24359

The condition of symmetrical sacroiliitis in axial spondyloarthritis: comment on the article by Coates et al

To the Editor:

I read the article by Coates et al (1), recently published in *Arthritis Care & Research*, with great interest as it relates to the radiographic phenotype of axial spondyloarthritis according to the presence of HLA–B27, regardless of the primary diagnosis of ankylosing spondylitis (AS) or psoriatic arthritis (PsA). The authors stated that the HLA–B27 gene is related to radiographic progression, syndesmophyte symmetry, and marginal syndesmophytes, but not to sacroiliac symmetry. I would like to offer some suggestions from a biomechanical aspect, based on their important findings. Chronic biomechanical stress and microdamage have been postulated as major triggering factors for lesion development, particularly in PsA. As compatible with a biomechanical standpoint, tissue-specific kinetic factors (applied forces during motion) may play a critical role in the development of psoriatic lesions. Thus, PsA can de defined as an inflammatory disorder of more mobile musculoskeletal structures when AS is considered.

Psoriatic spondylitis is characterized by asymmetrical involvement of the vertebral column, but the predominant form of sacroiliitis might be symmetric or bilateral, similar to AS (2). The spinal column moves in 3 dimensions, with combined concentric and eccentric muscle contractions. Consequently, the distribution of applied forces under gravity may often be asymmetric during motion.

Conversely, sacroiliac (SI) joints usually function in a symmetrical pattern, with limited but irregular multiaxial mobility, although mechanical load distribution frequently changes as a result of spinal mobility. The main functions of SI joints are to connect the vertebral column and pelvis, bear the trunk weight, and absorb the transmitted compressive forces (and shock waves) (3,4). Most importantly, SI joints especially are insufficient to withstand axial compressive forces as compared to the lumbar spine (4). During the stance phase of the gait cycle, the sacrum bears the body weight as the ilium withstands the upward ground reaction force. Consequently, the SI joints are repetitively exposed to vertical shearing stress during walking as an example of chronic physiologic mechanical stress (3,5,6). Obesity increases mechanical stress. Therefore, a symmetrical sacroiliac involvement pattern may also be compatible with the biomechanical standpoint of PsA.

On the other hand, AS is characterized by severe low back pain accompanying symmetrical involvement of both the vertebral column and SI joints that might further progress to a bamboo spine appearance. Rapidly progressive complete ankylosis may be the major differentiating radiographic feature of AS sacroiliitis. Most likely, progressive symmetrical ankylosis of SI joints, together with the symmetrical ligamentous ossification of the vertebral column, may point to a different pathophysiologic process for AS that is strongly catalyzed by HLA–B27. Notably, SI joint mobility is approximately 40% less for men than for women (7). Besides hormonal factors, male SI joint development has been suggested to be a functional adaptation to cope with major forces, and thickening of ligaments leads to low mobility (8). If the SI joint becomes markedly hypomobile, it may not be able to effectively dissipate forces, and load transmission is altered as well (4). However, effective load transmission and counterbalance of the shearing forces acting on SI joints are mainly dependent on adequate control of lumbopelvic muscles. Overrigid lumbopelvic stabilization may exist in AS, as indicated by axial myofascial hypertonicity, although it might be considered as a consequence (9). Interestingly, the scintigraphic SI joint index has been reported to be higher in healthy men than in healthy women, thus suggesting a potential stress condition for male SI joints (10).

Dr. Tönük has received speaking fees from Pfizer (less than $10,000). No other disclosures relevant to this article were reported.

Şükrü Burak Tönük, MD
Bolu Abant Izzet Baysal University School of Medicine
Bolu, Turkey

