Arthritis Care & Research

Aims and Scope

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

Volume 75 No. 12 December 2023

Arthritis Care & Research

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

Editor
Kelli D. Allen, PhD
University of North Carolina at Chapel Hill
Durham VA Medical Center

Deputy Editors
S. Sam Lim, MD, MPH
Emory University, Atlanta
Todd A. Schwartz, DrPH
University of North Carolina at Chapel Hill

Social Media Editor
Paul H. Sufka, MD, *Eagan*

Journal Publications Committee
Amr Sawalha, MD, Chair, *Pittsburgh*
Susan Boackle, MD, *Denver*
Aileen Davis, PhD, *Toronto*
Daniel Horton, MD, *New Brunswick*
Donnamarie Krause, PhD, OTR/L, *Las Vegas*
Brandon Law, MD, *Baltimore*
Michelle Ormseth, MD, *Nashville*
Suraj Rajasimhan, PharmD, *Wilmington*

Associate Editors
Joshua Baker, MD, MSCE, *Philadelphia*
Nancy Baker, ScD, MPH, OT, *Boston*
Cheryl C. M. Barnabe, MD, MSc, *Calgary*
Bonnie L. Bermas, MD, *Dallas*
Lorinda Chung, MD, MS, *Stanford*
Maria I. Danila, MD, MSc, MSPH, *Birmingham*
Robert F. DeVellis, PhD, *Chapel Hill*
Afton L. Hassett, PsyD, *Ann Arbor*
Puja P. Khanna, MD, MPH, *Ann Arbor*
Kanta Kumar, PhD, *Birmingham, UK*
Crystal MacKay, PhD, MHSc, BScPT, *Toronto*
Natalie McCormick, PhD, *Boston and Vancouver*
Kaleb Michaud, PhD, *Omaha*
Eli M. Miloslavsky, MD, *Boston*
Michael H. Weisman, MD, *Palo Alto and Los Angeles*
Pamela F. Weiss, MD, MSCE, *Philadelphia*
Daniel K. White, PT, ScD, MSc, *Newark*

Editorial Staff
Kimberly M. Murphy, *Senior Publishing Director, Delaware*
Maggie Parry, *Director, Quality and Production, Georgia*
Brian Robinson, *Director, Digital Content, Pennsylvania*
Chris Reynolds, *Product Manager, Georgia*
Christy Austin, *Publishing Coordinator, Washington*
Laura Bolte, *Managing Editor, North Carolina*

Editorial Board
Carolina Alvarez, MS, *Chapel Hill*
Liubov Arbeeva, PhD, *Carrboro*
Kathryn L. Bacon, MPH, PhD, *Concord*
Ram Bajpai, PhD, *Newcastle-under-Lyme*
Matthew Baker, MD, MS, *Stanford*
April Barnado, MD, MSCI, *Nashville*
Christie Bartels, MD, MS, *Madison*
Jennifer Barton, MD, *Portland*
Teresa J. Brady, PhD, *Atlanta*
Mayilee Canizares, PhD, *Toronto*
Robin Christensen, BSc, MSc, Phd, *Copenhagen*
Rebecca Cleveland, PhD, *Chapel Hill*
Jamie E. Collins, PhD, *Boston*
Delphine Courvoisier, PhD, *Geneva*
Cynthia Crowson, PhD, *Rochester*
John M. Davis, MD, *Rochester*
Robyn Domsic, MD, MPH, *Pittsburgh*
Cristina Drenkard, MD, PhD, *Atlanta*
Jeffrey Driban, PhD, *Boston*
Alyssa B. Dufour, PhD, *Boston*
Vicky Duong, DPT, PhD, *Sydney*
Bryant England, MD, *Omaha*
Titilola Falasinnu, PhD, *Stanford*
Candance H. Feldman, *Boston*
Elizabeth Ferucci, MD, MPH, *Anchorage*
John D. FitzGerald, MD, PhD, *Los Angeles*
Ivan Foeldvari, MD, *Hamburg*
Tracy Frech, MD, MS, *Salt Lake City*
Angelo Gaffo, MD, *Birmingham*
Michael George, MD, MSCE, *Philadelphia*
Yvonne Golightly, PhD, *Chapel Hill*
Jaime Guzman, MD, MSc, *Vancouver*
Reza Jafarzadeh, DVM, MPVM, PhD, *Boston*
Tate Johnson, MD, *Omaha*
Robert Keenan, MD, MPH, MBA, *Durham*
Andrea Knight, MD, MSCE, *Toronto*
Anna Kratz, PhD, *Ann Arbor*
Yvonne C. Lee, MD, MMSc, *Chicago*
Linda Li, PhD, *Vancouver*
Grace Hsaio-Wei Lo, MD, MSc, *Houston*
Una Makris, MD, *Dallas*
Hiral Master, PT, PhD, MPH, *Nashville*
Susan Murphy, ScD, OTR, *Ann Arbor*
Elena Myaseodova, MD, PhD, *Rochester*
Gulsen Ozen, MD, *Omaha*
Sofia Pedro, MSc, *Wichita*
Anthony Perruccio, PhD, *Toronto*
Paula S. Ramos, PhD, *Charleston*
Grant Schulert, MD, PhD, *Cincinnati*
Pascale Schwab, MD, *Portland*
Julia Simard, ScD, *Stanford*
Jeffrey A. Sparks, MD, MMSc, *Boston*
Joshua Stefanik, MSPT, PhD, *Boston*
Sara Tedeschi, MD, MPH, *Boston*
Lauren Terhorst, PhD, *Pittsburgh*
Louise Thoma, PhD, *Chapel Hill*
Zahi Touma, MD, PhD, *Toronto*
Manuel Ugarte-Gil, MD, MSc, *Lima*
Suzanne Verstappen, PhD, *Manchester*
Ernest Vina, MD, MSc, *Tucson*
Dana Voinier, MS, DPT, *Newark*
Zachary Wallace, MD, MSc, *Boston*
Jessica Widdifield, PhD, *Toronto*

Association of Rheumatology Professionals
2022–2023 Executive Committee
Janell Martin, CAE, Executive Director, *Atlanta*

Kori Dewing, ARNP, DNP, *Everett*, President
Aileen Ledingham, MS, PhD, PT, *Waltham*, President-Elect
Adam Goode, DPT, PhD, PT, *Durham,* Member at Large-Secretary
Anna Lawrence, MBA, MS, *Lawrenceville,* Member at Large-Finance
Adena Batterman, MSW, *New York*, eLearning Subcommittee Chair
Kaleb Michaud, PhD, *Omaha*, Committee on Research Representative
Kimberly Steinbarger, DHSc, MHS, PT, *Dixmont,* Practice Committee Chair
Don Miller, PharmD, *Fargo*, Government Affairs Representative
Jillian Rose-Smith, MPH, PhD, MSW, *New York*, Annual Meeting Planning Subcommittee Chair
Barbara Slusher, PA-C, MSW, Summerton, Nominations Chair
Michael LaValley, PhD, *Boston*, Member-at-Large
Susan Bartlett, PhD*, Montreal,* Member-at-Large
Sindhu Johnson, MD, PhD, *Toronto,* ACR Member Representative

Copyright © 2023 American College of Rheumatology. All rights reserved. No part of this publication may be reproduced, stored or transmitted in any form or by any means without the prior permission in writing from the copyright holder. Authorization to copy items for internal and personal use is granted by the copyright holder for libraries and other users registered with their local Reproduction Rights Organization (RRO), e.g. Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923, USA (www.copyright.com), provided the appropriate fee is paid directly to the RRO. This consent does not extend to other kindsof copying such as copying for general distribution, for advertising or promotional purposes, for creating new collective works or for resale. Special requests should be addressed to: permissions@wiley.com

Access Policy: Subject to restrictions on certain backfiles, access to the online version of this issue is available to all registered Wiley InterScience users 12 months after publication. Subscribers and eligible users at subscribing institutions have immediate access in accordance with the relevant subscription type. Please go to onlinelibrary.wiley.com for details.

The views and recommendations expressed in articles, letters, and other communications published in *Arthritis Care & Research* are those of the authors and do not necessarily reflect the opinions of the editors, publisher, or American College of Rheumatology. The publisher and the American College of Rheumatology do not investigate the information contained in the classified advertisements in this journal and assume no responsibility concerning them. Further, the publisher and the American College of Rheumatology do not guarantee,warrant, or endorse any product or service advertised in this journal.

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 75 • DECEMBER 2023 • NO. 12

Cover image: The image on the cover (from Sondhi et al, page 2423) shows intensely fluorodeoxyglucose-avid right adrenal mass measuring 3.5 cm in length and left adrenal mass measuring 2.4 cm in length, along with retroperitoneal nodes.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2405–2419
DOI 10.1002/acr.25240
© 2023 American College of Rheumatology. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

American College of Rheumatology
Empowering Rheumatology Professionals

2022 American College of Rheumatology Guideline for the Prevention and Treatment of Glucocorticoid-Induced Osteoporosis

Mary Beth Humphrey,[1*] Linda Russell,[2*] Maria I. Danila,[3] Howard A. Fink,[4] Gordon Guyatt,[5] Michael Cannon,[6] Liron Caplan,[7] Sara Gore,[8] Jennifer Grossman,[9] Karen E. Hansen,[10] Nancy E. Lane,[11] Nina S. Ma,[12] Marina Magrey,[13] Tim McAlindon,[14] Angela Byun Robinson,[15] Sumona Saha,[10] Charles Womack,[8] Basma Abdulhadi,[3] Julia F. Charles,[16] Jonathan T. L. Cheah,[17] Sharon Chou,[16] Itivrita Goyal,[1] Katherine Haseltine,[2] Lesley Jackson,[3] Reza Mirza,[5] Iram Moledina,[3] Emma Punni,[1] Tim Rinden,[18] Marat Turgunbaev,[19] Katherine Wysham,[20] Amy S. Turner,[19] and Stacey Uhl[21]

Guidelines and recommendations developed and/or endorsed by the American College of Rheumatology (ACR) are intended to provide guidance for patterns of practice and not to dictate the care of a particular patient. The ACR considers adherence to the recommendations within this guideline to be voluntary, with the ultimate determination regarding their application to be made by the clinician 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, and drug formularies or other third-party analyses that cite ACR guidelines should state this. 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. The objective is to update recommendations for prevention and treatment of glucocorticoid-induced osteoporosis (GIOP) for patients with rheumatic or nonrheumatic conditions receiving >3 months treatment with glucocorticoids (GCs) ≥2.5 mg daily.

Methods. An updated systematic literature review was performed for clinical questions on nonpharmacologic, pharmacologic treatments, discontinuation of medications, and sequential therapy. Grading of Recommendations Assessment, Development and Evaluation approach was used to rate the certainty of evidence. A Voting Panel achieved ≥70% consensus on the direction (for or against) and strength (strong or conditional) of recommendations.

Results. For adults beginning or continuing >3 months of GC treatment, we strongly recommend as soon as possible after initiation of GCs, initial assessment of fracture risks with clinical fracture assessment, bone mineral density with vertebral fracture assessment or spinal x-ray, and Fracture Risk Assessment Tool if ≥40 years old. For adults at medium, high, or very high fracture risk, we strongly recommend pharmacologic treatment. Choice of oral or intravenous bisphosphonates, denosumab, or parathyroid hormone analogs should be made by shared decision-making. Anabolic agents are conditionally recommended as initial therapy for those with high and very high fracture risk. Recommendations are made for special populations, including children, people with organ transplants, people who may become pregnant, and people receiving very high-dose GC treatment. New recommendations for both discontinuation of osteoporosis therapy and sequential therapies are included.

Conclusion. This guideline provides direction for clinicians and patients making treatment decisions for management of GIOP. These recommendations should not be used to limit or deny access to therapies.

INTRODUCTION

Glucocorticoids (GCs) remain a common therapeutic modality for patients with a variety of diseases. Prevention of GC-induced bone loss and fractures has been a focus of the American College of Rheumatology (ACR) for many years because patients with osteoporotic fractures have increased risk of morbidity and mortality (1–4). It is estimated that 1% of the US population is treated with long-term GCs (5). GC doses ≥2.5 mg/day increase fracture at both the spine and hip, and GC <2.5 mg/day increase the risk of spinal fractures (6). Both high daily (≥30 mg/day) and high cumulative (≥5 g/year) doses of GCs further increase the risk of fragility fractures, with peak incidence at 12 months (7–11). The highest rate of bone loss occurs within the first 3 to 6 months of GC treatment, due to early osteoclast activation followed by decreased osteoblast proliferation and increased apoptosis of osteoblasts and osteocytes (12). In children, GCs adversely affect bone strength, growth, and peak bone mass, with increased fracture risk (11,13–15). However, children (16) and young adults often regain lost bone when GCs are discontinued (17).

Despite increasing treatment options to prevent and treat glucocorticoid-induced osteoporosis (GIOP), many GC-treated patients are not evaluated or treated, resulting in preventable fractures (18,19). Risk calculators provide estimates of the 10-year risk of major osteoporotic fractures (MOFs) and hip fractures among individuals ≥40 years of age, with adjustment for GC doses >7.5 mg/day or <2.5 mg/day in some calculators (20–22). Of note, the Fracture Risk Assessment Tool (FRAX) is not validated for adults <40 years. These calculators underestimate fracture risk for patients on very high doses of GC therapy (eg, ≥ 30 mg/day) and do not adequately include frailty, multiple fractures, or fall history.

The ACR first published recommendations for prevention and treatment of GIOP in 1996 (23). ACR updated these guidelines in 2001, 2010, and 2017 as new techniques for assessing fracture risk, risk factors, and therapies became available (23–26). This guideline updated the literature search from April 23, 2016, through January 24, 2022, and it includes two medications newly US Food and Drug Administration (FDA)-approved for OP treatment since the 2017 guideline.

METHODS

This guideline follows the ACR guideline development process and ACR policy guiding management of conflicts of interest and disclosures (https://www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines), including use of Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology (27,28) and adherence to Appraisal of Guidelines for Research and Evaluation (AGREE) criteria (29). Supplementary Appendix 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25240/abstract, includes a detailed description of the methods. Briefly, the Core Leadership Team (MBH, LR, MID, HAF, GG, SU) reviewed the 2017 ACR GIOP guideline clinical Patient/Intervention/Comparator/Outcomes (PICO) questions, modified and drafted new PICO questions in topic areas not covered previously (eg, abaloparatide, romosozumab, combination and sequential therapy) (see Supplementary Appendix 2, http://onlinelibrary.wiley.com/doi/10.1002/acr.25240/abstract). The Literature Review Team updated the systematic literature reviews for each of the previous PICO questions and/or performed new ones for new questions, graded the quality of evidence (high, moderate, low, very low), and produced the evidence report (see Supplementary Appendix 3, http://onlinelibrary.wiley.com/doi/10.1002/acr.25240/abstract). The resulting evidence was reviewed, and recommendations were formulated and voted on by an expert Voting Panel. A virtual Patient Panel of three patients with GIOP and one parent of a child treated with GCs reviewed the evidence with a co-principal investigator (LR) and provided patient perspectives and preferences for consideration by the Voting Panel. Voting Panel consensus required ≥70% agreement on both direction (for or against) and strength (strong or conditional) of each recommendation. Rosters of the Core Leadership Team,

This article is published simultaneously in *Arthritis & Rheumatology*.
Supported by the American College of Rheumatology.