1. Coates LC, Baraliakos X, Blanco FJ, Blanco-Morales E, Braun J, Chandran V, et al. Phenotype of axial spondyloarthritis: is it dependent on HLA–B27 status? Arthritis Care Res (Hoboken) 2021;73:856–60.
2. Jadon DR, Sengupta R, Nightingale A, Lindsay M, Korendowych E, Robinson G, et al. Axial disease in psoriatic arthritis study: defining the clinical and radiographic phenotype of psoriatic spondyloarthritis. Ann Rheum Dis 2017;76:701–7.
3. Vleeming A, Schuenke M. Form and force closure of the sacroiliac joints. PM R 2019;11 Suppl 1:S24–31.
4. Dreyfuss P, Dreyer SJ, Cole A, Mayo K. Sacroiliac joint pain. J Am Acad Orthop Surg 2004;12:255–65.
5. Murakami E. Anatomy. In: Sacroiliac joint disorder: accurately diagnosing low back pain. Singapore: Springer Nature; 2019. p. 7–11.
6. Pel JJ, Spoor CW, Pool-Goudzwaard AL, Hoek van Dijke GA, Snijders CJ. Biomechanical analysis of reducing sacroiliac joint shear load by optimization of pelvic muscle and ligament forces. Ann Biomed Eng 2008;36:415–24.
7. Vleeming A, Schuenke MD, Masi AT, Carreiro JE, Danneels L, Willard FH. The sacroiliac joint: an overview of its anatomy, function and potential clinical implications. J Anat 2012;221:537–67.
8. Vleeming A, Stoeckart R, Snijders CJ. General introduction: developmental biology of the sacroiliac joint. In: Vleeming A, Mooney V,

Snijders C, editors. Low back pain and its relation to the sacroiliac joint. Rotterdam: ECO; 1992. p. 13–24.
9. Masi AT, Nair K, Andonian BJ, Prus KM, Kelly J, Sanchez JR, et al. Integrative structural biomechanical concepts of ankylosing spondylitis. Arthritis 2011;2011:205904.
10. Miron SD, Khan MA, Wiesen EJ, Kushner I, Bellon EM. The value of quantitative sacroiliac scintigraphy in detection of sacroiliitis. Clin Rheumatol 1983;2:407–14.

DOI 10.1002/acr.24457

Ustekinumab in giant cell arteritis: comment on the article by Matza et al

To the Editor:

We read with interest the article by Matza et al on ustekinumab in giant cell arteritis (GCA) recently published in *Arthritis Care and Research* (1). The rationale for using ustekinumab in GCA is that it targets the Th1 and Th17 pathways felt to be centrally involved in GCA pathogenesis, potentially impacting at a more proximal stage than other agents such as tocilizumab (2). It is encouraging to see efforts to explore alternative treatment options for GCA, but we believe a number of aspects of this study require further consideration to inform the potential role of ustekinumab in GCA.

The authors enrolled patients with new-onset or relapsing GCA and administered ustekinumab in combination with a modified accelerated glucocorticoid taper similar to that employed in the GiACTA study (3). They enrolled 13 patients, but the study was terminated after 7 of the first 10 patients enrolled relapsed. This relapse rate is high and in excess of that normally experienced in clinical practice (4). We suspect that the relapse rate is likely due to the accelerated taper, which may not be appropriate for relapsing patients. Additionally, the use of an accelerated taper will potentially favor more rapidly acting treatments. An effective treatment with a slower onset of action may appear falsely ineffective in this design. Therefore, the use of the accelerated glucocorticoid taper in the current study may have biased the results to the null. Only 3 patients (23%) exhibited cranial symptoms potentially consistent with active GCA at the time of relapse. Reliably assessing efficacy of a medication in the absence of a control group is not possible. The decision to discontinue the study is regrettable. Were there prespecified criteria for study discontinuation?

One patient relapsed while taking a 9-mg dose of prednisone once a day. This patient's disease had been refractory to multiple other therapies, including tocilizumab, abatacept, and methotrexate. The other 6 patients with clinical relapse did so while taking a median dose (2 mg) of prednisone once a day, mostly with polymyalgic symptoms. This low dose of prednisone is below the level that would be expected to contribute to the known significant burden of glucocorticoid toxicity in GCA patients (5). In many patients with relapsing GCA, a rapid taper and total discontinuation of glucocorticoids may be difficult due to factors unrelated to GCA or polymyalgia rheumatica (PMR). Bursitis, osteoarthritis, calcium pyrophosphate arthropathy, and adrenal insufficiency are common steroid-responsive comorbidities (6,7) that can be unmasked by tapering off low doses of prednisolone, making the assessment of disease activity challenging.