[1]Mary Beth Humphrey, MD, PhD, Itivrita Goyal, MD, Emma Punni, MD: University of Oklahoma Health Sciences Center and U.S. Department of Veterans Affairs, Oklahoma City, Oklahoma; [2]Linda Russell, MD, Katherine Haseltine, MD: Hospital for Special Surgery, New York, New York; [3]Maria I. Danila, MD, MSc, MSPH, Basma Abdulhadi, MD, Lesley Jackson, MD, Iram Moledina, MD: University of Alabama at Birmingham and Birmingham VA Medical Center, Birmingham, Alabama; [4]Howard A. Fink, MD, MPH: Geriatric Research Education and Clinical Center, VA Health Care System, Minneapolis, Minnesota; [5]Gordon Guyatt, MD, Reza Mirza, MD: McMaster University, Hamilton, Ontario, Canada; [6]Michael Cannon, MD: Arthritis Consultants of Tidewater, Virginia Beach, Virginia; [7]Liron Caplan, MD, PhD: University of Colorado, Denver; [8]Sara Gore, ••, Charles Womack, ••: Oklahoma City, Oklahoma; [9]Jennifer Grossman, MD: UCLA Health, Los Angeles, California; [10]Karen E. Hansen, MD, MS, Sumona Saha, MD, MS: University of Wisconsin, Madison; [11]Nancy E. Lane, MD: UC Davis Health, Sacramento, California; [12]Nina S. Ma, MD: Children's Hospital Colorado, University of Colorado School of Medicine, Aurora; [13]Marina Magrey, MD: Case Western Reserve University, MetroHealth, Cleveland, Ohio; [14]Tim McAlindon, MD, MPH: Tufts University, Boston, Massachusetts; [15]Angela Byun Robinson, MD, MPH: Cleveland Clinic Foundation, Cleveland, Ohio; [16]Julia F. Charles, MD, PhD, Sharon Chou, MD: Brigham and Women's Hospital, Boston, Massachusetts; [17]Jonathan T. L. Cheah, MBBS: UMass Memorial Health and UMass Chan Medical School, Worcester, Massachusetts; [18]Tim Rinden, MD: University of Minnesota, Minneapolis; [19]Marat Turgunbaev, MD, MPH, Amy S. Turner, BS: American College of Rheumatology, Atlanta, Georgia; [20]Katherine Wysham, MD: VA Puget Sound Health Care System and University of Washington School of Medicine, Seattle, Washington; [21]Stacey Uhl, MS: ECRI Institute, Plymouth Meeting, Pennsylvania.

*Drs. Humphrey and Russell contributed equally to this work.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25240.

Address correspondence via email to Mary Beth Humphrey, MD, PhD, at marybeth-humphrey@ouhsc.edu.

Submitted for publication December 1, 2022; accepted in revised form July 3, 2023.

Literature Review Team, Voting Panel, and Patient Panel are included in Supplementary Appendix 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25240/abstract. This study did not involve human subjects and, therefore, approval from Human Studies Committees was not required.

RECOMMENDATIONS

How to interpret the recommendations

According to GRADE, a strong recommendation is usually supported by moderate- to high-certainty evidence, including randomized control trials, the recommended course of action would apply to all or almost all patients, and there is high confidence that the benefits of the intervention clearly outweigh the harms (or vice versa). In rare instances, a strong recommendation or best practices may be made with very-low certainty evidence if the recommendation is considered benign, low cost, and without harms.

A conditional recommendation is supported by lower certainty evidence, has uncertainty regarding the balance of benefits and harms, is sensitive to individual patient preferences, or has costs expected to impact the decision. Thus, conditional recommendations warrant shared decision-making with the patient. Notably, most evidence reviewed in this guideline is downgraded for indirectness because 1) identified studies in GIOP rely on a surrogate fracture risk marker, bone mineral density (BMD), because they were not powered for fracture outcomes and 2) available fracture data were exclusively or predominantly from general osteoporosis (OP) studies.

Key recommendations

1. As soon as possible after initiation of ≥2.5 mg/day GC treatment for >3 months, screening for fracture risk in patients ≥40 years of age should be assessed by using FRAX and by performing BMD using dual-energy x-ray absorptiometry (DXA) with vertebral fracture assessment (VFA) testing or spinal x-rays. BMD with VFA testing or spinal x-ray is advised in patients <40 years, as FRAX is not validated in this population.
2. Adequate age-appropriate dietary and supplemental intake of calcium and vitamin D, weight-bearing exercise, and avoidance of smoking and excessive alcohol intake is encouraged for all patients receiving GCs.
3. All adult patients with medium, high, or very high fracture risk should be offered OP therapy.
4. Oral bisphosphonates (BP) are strongly recommended over no treatment in high or very high fracture risk adults.
5. For adults with very high fracture risk, anabolic agents (parathyroid hormone [PTH] and PTH-related protein [PTHrP]) are conditionally recommended over antiresorptive agents (BP or denosumab [DEN]).
6. In adults ≥40 years of age at high risk of fracture, DEN or PTH/PTHrP are conditionally recommended over BP.
7. In adults at moderate risk of fracture, oral or intravenous (IV) BP, DEN, and PTH/PTHrP are conditionally recommended.
8. Include in decision-making that sequential OP treatment is recommended to prevent rebound bone loss and vertebral fractures after discontinuation of DEN, romosozumab, and PTH/PTHrP.

Table 1 presents the definitions of terms used in the recommendations and a synopsis of the age-based recommendations for fracture risk assessment and treatments.

Recommendations for fracture risk assessment (Figure 1)

For all adults (≥18 years old) initiating or continuing GC therapy ≥2.5 mg/day for >3 months, we strongly recommended initial clinical fracture risk assessment including symptomatic and asymptomatic fracture history, FRAX (age ≥40 only), and BMD with VFA or spine x-rays over no assessment (PICO 8.1–8.4).

These strong recommendations are based on good clinical practice and the need for clinicians to risk stratify patients beginning or continuing GC therapy, despite the low certainty of the evidence. Initial assessment should occur as soon as possible within 6 months of GC therapy initiation. Clinical fracture risk assessment includes dose, duration, and pattern of GC use, alcohol use, smoking history, hypogonadism, history of prior fractures (traumatic, fragility, asymptomatic), low body weight, significant weight loss, parental history of hip fracture, fall history, thyroid disease, hyperparathyroidism, rheumatoid arthritis, malabsorption, chronic liver disease, and inflammatory bowel disease (Figure 1). BMD with VFA or spinal x-rays are strongly recommended, and, for adults ≥40 years old, FRAX analysis is also recommended. (Figure 1). If prednisone dose is >7.5 mg daily, FRAX GC correction is recommended (Table 1, Figure 1) (21); however, even this adjustment may not correct for very high doses of GC (≥30 mg/day) (30). Additionally, FRAX does not incorporate falls, number or timing of fractures, or frailty that may put a person at higher risk of fracture. BMD assessment provides a strong predictor of fracture risk and serves as a baseline for reassessment because FRAX analysis is not validated for fracture risk reassessment during OP therapy. Trabecular bone score (TBS) provides a more sensitive measure of therapeutic responses to OP treatment (31). BMD measurement is strongly recommended for patients <40 years on GCs ≥2.5 mg/day with one or more osteoporotic risk factors. In this age group, z-scores ≤ –2.0 indicates low bone mass for age. Unlike *t*-scores, z-scores do not provide an