Three patients were deemed to have not achieved the primary end point because of abnormal acute-phase reactants (APRs). Isolated elevations of APRs are not typically considered as evidence of disease flare in the absence of symptoms or signs referable to active GCA/PMR. These APR elevations could also conceivably relate to intercurrent infection or other pathology. This difference from the experience with tocilizumab is expected, where normalization of APRs is expected and improvement of polymyalgic symptoms and a general sense of well-being is enhanced, independent of its impact on disease activity.

An alternative, more optimistic assessment of the available data is that all but 1 of the patients in this study were able to taper to a <5-mg daily dose of prednisolone without relapse. This tapering is not achieved by many GCA patients with the existing standard of care.

We previously reported a prospective study of 25 patients with refractory GCA treated with ustekinumab. This group of patients represents a common clinical scenario, patients who require a sustained-maintenance glucocorticoid dose and are unable to taper below this level due to recurrent symptoms, with a subsequent accumulation of glucocorticoid-related adverse events. Clearly, in this patient cohort a rapid glucocorticoid taper is not appropriate and is unlikely to be successful. To comprehensively assess treatment options in GCA, such as ustekinumab, alternative complementary efficacy evaluations are needed; evaluating the effect on long-term maintenance glucocorticoid requirements is one such option. We demonstrated a significant reduction in prednisolone dose over 52 weeks from a median of 20 mg to 5 mg with an accompanying decrease in APRs (8,9).

There is an urgent unmet need for new treatment options in GCA to minimize treatment-related adverse events and improve long-term outcomes for patients. Our pilot study of ustekinumab in GCA, while suggestive of a benefit, did not conclusively prove that ustekinumab is an effective treatment in GCA. Similarly, the study by Matza et al does not prove ustekinumab is not an effective treatment for GCA, and the limitations of their study preclude dismissal of its potential utility in GCA. Contrasting the 2 studies would suggest that a rapid prednisolone taper is not advisable in patients with relapsing GCA being treated with ustekinumab. However, a randomized controlled trial of ustekinumab is required to properly evaluate its efficacy for GCA.

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

Richard Conway, MB, BCh, BAO, PhD
St. James's Hospital
Dublin, Ireland
Eamonn S. Molloy, MB, BCh, BAO, MD, MS
St. Vincent's University Hospital
and Dublin Academic Medical Centre
Dublin, Ireland

1. Matza MA, Fernandes AD, Stone JH, Unizony SH. Ustekinumab for the treatment of giant cell arteritis. Arthritis Care Res (Hoboken) 2021;73:893–7.
2. Conway R, O'Neill L, McCarthy GM, Murphy CC, Fabre A, Kennedy S, et al. Interleukin 12 and interleukin 23 play key pathogenic roles in inflammatory and proliferative pathways in giant cell arteritis. Ann Rheum Dis 2018;77:1815–24.
3. Stone JH, Tuckwell K, Dimonaco S, Klearman M, Aringer M, Blockmans D, et al. Trial of tocilizumab in giant-cell arteritis. N Engl J Med 2017;377:317–28.
4. Mainbourg S, Addario A, Samson M, Puéchal X, François M, Durupt S, et al. Prevalence of giant cell arteritis relapse in patients treated with glucocorticoids: a meta-analysis. Arthritis Care Res (Hoboken) 2020;72:838–49.
5. Proven A, Gabriel SE, Orces C, O'Fallon WM, Hunder GG. Glucocorticoid therapy in giant cell arteritis: duration and adverse outcomes. Arthritis Rheum 2003;49:703–8.
6. Jamilloux Y, Liozon E, Pugnet G, Nadalon S, Heang Ly K, Dumonteil S, et al. Recovery of adrenal function after long-term glucocorticoid therapy for giant cell arteritis: a cohort study. PLoS One 2013;8:e68713.
7. Kroon FP, Kortekaas MC, Boonen A, Böhringer S, Reijnierse M, Rosendaal FR, et al. Results of a 6-week treatment with 10 mg prednisolone in patients with hand osteoarthritis (HOPE): a double-blind, randomised, placebo-controlled trial. Lancet 2019;394:1993–2001.
8. Conway R, O'Neill L, Gallagher P, McCarthy GM, Murphy CC, Veale DJ, et al. Ustekinumab for refractory giant cell arteritis: a prospective 52-week trial. Semin Arthritis Rheum 2018;48:523–8.
9. Conway R, O'Neill L, O'Flynn E, Gallagher P, McCarthy GM, Murphy CC, et al. Ustekinumab for the treatment of refractory giant cell arteritis. Ann Rheum Dis 2016;75:1578–9.

DOI 10.1002/acr.24454

Reply

To the Editor:

We appreciate the opportunity to respond to the letter by Conway and Molloy regarding our recent article. Our results demonstrated that ustekinumab (UST) administered every 8 weeks was well tolerated but did not prevent disease relapse in a significant proportion of patients with GCA when used in combination with 6 months of prednisone. We acknowledge that a small sample size and the lack of a control group are important limitations of our study and agree with Conway and Molloy that further research is needed to assess the utility of UST more definitively in GCA.

The design of our study responded to the 2 most frequent problems encountered by GCA patients in general, disease relapse and prolonged glucocorticoid exposure leading to toxicity. Therefore, our prospective, single-arm pilot trial enrolled patients without preselection (e.g., new-onset versus relapsing disease) and employed a prespecified prednisone taper over 6 months. Our goal was to gather preliminary data on the safety and efficacy of UST using this protocol, to then conduct a larger study if we had found a positive signal. Unfortunately, the results of our study demonstrated that UST was not as effective as originally hoped. We agree with Conway and Molloy that other studies can be designed to better define whether there is a role for UST in the treatment of GCA. Limiting enrollment to relapsing patients, using longer prednisone tapers, or leaving patients on a low prednisone maintenance dose are some possibilities. Ultimately, we agree that to rigorously test the efficacy and safety of UST for GCA, an adequately powered randomized, placebo-controlled trial is required.

Our trial may be insightful in understanding that a steroid taper of 6 months may indeed be too short when used with UST. Although shorter than tapers commonly used in clinical practice, the 6-month taper employed in our trial (Supplementary Table 1 of our article, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract) is in accordance with those used in most GCA clinical trials (1–4) (ClinicalTrials.gov identifiers: NCT03827018, NCT03600805, NCT02531633, NCT03725202). Such tapers have had the objective of assessing whether an intervention offers unequivocal glucocorticoid-sparing potential, which was one of the main drivers of our investigation. Furthermore, when planning the prednisone regimen used in our study, we considered that between 8 and 12 weeks are required for UST to exert its immunomodulatory effects (5). To ensure enough prednisone treatment until the onset of action of UST, the duration of the taper was 6 months in all cases, leading to prednisone doses of 12.5 mg/day, 10 mg/day, and 6 mg/day at week 12 for patients starting the taper at 60 mg, 40 mg, and 20 mg, respectively. In addition, we opted for a higher UST dose (90 mg at baseline, week 4, and every 8 weeks) than the one used in the prior study (6) to achieve higher serum concentrations and increase the cumulative exposure to this agent. As was discussed in our prior response to the letter by Samson and Bonnotte, a posteriori analysis of the flares occurring in our study (Supplementary Table 2 of our article, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract) further suggests, in our opinion, that the prednisone taper chosen was adequate if indeed UST was to be effective. All disease flares occurred at or after week 16 following an adequate duration of UST treatment. In addition, the individual prednisone doses at week 12 were 10 mg/day or greater for all except 1 patient, who experienced a flare, indicating that the great majority of the patients with flares received a daily prednisone dose commonly associated with remission maintenance (7–9) by the onset of action of UST. Nevertheless, we also believe it is important to understand how UST would fare as a glucocorticoid-sparing agent with a slightly longer taper or when patients remain on low, nontoxic glucocorticoid doses.

The relapse rate seen in our study is within the range reported in some cohort studies of patients with GCA treated only with glucocorticoids (7,10–12), clinical trials of drugs proven to be ineffective, and placebo arms of clinical trials of drugs

that showed efficacy (2–4,13). In addition, although we agree that disease activity monitoring in GCA is often confounded by unrelated pathology (e.g., adrenal insufficiency or noninflammatory musculoskeletal conditions), the flares observed in this cohort of patients were fairly classic in terms of symptoms and also demonstrated re-elevation of the serum inflammatory markers (Supplementary Table 2 of our article). To account for the problem of elevated inflammatory markers without obvious clinical manifestations (frequently seen in clinical practice and of unclear significance), we also completed a sensitivity analysis excluding the erythrocyte sedimentation rate and C-reactive protein values from the definition of sustained remission. Such analysis demonstrated that more than half of the patients did not benefit from UST following our study protocol.