Table 1. Definitions of selected terms used in the recommendations and upgraded position statements for GIOP*

Term	Adults ≥40 years of age	Adults <40 years of age
MOF	Nontraumatic or pathological fractures of the spine, hip, wrist, or humerus	Nontraumatic or pathological fractures of the spine, hip, wrist, or humerus
Clinical fracture risk assessment	History of GC use, evaluation for falls, fractures, frailty, secondary causes of OP, FRAX with GC adjustment, BMD with VFA or spinal x-ray	History of GC use, evaluation for falls, fractures, frailty, secondary causes of OP, BMD with VFA or spinal x-ray (FRAX not validated at age <40 years)
Follow-up risk assessment during GC treatment	BMD with VFA or spinal x-ray every 1–2 years during OP therapy; BMD with VFA or spinal x-ray every 1–2 years after OP therapy is discontinued	BMD with VFA or spinal x-ray every 1–2 years during treatment; BMD with VFA or spinal x-ray every 1–2 years after OP therapy is discontinued
FRAX GC correction	If GC dose is >7.5 mg/day, multiply the 10-year risk of MOF by 1.15 and the hip fracture risk by 1.2†	Not applicable as FRAX is not validated in this age group
Very high fracture risk	Prior OP fracture(s) OR BMD *t*-score ≤−3.5 OR FRAX (GC-Adjusted) 10-year risk of MOF ≥30% or hip ≥4.5% OR high GC ≥30 mg/day for >30 days OR cumulative doses ≥5 g/y	Prior fracture(s) OR GC ≥30 mg/day OR cumulative ≥5 g/y
High fracture risk	BMD *t*-score ≤−2.5 but >−3.5 OR FRAX (GC Adjusted) 10-year risk of MOF ≥20% but <30% or hip ≥3% but <4.5%	–
Moderate fracture risk	FRAX (GC-Adjusted) 10-year risk of MOF ≥10 and <20%, hip >1 and <3% OR BMD *t*-score between −1 and −2.4	Continuing GC treatment ≥7.5 mg/day for ≥6 months AND BMD *z*-score < −3 OR significant BMD loss (more than the least significant change of DXA)
Low fracture risk	FRAX (GC-Adjusted) 10-year risk of MOF <10%, hip <1 %, BMD >−1.0	None of the above risk factors other than GC treatment
Recommended treatment strategy	**Adults ≥40 years at moderate, high, or very high risk of fracture**	**Adults <40 years at moderate or very high risk of fracture**
Calcium and vitamin D	Optimized intake of dietary and supplemental calcium and vitamin D based on age-appropriate US Recommended Dietary Allowances	
BP (Alendronate [oral], Risedronate [oral]; Ibandronate [oral/ IV], Zoledronic acid [IV])	We strongly recommend OP treatment for those at moderate, high, or very high risk of fracture. We **s**trongly recommend oral BP over no treatment in high and very high fracture risk due to fracture reduction in GIOP. We conditionally recommend IV BP, ROM, RAL over no treatment in high and very high risk of fracture. In moderate risk, we conditionally recommend BP, DEN, or PTH/PTHrP in no preferred order among these agents.	We conditionally recommend treatment for those at moderate or very high risk of fracture with oral or IV BP,‡ PTH/PTHrP,§ or DEN§#
PTH/PTHrP Agonists (TER, ABL, Anti-RANKL, DEN)	We conditionally recommend PTH/PTHrP over anti-resorptives in patients at very high risk of fracture. We conditionally recommend DEN§# or PTH/PTHrP over oral and IV BP in high risk of fracture. In moderate risk, we conditionally recommend BP, DEN, or PTH/PTHrP in no preferred order among these agents.	–
Selective estrogen receptor modifier (RAL), Anti-sclerostin (ROM)	We conditionally recommend IV BP, ROM, RAL over no treatment in high and very high risk of fracture. Except in patients intolerant of other agents, we conditionally recommend against RAL due to harms of VTE and fatal stroke or ROM due to uncertain harms with increased myocardial infarction, stroke and death.	We conditionally recommended against RAL due to harms of VTE and fatal stroke or ROM due to uncertain harms including increased myocardial infarction, stroke and death

* ABL = Abaloparatide; BMD = bone mineral density; BP = bisphosphonate; DEN = Denosumab; FRAX = Fracture Risk Assessment Tool; GC = glucocorticoid; MOF = major osteoporotic fracture; PTH = parathyroid hormone; PTHrP = PTH-related protein; RAL = Raloxifene; RANKL = Receptor activator of NF-κβ-Ligand; ROM = Romosozumab; TER = Teriparatide; VTE = venous thromboembolism.
† FRAX GC correction example: if hip fracture risk is 2.0% multiply by 1.2 for adjusted risk = 2.4%.
‡ Use with caution in patients who may become pregnant due higher potency and longer half-life in fetal bones.
§ Avoid in young adults with open growth plates.
# Use with caution in patients of child-bearing potential due to potential fetal harm. Avoid pregnancy for 5 months after last dose.

estimate of fracture risk because adults <40 years have low fracture risk at baseline.

This guideline did not include specific PICO questions concerning DXA or spinal imaging in children beginning or continuing chronic GC therapy, but the Voting Panel discussed this population. Despite uncertainty about initial DXA or screening spine radiographs, we recommend spine x-ray in children with back pain (32). However, the totality of a child's clinical presentation