We had prespecified criteria to stop the trial based on safety, but not based on efficacy, given the small sample planned for this pilot study. However, after seeing that 70% of the first 10 patients recruited had experienced a flare, and knowing that better treatment alternatives were available at that time, we made what we thought was a rational and ethical decision to stop enrollment. We do regret having to terminate the trial prematurely.

Ultimately, we agree with Conway and Molloy that UST could have a role in the treatment of GCA. Our study was an initial approach to assess whether this agent would maintain disease remission in the absence of glucocorticoids, given the unmet need for more glucocorticoid-sparing agents in GCA. Perhaps a different study population or trial design will demonstrate efficacy, as the report by Conway et al suggests (6). Our conclusion is not that UST is ineffective, but rather that it did not allow for acceptable rates of remission when used with a prednisone taper over 6 months.

Mark A. Matza, MD, MBA
Ana D. Fernandes, BA
John H. Stone, MD, MPH
Sebastian H. Unizony, MD
Massachusetts General Hospital and Harvard Medical School
Boston, Massachusetts

1. Hoffman GS, Cid MC, Hellmann DB, Guillevin L, Stone JH, Schousboe J, et al. A multicenter, randomized, double-blind, placebo-controlled trial of adjuvant methotrexate treatment for giant cell arteritis. Arthritis Rheum 2002;46:1309–18.
2. Langford CA, Cuthbertson D, Ytterberg SR, Khalidi N, Monach PA, Carette S, et al. A randomized, double-blind trial of abatacept (CTLA-4Ig) for the treatment of giant cell arteritis. Arthritis Rheumatol 2017; 69:837–45.
3. Stone JH, Tuckwell K, Dimonaco S, Klearman M, Aringer M, Blockmans D, et al. Trial of tocilizumab in giant-cell arteritis. N Engl J Med 2017;377:317–28.
4. Villiger PM, Adler S, Kuchen S, Wermelinger F, Dan D, Fiege V, et al. Tocilizumab for induction and maintenance of remission in giant cell arteritis: a phase 2, randomised, double-blind, placebo-controlled trial. Lancet 2016;387:1921–7.
5. Stelara [package insert]. Horsham PA: Janssen Pharmaceutical Companies; 2016.
6. Conway R, O'Neill L, Gallagher P, McCarthy GM, Murphy CC, Veale DJ, et al. Ustekinumab for refractory giant cell arteritis: A prospective 52-week trial. Semin Arthritis Rheum 2018;48: 523–8.
7. Alba MA, Garcia-Martinez A, Prieto-Gonzalez S, Tavera-Bahillo I, Corbera-Bellalta M, Planas-Rigol E, et al. Relapses in patients with giant cell arteritis: prevalence, characteristics, and associated clinical findings in a longitudinally followed cohort of 106 patients. Medicine (Baltimore) 2014;93:194–201.
8. Stone JH, Tuckwell K, Dimonaco S, Klearman M, Aringer M, Blockmans D, et al. Glucocorticoid dosages and acute-phase reactant levels at giant cell arteritis flare in a randomized trial of tocilizumab. Arthritis Rheumatol 2019;71:1329–38.
9. Kermani TA, Warrington KJ, Cuthbertson D, Carette S, Hoffman GS, Khalidi NA, et al. Disease relapses among patients with giant cell arteritis: a prospective, longitudinal cohort study. J Rheumatol 2015;42:1213–7.
10. Labarca C, Koster MJ, Crowson CS, Makol A, Ytterberg SR, Matteson EL, et al. Predictors of relapse and treatment outcomes in biopsy-proven giant cell arteritis: a retrospective cohort study. Rheumatology (Oxford) 2016;55:347–56.
11. Restuccia G, Boiardi L, Cavazza A, Catanoso M, Macchioni P, Muratore F, et al. Flares in biopsy-proven giant cell arteritis in Northern Italy: characteristics and predictors in a long-term follow-up study. Medicine (Baltimore) 2016;95:e3524.
12. Hernández-Rodríguez J, García-Martínez A, Casademont J, Filella X, Esteban MJ, López-Soto A, et al. A strong initial systemic inflammatory response is associated with higher corticosteroid requirements and longer duration of therapy in patients with giant-cell arteritis. Arthritis Rheum 2002;47:29–35.
13. Seror R, Baron G, Hachulla E, Debandt M, Larroche C, Puechal X, et al. Adalimumab for steroid sparing in patients with giant-cell arteritis: results of a multicentre randomised controlled trial. Ann Rheum Dis 2014;73:2074–81.