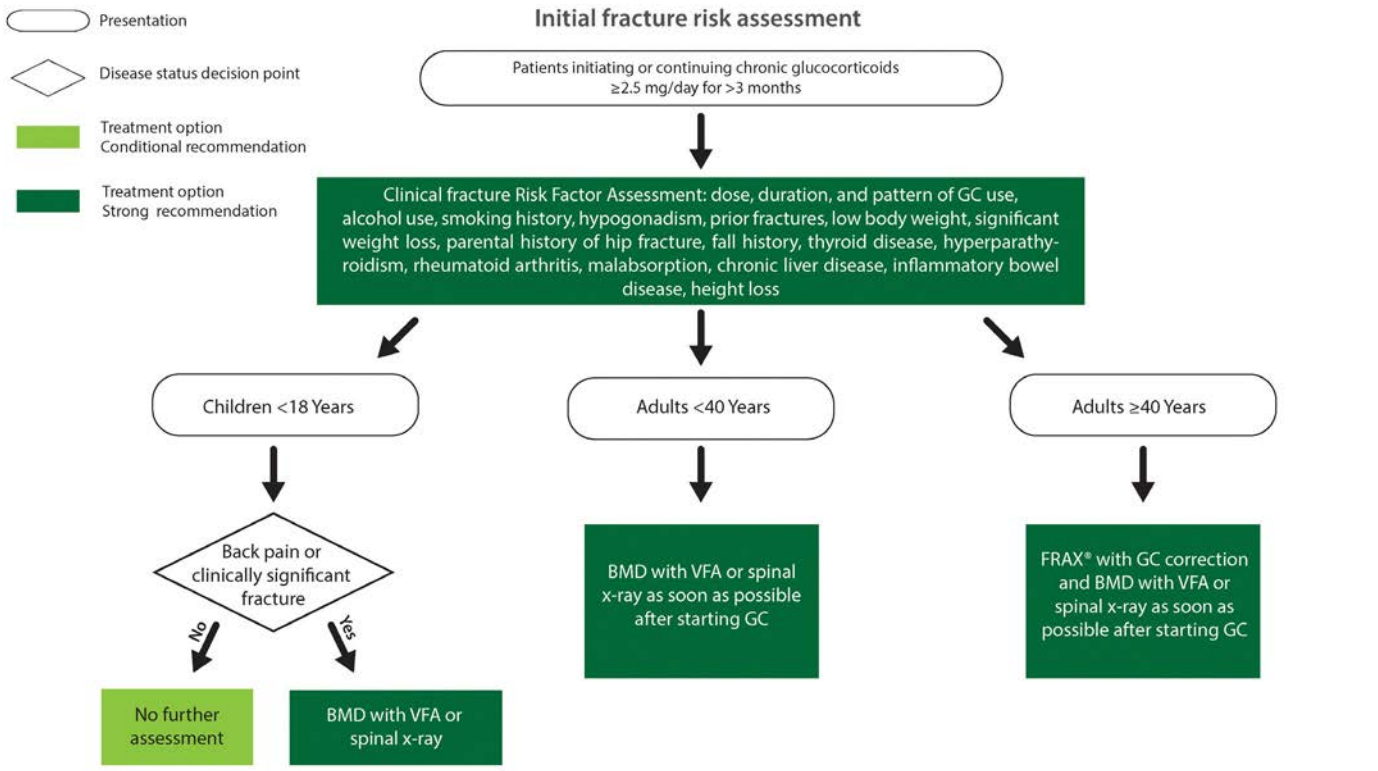


Figure 1. Initial fracture risk assessment. GC = glucocorticoid; MOF = major osteoporotic fracture; VFA = vertebral fracture assessment.

(eg, age at diagnosis, growth, body mass index [BMI], disease severity, GC dosing, BMD, symptomatic or asymptomatic vertebral compression fractures) should be taken into account when considering assessment for OP therapy (16).

As in prior guidelines, we used risk categories of low, moderate, and high using DXA and/or FRAX assessments (see Table 1). Similar to other recent OP guidelines (33-35) (United Kingdom National Osteoporosis Guideline Group [NOGG], American Association of Clinical Endocrinologists [AACE], Brazilian Society of Endocrinology and Metabolism [SBEM]), we further identified a very high risk group with prior osteoporotic fractures, very low BMDs, very high FRAX risks, or high daily dose or high cumulative doses of glucocorticoids.

Recommendations for reassessment of fracture risk (Figure 2)

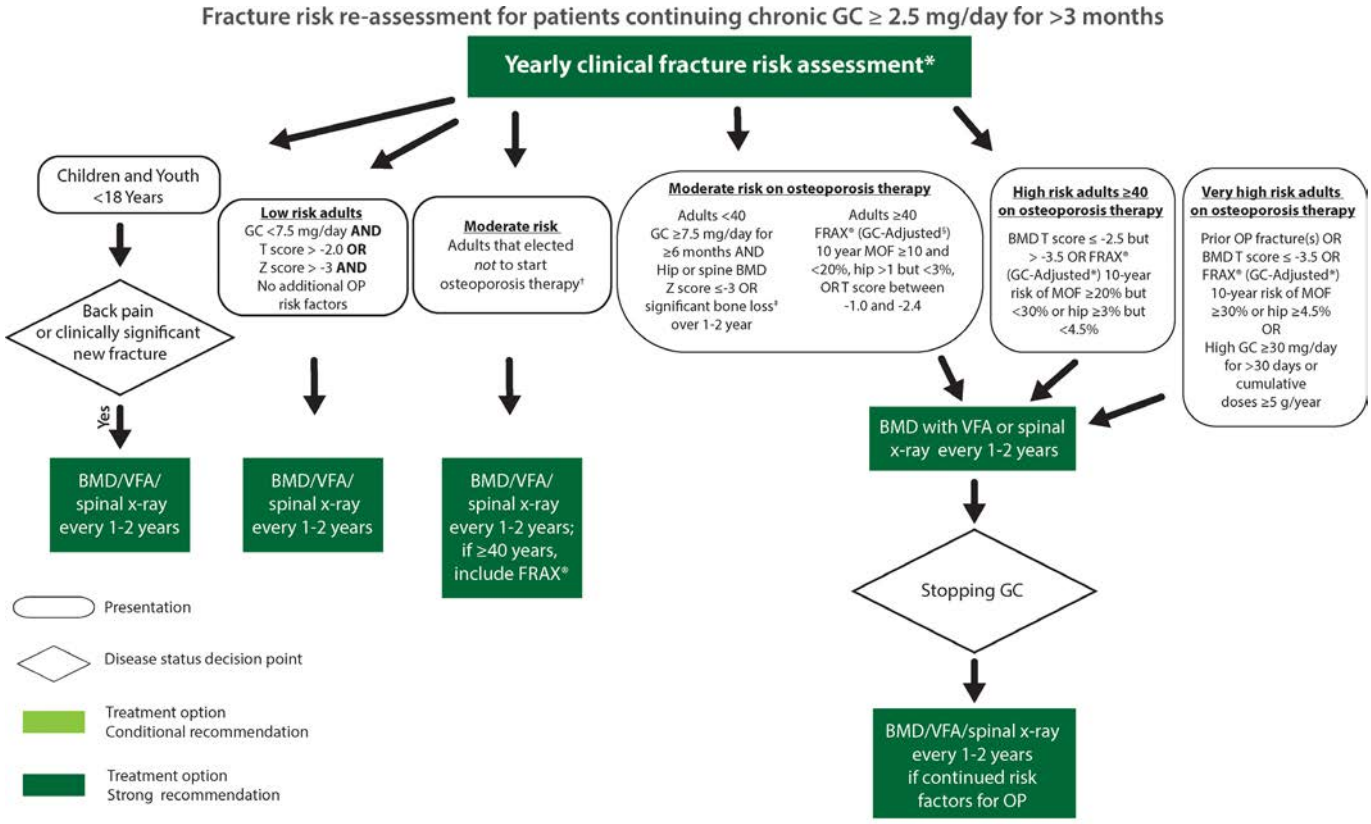
For adults continuing chronic GC ≥2.5 mg/day but <7.5 mg/day and assessed as low fracture risk, who were not recommended to start therapy, or moderate fracture risk who chose not to start OP therapy (except calcium and vitamin D), we strongly recommend fracture risk reassessment every 1 to 2 years (PICO 9.1–9.4).

Despite the low certainty of the evidence, this is a strong recommendation as good clinical practice. Fracture risk reassessment includes clinical fracture risk history, new symptomatic fractures, FRAX, BMD, VFA, and/or spine x-rays. Repeating DXA assessment every 1 to 2 years allows providers to detect the least significant BMD change according to their DXA machine, triggering the need to start OP therapy.

For adults continuing chronic GC ≥2.5 mg/day and assessed as moderate, high, or very high fracture risk who are continuing OP therapy ≥1 year, we strongly recommend fracture risk re-assessment every 1 to 2 years over no risk reassessment (PICO 9.5–9.12).

Despite the low certainty of the evidence, this is a strong recommendation as good clinical practice. Reassessment allows providers to determine if patients continuing GC and OP therapy are maintaining, gaining, or losing BMD, warranting possible changes in OP therapy. Yearly BMD assessment until a stable BMD is reached may be preferred in very high fracture risk patients.

For adults stopping GC and remaining at moderate, high, or very high fracture risk, we strongly recommend continuing OP therapy (PICO 12.1–12.6).



OP = osteoporosis; GC = glucocorticoids; FRAX® = Fracture risk assessment tool can only be used in adults ≥40 years; BMD = bone mineral density testing; *Clinical fracture risk assessment: dose duration and pattern of GC use, alcohol use, smoking history, hypogonadism, prior fractures, low body weight, significant weight loss, parental history of hip fracture, fall history, thyroid disease, hyperparathyroidism, rheumatoid arthritis, malabsorption, chronic liver disease, inflammatory bowel disease, height; †Moderate risk adults should be offered therapy but may choose not to be treated; ‡ > least significant decline according to DXA machine (typically 3-5%); §FRAX® GC correction for GC ≥7.5 mg/day example: if hip fracture risk is 2.0% multiply by 1.2 for adjusted risk = 2.4%

Figure 2. Fracture risk re-assessment for patients continuing chronic GC ≥2.5 mg/day for >3 months. DXA = dual-energy x-ray absorptiometry; MOF = major osteoporotic fracture; VFA = vertebral fracture assessment. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25240/abstract.