DOI 10.1002/acr.24376

Ustekinumab for the treatment of giant cell arteritis: comment on the article by Matza et al

To the Editor:

We read with great interest the article by Matza et al, recently published in *Arthritis Care & Research*, that reports on the authors' prospective, open-label, single-arm, single-center trial that was designed to evaluate the efficacy and safety of ustekinumab (UST) in combination with prednisone in patients with active giant cell arteritis (GCA) (ClinicalTrials.gov NCT02955147) (1). The authors concluded that UST is an ineffective treatment for GCA, since only 23% of the 13 patients enrolled achieved the primary end point (i.e., absence of relapse through week 52 and normalization of the erythrocyte sedimentation rate and C-reactive protein level), which is inferior to previous results obtained in patients treated with tocilizumab (TCZ) (2).

Although this study suggests that, unlike TCZ (2), UST is ineffective in controlling GCA unless it is associated with glucocorticoids, we have several reasons to believe it is too early to draw conclusions regarding the use of UST as a glucocorticoid-sparing drug in GCA. First, this trial was not controlled, and more importantly, the prednisone regimen used in this study was short. Unlike TCZ and other biologics that directly target proinflammatory cytokines, resulting in rapid resolution of inflammation, UST modulates T cell homeostasis by targeting both interleukin-12 (IL-12) and IL-23 pathways through a blockade of the p40 subunit, resulting in an onset of action 8 to 12 weeks later (3). In the study by Matza et al (1), patients were receiving 12.5, 10, or 6 mg/day of prednisone at 12 weeks (depending on their starting dose: 60 mg [3 patients], 40 mg [9 patients] or 20 mg [1 patient]). We therefore assume that many patients relapsed because the prednisone dose was too low to control GCA when UST became effective, which may explain why the authors did not observe the effectiveness of UST as in previous reports (4–6).

Recent advances in our understanding of the pathophysiology of GCA have shown the critical implication of Th1 and Th17 cells, which produce interferon-γ (IFNγ) and IL-17, respectively (7). Unlike Th17 cells, Th1 cells are resistant to glucocorticoids (8); this resistance results in the chronic production of IFNγ in GCA, which plays a major role in triggering vascular remodeling and for which there is currently no effective treatment (9). By blocking IL-12 and IL-23, the 2 key cytokines involved in Th1 and Th17 polarization, UST provides a great opportunity to disrupt both Th1 and Th17 pathways (6), potentially leading to better control of vascular remodeling and inflammation in GCA. In addition, UST has shown an excellent safety profile in clinical trials (10,11), which is a considerable advantage for patients experiencing relapsing GCA, who are often elderly and frail with a high risk of infection. For all these reasons, we believe that UST remains an interesting treatment that should particularly be evaluated for patients with relapsing GCA, especially if they are not eligible for TCZ.

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

Maxime Samson, MD, PhD
Hôpital François Mitterrand, CHU Dijon
Dijon, France
Bernard Bonnotte, MD, PhD
University of Bourgogne Franche-Comté, INSERM
Dijon, France