Recommendations for initial treatment (Table 2, Figure 3)

For all adults and children beginning or continuing chronic GC at a dose of ≥2.5 mg/day for >3 months, we conditionally recommended optimizing age appropriate dietary and supplemental calcium and vitamin D, in addition to lifestyle modifications (PICO 1.1–1.3, 2.1–2.3, 3.1–3.3, 4.1–4.3, 5.1–5.3, 6.1–6.3, 7.1–7.4).

The evidence for calcium and vitamin D supplementation for fracture reduction in GIOP is low to very low. Dietary and supplemented elemental calcium intake of up to 1,000 to 1,200 mg daily is recommended for adults (36) and between 1,000 and 1,300 mg daily based on age of the child. Serum vitamin D levels should be monitored, and vitamin D supplemented to maintain serum vitamin D 25(OH)D levels ≥30 to 50 ng/mL; 600 to 800 IU daily or more is typically required. Lifestyle modifications include smoking cessation, limiting alcohol to ≤2 servings a day, eating a balanced diet, maintaining weight in the recommended range, and performing regular weight-bearing or resistance training exercises. All subsequent recommendations refer to adults and children beginning or continuing chronic GCs at a dose of ≥2.5 mg/day for >3 months and assume the use of calcium, vitamin D, and lifestyle modifications.

For adults ≥40 years with high or very high fracture risk, we strongly recommended treatment with OP therapy over treatment with calcium and vitamin D alone (PICO 1).

For adults ≥40 years with very high fracture risk, we conditionally recommend PTH/PTHrP over anti-resorptives (BP or DEN) (PICO 1.13c, 1.14c, 1.15c, 1.18c, 1.19c, 1.20c).

Compared to alendronate, teriparatide increased lumbar and hip BMD and decreased vertebral but not nonvertebral fractures at 36 months in GIOP (37,38). Bone anabolic effect is blunted when treatment follows anti-resorptive therapy.

For adults ≥40 years with high or very high fracture risk, we strongly recommended oral BP (16) over no treatment (PICO 1).

A strong recommendation for oral BP is based on studies showing a reduction in total and vertebral fractures at 24 months and increased hip and lumbar spine BMD compared to calcium and vitamin D alone in GIOP (evidence report, Appendix S3, page 16).

Table 2. Recommendations for initial treatment for prevention of GIOP in adults beginning long-term GC therapy*

Recommendations for patients taking prednisone ≥2.5 mg/day for >3 months	Certainty of evidence	PICO evidence report basis	Evidence Report, pp
For adults and children beginning or continuing chronic GC treatment at low, moderate, high, or very high risk of fracture, we conditionally recommend optimizing dietary and supplemental calcium and vitamin D in addition to lifestyle modifications	Low or very low	1.1a,b,c–1.3a,b,c, 2.1–2.3, 7.16–7.26	6–8, 47–48, 63–65, 141–144,148–151
In adults ≥40 years†			
For adults ≥40 years with high or very high fracture risk, we strongly recommend OP therapy over no treatment. Agents to use include oral BP,‡ IV BP,§ PTH/PTHrP,§ DEN,§ RAL, or ROM.	Low or very low	1.4c–1.28c	6–50
For adults ≥40 years with very high fracture risk, we conditionally recommend PTH/PTHrP over anti-resorptive (DEN, BP) treatment.	Low	1.13c–1.20c	49–50
For adults ≥40 years with high fracture risk, we conditionally recommend PTH/PTHrP or DEN over BP treatment.	Low	1.13c–1.20c	49–50
For adults ≥40 years with high or very high fracture risk, we strongly recommend oral BP over no treatment.	Low	1.4c	8–18
For adults ≥40 years with high or very high fracture risk, we conditionally recommend using ROM or RAL in patients intolerant of other agents.	Very low	1.16c, 1.21c, 1.28c	50
For adults ≥40 years with high or very high fracture risk, we conditionally recommend against using two different OP medications.	Very low	1.29–1.35	53–62
For adults ≥40 years with moderate fracture risk, we conditionally recommend against ROM except for in patients intolerant of other agents, due to risk of myocardial infarction, stroke, or death.	Very low	1.12b, 1.16b, 1.17b, 1.21b–1.25b, 1.28b	40–41, 44–47
For adults ≥40 years with low fracture risk, we strongly recommend against OP medications due to known risk of harms and no evidence of benefit.	Very low	4.4a–4.13a	91–101
Adults receiving high-dose GC (initial dose ≥30 mg/day for >30 days or cumulative dose ≥5 g in 1 year)			
We conditionally recommend treating with PTH/PTHrP over anti-resorptives.	Low	6.1b–6.19a	120–141
Oral BP are strongly recommended over no treatment.	Low	6.1b–6.19a	120–141
IV BP and DEN are conditionally recommended over no treatment.	Low	6.1b–6.19a	120–141
RAL and ROM are conditionally recommended in those intolerant of other agents.	Low	6.1b–6.19a	120–141
In adults <40 years†			
Adults <40 years with moderate fracture risk, we conditionally recommend oral or IV BP,¶ DEN,¶ or PTH/PTHrP therapy.	Low or very low	2.4–2.22, 3.4–3.17	65–76, 79–84
Adults <40 years with moderate fracture risk, we conditionally recommend against using ROM due to risk of myocardial infarction, stroke, or death.	Very low	2.9, 3.9	70, 87
For adults with solid organ transplants, glomerular filtration rate ≥35 mL/min, and no evidence of CKD-MBD# or hyperparathyroidism			
We conditionally recommend expert evaluation for CKD-MBD in renal transplant recipients.	Low	5.1–5.26	103–118
We conditionally recommend treatment with oral or IV BP, DEN, PTH/PTHrP, or RAL based on individual patient factors.	Low	5.1–5.26	103–118
We conditionally recommend against using ROM due to risk of myocardial infarction, stroke, or death.	Very low	5.9	112
Children ages 4–17 years treated with GCs for >3 months (low and moderate risk)			
We conditionally recommend optimization of dietary and supplementation of calcium and vitamin D as recommended by the US RDA depending on the age of the child.	Very low	7.1a–7.4a	141–144
We conditionally recommend against starting oral or IV BP due to low risk of OP fractures in this age group.	Very low	7.5a	144
Children ages 4-17 years with an osteoporotic fracture who are continuing treatment with GCs at a dose of ≥0.1 mg/kg/day for >3 months (high risk)			
We conditionally recommend treating with an oral or IV BP.	Very low	7.1b–7.2b	148–153

* BP = bisphosphonate; CKD-MBD = chronic kidney disease–mineral and bone disorder; DEN = denosumab; GC = glucocorticoid; GIOP = GC-induced OP; IV = intravenous; OP = osteoporosis; PICO = Patients, Intervention, Comparison, Outcome; PTH/PTHrP = parathyroid hormone/parathyroid hormone–related protein; RAL = raloxifene; RDA = Recommended Dietary Allowances; ROM = romosozumab.
† In addition to calcium, vitamin D, and lifestyle modifications.
‡ Strong recommendation based on fracture data.
§ Conditional due to a lack of fracture data.
¶ Only for patients who are not planning on pregnancy during the OP treatment period or are using effective birth control if sexually active.
# Includes osteomalacia, adynamic bone disease, osteitis fibrosa cystica, mixed uremic osteodystrophy.