1. Matza MA, Fernandes AD, Stone JH, Unizony SH. Ustekinumab for the treatment of giant cell arteritis. Arthritis Care Res (Hoboken) 2021;73:893–7.
2. Stone JH, Tuckwell K, Dimonaco S, Klearman M, Aringer M, Blockmans D, et al. Trial of tocilizumab in giant-cell arteritis. N Engl J Med 2017;377:317–28.
3. Ritchlin C, Rahman P, Kavanaugh A, McInnes IB, Puig L, Li S, et al. Efficacy and safety of the anti-IL-12/23 p40 monoclonal antibody, ustekinumab, in patients with active psoriatic arthritis despite conventional non-biological and biological anti-tumour necrosis factor therapy: 6-month and 1-year results of the phase 3, multicentre, double-blind, placebo-controlled, randomised PSUMMIT 2 trial. Ann Rheum Dis 2014;73:990–9.
4. Conway R, O'Neill L, Gallagher P, McCarthy GM, Murphy CC, Veale DJ, et al. Ustekinumab for refractory giant cell arteritis: a prospective 52-week trial. Semin Arthritis Rheum 2018;48:523–8.
5. Conway R, O'Neill L, O'Flynn E, Gallagher P, McCarthy GM, Murphy CC, et al. Ustekinumab for the treatment of refractory giant cell arteritis. Ann Rheum Dis 2016;75:1578–9.
6. Samson M, Ghesquière T, Berthier S, Bonnotte B. Ustekinumab inhibits Th1 and Th17 polarisation in a patient with giant cell arteritis. Ann Rheum Dis 2018;77:e6.
7. Samson M, Corbera-Bellalta M, Audia S, Planas-Rigol E, Martin L, Cid MC, et al. Recent advances in our understanding of giant cell arteritis pathogenesis. Autoimmun Rev 2017;16:833–44.
8. Deng J, Younge BR, Olshen RA, Goronzy JJ, Weyand CM. Th17 and Th1 T-cell responses in giant cell arteritis. Circulation 2010;121:906–15.
9. Samson M, Espígol-Frigolé G, Terrades-García N, Prieto-González S, Corbera-Bellalta M, Alba-Rovira R, et al. Biological treatments in giant cell arteritis & Takayasu arteritis. Eur J Intern Med 2018;50:12–19.
10. McInnes IB, Kavanaugh A, Gottlieb AB, Puig L, Rahman P, Ritchlin C, et al. Efficacy and safety of ustekinumab in patients with active psoriatic arthritis: 1 year results of the phase 3, multicentre, double-blind, placebo-controlled PSUMMIT 1 trial. Lancet 2013;382:780–9.
11. Gordon KB, Papp KA, Langley RG, Ho V, Kimball AB, Guzzo C, et al. Long-term safety experience of ustekinumab in patients with moderate to severe psoriasis (Part II of II): results from analyses of infections and malignancy from pooled phase II and III clinical trials. J Am Acad Dermatol 2012;66:742–51.

DOI 10.1002/acr.24378

Reply

To the Editor:

Thank you for allowing us to respond to the letter by Samson and Bonnotte regarding our recent article. Our results demonstrated that UST was well tolerated but did not prevent disease relapse in a significant proportion of patients with GCA after prednisone was discontinued or tapered to a low daily dose. Samson and Bonnotte's letter suggests that it may be too early to draw conclusions regarding the efficacy of UST as a glucocorticoid-sparing drug in GCA for 2 reasons. First, the study was not controlled, and second, the prednisone taper used was short, perhaps not allowing UST to exert its putative corrective actions in GCA (1), which would be expected 8–12 weeks after treatment initiation.

As stated in our article, we acknowledge the limitations of our study and believe that firm conclusions about the efficacy of IL-12/23p40 blockade in GCA cannot be drawn until more robust data are available. We agree with Samson and Bonnotte that an adequately powered, randomized, double-blind, placebo-controlled trial would be one of the best approaches to assess the value of UST for treatment of this disorder.

Our opinion regarding whether the prednisone taper employed in our trial was too short to allow UST to work, however, differs from that of Samson and Bonnotte. First, when planning this trial, we decided to use 6 months of prednisone in all cases (Supplementary Table 1 of our article, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract) regardless of the prednisone dose at baseline to ensure that UST reached steady-state levels (2) by the time of glucocorticoid discontinuation. In addition, unlike a prior study that dosed UST every 12 weeks (3), we chose to administer UST every 8 weeks to achieve higher serum concentrations and ultimately increase the cumulative exposure to this agent in our cohort of patients. Second, all disease flares occurred at week 16 or later, with a mean time to flare of 23 weeks (Table 2 and Supplementary Table 2 of our article, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24200/abstract), which is longer than the 8–12 weeks proposed as the time period required for UST to modulate the adaptive immune pathways playing a role in GCA. Third, the daily prednisone doses of the patients who had flares were 17.5 mg (n = 1), 15 mg (n = 5), and 9 mg (n = 1) at week 8, and 12.5 mg (n = 1), 10 mg (n = 5), and 6 mg (n = 1) at week 12. Thus, for all except 1 of the patients who relapsed in our study, the individual prednisone doses at week 12 were ≥10 mg/day, which are doses typically enough to prevent flare in the majority of cases treated only with glucocorticoids (4–6). Therefore, one would think that the duration of the glucocorticoid taper and the daily prednisone dose of the patients at week 12 should allow UST to take over remission maintenance 12 weeks after UST initiation if indeed this agent was effective. Of note, by the time of flare, patients had received an average of four 90-mg injections of UST.