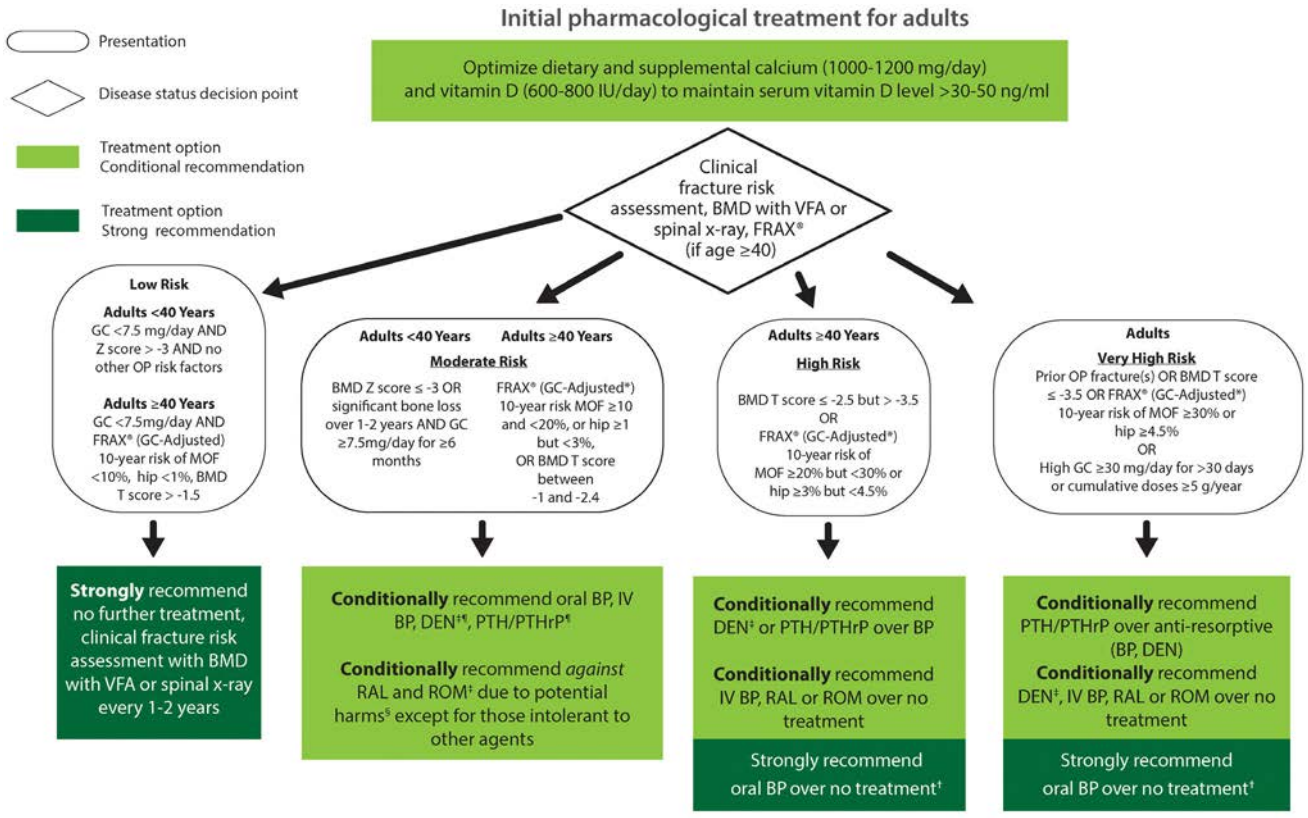


Figure 3. Initial pharmacological treatment for adults. BMD = bone mineral density; DVT = deep vein thrombosis; GC = glucocorticoid; GIOP = GC-induced OP; PE = pulmonary embolism; VFA = vertebral fracture assessment. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25240/abstract.

For adults ≥40 years with high fracture risk, we conditionally recommend PTH/PTHrP or DEN over BP (PICO 1.4c–1.28.c).

For adults ≥40 years with high fracture risk, we conditionally recommend IV or oral BP, PTH/PTHrP, or DEN over Raloxifene (RAL) or Romosozumab (ROM) (PICO 1.4c–1.28.c).

High-certainty evidence indicates that oral BP prevents vertebral fractures in GIOP (39) and warrants a strong recommendation for use here. Compared to oral BP, PTH is superior at increasing BMD 24 and 36 months and prevented vertebral fractures at 36 months (37). In the very high risk group, providers may recommend PTH/PTHrP as initial treatment because anabolism is blunted in patients previously treated with BP (40). IV BP and DEN GIOP trials have not been powered to detect reductions of GIOP fractures and instead use a surrogate endpoint of BMD changes (41–43). However, the relationship between increases in BMD and a decrease in vertebral fractures is inconsistent and may account for only 25% of overall reduction in fracture risk (44). Evidence for fracture reduction of PTHrP, DEN, RAL, and ROM therapies have been demonstrated in general OP but not GIOP, leading to downgrading the evidence to low or very low certainty evidence. However, DEN and PTH show superior BMD gains in GIOP compared to BP and may be preferred in patients with high risk.

Compared to BP and RAL, PTH/PTHrP, DEN, and ROM require sequential therapy with an anti-resorptive agent to prevent bone losses. Discontinuation of DEN must be followed by a BP beginning at 6 to 7 months after the last DEN dose to prevent rapidly progressive vertebral fractures. Additionally, IV BP, DEN, and ROM have increased risk of atypical femur fractures and osteonecrosis of the jaw compared to oral BP (45). Due to RAL harms of venous thrombotic embolism events (pulmonary embolism/deep vein thrombosis [PE/DVT]) and fatal stroke and association of ROM with increased myocardial infarction, stroke, and death, these therapies should be reserved for those unable to tolerate other agents (46,47). The panel recommends initial treatment choice be informed by patient co-morbidities and preferences regarding costs, burden of injections, and the need for sequential therapy (48).

In adults ≥40 years with high and very high fracture risk, we conditionally recommend against using multiple OP therapies at the same time (PICO 1.29–1.35).

Very low level evidence does not support using combination therapy (eg, PTH/PTHrP and DEN, PTH/PTHrP and BP) in GIOP. In patients with postmenopausal OP, studies have shown synergistic increases in BMD with combination of PTH with IV BP (49), PTH with RAL (50), and PTH and DEN (51). However, based on the added cost, the possibility of greater side effects, and the lack of fracture evidence, combination therapy is not currently recommended.

For all adults with moderate fracture risk, we conditionally recommend oral or IV BP, PTH/PTHrP, or DEN over no treatment (PICO 1.4b–1.28.b, 2.4b,c–2.17b,c).

In all adults with moderate fracture risk, we conditionally recommend against ROM and RAL therapies except in those intolerant of other OP medications, due to possible life-threatening harms, including thrombosis, fatal stroke, major cardiovascular events, and death (PICO 1.6b, 1.10b, 1.12b, 1.16b, 1.17b, 1.21b, 1.22b, 1.23b, 1.24b, 1.25b, 1.28b, 2.9, 2.14, 2.18, 2.21).