We speculate in the discussion of our article about possible reasons behind the discrepancies between our study and the prior study of UST for GCA conducted by Conway et al (3), in which patients with relapsing or refractory disease responded favorably. The most likely explanation in our opinion is that approximately 75% of patients in the study by Conway et al were still receiving glucocorticoids at week 52 when the outcomes were assessed (3). Nevertheless, we believe that the major unmet need in GCA is to maintain disease remission and at the same time minimize the cumulative glucocorticoid exposure and its associated toxicity (7). Like recent landmark clinical trials in the field (8,9), we also followed this premise when designing the prednisone taper of our study protocol.

We understand that GCA patients are in need of more and better treatment options, which should address the historical problem of relying excessively on glucocorticoids to prevent relapse. Unfortunately, we think that our study does give reason to believe that UST may not be as effective as initially hoped in GCA. We agree with Samson and Bonnotte that researchers need to gather more data on the effects of IL-12 and IL-23 modulation with UST (perhaps in relapsing patients) and other molecules (e.g., IL-23p19 inhibitors) as part of the search that will hopefully diversify our therapeutic arsenal against GCA in the coming years.

Mark A. Matza, MD, MBA
Ana D. Fernandes, BA
John H. Stone, MD, MPH
Sebastian H. Unizony, MD
Massachusetts General Hospital and Harvard Medical School
Boston, Massachusetts

1. Samson M, Ghesquiere T, Berthier S, Bonnotte B. Ustekinumab inhibits Th1 and Th17 polarisation in a patient with giant cell arteritis. Ann Rheum Dis 2018;77:e6.
2. Stelara [package insert]. Horsham PA: Janssen Pharmaceutical Companies; 2016.
3. Conway R, O'Neill L, Gallagher P, McCarthy GM, Murphy CC, Veale DJ, et al. Ustekinumab for refractory giant cell arteritis: a prospective 52-week trial. Semin Arthritis Rheum 2018;48:523–8.
4. Alba MA, Garcia-Martinez A, Prieto-Gonzalez S, Tavera-Bahillo I, Corbera-Bellalta M, Planas-Rigol E, et al. Relapses in patients with giant cell arteritis: prevalence, characteristics, and associated clinical findings in a longitudinally followed cohort of 106 patients. Medicine (Baltimore) 2014;93:194–201.
5. Stone JH, Tuckwell K, Dimonaco S, Klearman M, Aringer M, Blockmans D, et al. Glucocorticoid dosages and acute-phase reactant levels at giant cell arteritis flare in a randomized trial of tocilizumab. Arthritis Rheumatol 2019;71:1329–38.
6. Kermani TA, Warrington KJ, Cuthbertson D, Carette S, Hoffman GS, Khalidi NA, et al. Disease relapses among patients with giant cell arteritis: a prospective, longitudinal cohort study. J Rheumatol 2015;42:1213–7.
7. Kermani TA, Sreih AG, Cuthbertson D, Carette S, Hoffman GS, Khalidi NA, et al. Evaluation of damage in giant cell arteritis. Rheumatology (Oxford) 2018;57:322–8.
8. Langford CA, Cuthbertson D, Ytterberg SR, Khalidi N, Monach PA, Carette S, et al. A randomized, double-blind trial of abatacept (CTLA-4Ig) for the treatment of giant cell arteritis. Arthritis Rheumatol 2017;69:837–45.
9. Stone JH, Tuckwell K, Dimonaco S, Klearman M, Aringer M, Blockmans D, et al. Trial of tocilizumab in giant-cell arteritis. N Engl J Med 2017;377:317–28.