Multiple studies have shown that 12 months of ROM followed by an anti-resorptive agent (BP or DEN) for 12 month prevents fractures in patients with postmenopausal OP when compared to anti-resorptive agent only (52–54). There is uncertainty concerning the cardiovascular risk, including myocardial infarction, stroke, and death related to ROM (47,55). However, until longer-term pharmacovigilance data become available, ROM should not be started in patients with a myocardial infarction or stroke within 12 months. Shared decision-making between patients and clinicians is needed to determine if benefits outweigh the risks in patients with other cardiovascular risk factors that may be untreated including hyperlipidemia, hypertension, and smoking. For RAL, a meta-analysis of nine trials (24,523 postmenopausal women) found that raloxifene was associated with an increased risk of DVT and PE (odds ratio [OR] 1.5, 95% confidence interval [CI] 1.1–2.1 and OR 1.9, 95% CI 1.0–3.5, respectively) (56). In the Raloxifene use for the heart (RUTH) trial, RAL were not associated with overall stroke risk but was associated with fatal stroke (59 vs 39 events, hazard ratio [HR] 1.49, 95% CI 1.0–2.2, absolute risk increase of 0.7 per 1000 woman-years) compared with placebo (46).

In adults with low fracture risk, we strongly recommend against adding oral or IV BP, PTH/PTHrP, RAL, DEN, or ROM (PICO 1.4a–1.28a).

Adults <40 years have low fracture risk and have significant capacity to rebuild BMD losses induced by chronic GC therapy. OP therapy should not be started in this low-risk group (17,57). This strong recommendation is based on low certainty evidence of anti-fracture benefit in this low fracture risk group, coupled with clear potential harms such as osteonecrosis of the jaw (BP, DEN, ROM), atypical femur fractures (BP, DEN, ROM), PE, DVT, and fatal stroke (RAL), myocardial infarction, stroke, and death (ROM), or requirements for sequential therapy (PTH/PTHrP, DEN, ROM). Adults >40 years on low-dose steroids that meet low risk criteria have uncertain benefit from osteoporosis therapy.

Recommendations for special populations of patients beginning long-term GC therapy at very high risk for fracture (Table 2)

For adults ≥40 years at very high fracture risk due to treatment with one or more courses of high-dose GC therapy (mean dose prednisone equivalent ≥30 mg daily for ≥30 days) or cumulative GC dose ≥5 g over 1 year, we conditionally recommend treating with PTH/PTHrP over anti-resorptive agents regardless of FRAX score or BMD. We strongly recommend oral BP over no treatment and conditionally recommend an IV BP, DEN, RAL or ROM over no treatment.

The relative risk for vertebral fracture was 14 and for hip fractures was 3 with a dose of ≥30 mg per day and ≥ 5 g of cumulative use (10).

For adults <40 years receiving one or more courses of high-dose GC therapy (mean dose prednisone equivalent ≥30 mg daily for ≥30 days) or cumulative GC dose ≥5 g over 1 year, we conditionally recommend oral or IV BP, PTH/PTHrP, DEN. We conditionally recommended against RAL/ROM (PICO 6.4a,b–6.24a,b).

In this younger population, PTH/ PTHrP and ROM should only be used in adults with closed growth plates. DEN should be used with caution in patients with open growth plates.

For patients who can become pregnant at moderate or high risk of fracture, we conditionally recommend treating with oral or IV BP, DEN, or PTH/PTHrP (PICO 2).

OP therapy is not contraindicated in patients who can become pregnant but should be used with effective birth control if sexually active. BP are avidly taken up by the fetal skeleton as shown in animal models and have a long half-life of BP in adult bones with unclear side effects for the fetal skeleton (58). Risedronate and ibandronate have shorter skeletal half-lives among BP and may be preferred in this setting. DEN and PTH/PTHrP may also be used if growth plates have closed. However, DEN may cause fetal harm and is contraindicated in pregnancy. Avoid pregnancy for 5 months after the last dose of DEN.

For adults with solid organ transplants and an estimated glomerular filtration rate (eGFR) ≥35 mL/min who are continuing chronic GC treatment, we conditionally recommend treatment with BP, DEN, PTH/PTHrP, or RAL, based on individual patient factors over no treatment (PICO 5.4-5.26).

In this solid organ transplant population, we conditionally recommend against using ROM due to potential harms in this population (PICO 5.9, 5.21, 5.16).

This group of patients is typically considered at increased risk of fracture regardless of BMD, due to the known risk of OP associated with solid organ transplantation and anti-rejection medications. The overall certainty of evidence for treatment in this population is low, and numerous potentially influential individual patient factors need to be weighed when selecting treatment.

For adult renal transplant recipients on chronic GC treatment, we conditionally recommend metabolic bone disease expert evaluation for chronic kidney disease–mineral and bone disorder (CKD-MBD).

In patients with stage IV and V CKD, renal osteodystrophy, including adynamic bone disease, osteomalacia, osteitis fibrosa cystica, and mixed uremic osteodystrophy, is nearly universal (59). Bone-specific alkaline phosphatase, intact PTH, and bone biopsy may exclude renal osteodystrophy. BP should generally not be used if eGFR <35 mL/min. Once renal osteodystrophy and hyperparathyroidism is excluded, no dose adjustment is needed when prescribing DEN, PTH/PTHrP, or ROM. However, if eGFR is <30 mL/min, DEN is not contraindicated but induces prolonged and more severe hypocalcemia (60).

The panel recommended that patients without hyperparathyroidism and eGFR ≥30 mL/min could use vitamin D3 (cholecalciferol) or vitamin D2 (ergocalciferol) instead of biologically active forms of vitamin D (calcitriol, paricalticol, or doxercalciferol). Patients with GFR <30 mL/min might require biologically active VitD to maintain neutral calcium balance.

For children and youth ages 4 to 17 years treated with GCs for >3 months who are at low or moderate risk for fracture, optimization of age-appropriate dietary and supplemental calcium and vitamin D to fulfill the Recommended Daily Allowance is conditionally recommended in addition to an exercise program. We conditionally recommend against starting OP therapy due to the low risk of osteoporotic fractures in children and youth ages 4 to 17 years (PICO 7.1a–7.5a).

For children and youth ages 4 to 17 years with an osteoporotic fracture who are continuing treatment with chronic GC at a dose of ≥0.1 mg/kg/day for >3 months, treating with an oral or IV BP is conditionally recommended over no treatment. (PICO 7.1b-7.2.b)

This conditional recommendation to treat with oral or IV BP to prevent recurrent fractures is based on low-certainty evidence. Depending on the specific disease or cause of pediatric OP, there is uncertainty about when and how to screen, and depending on the guidelines, it requires a history of clinically significant fracture(s), defined as ≥1 vertebral fractures, ≥2 long bone fractures prior to age 10 years, or ≥3 long bone fractures up to age 19 years (61,62). Twelve percent of children with rheumatic conditions on chronic GC averaging doses of 0.94 ± 0.84 mg/kg/day for 6 months who then tapered to 0.06 ± 0.12 mg/kg/day between 30 months and 36 months had vertebral fracture in the three years following GC initiation (14). The same study found that every 0.5 mg/kg increase in average daily GC dose was associated with a two-fold increased fracture risk (HR 2.0, 95% CI 1.1–3.5). Other OP therapies are understudied in this young age group with open growth plates.

Recommendations for initial treatment failure

For adults continuing GC treatment who have had an osteoporotic fracture ≥12 months after starting OP therapy, or who have had a significant loss of BMD (eg, greater than the least significant change per their DXA machine) after 1 to 2 years of OP treatment, we conditionally recommend changing to another class of OP medication over not switching the class of OP medication (PICO 10.1–10.9).

If oral BP is the first OP therapy and suboptimal adherence or poor absorption is suspected, based on low certainty evidence, we conditionally recommend treatment with IV BP, DEN, ROM, or PTH/PTHrP. Of note, use of PTH/PTHrP after long-term BP treatment has blunted anabolic response but still increases BMD. If DEN is the first agent, switching to PTH/PTHrP may lead to transient bone losses in the hip and spine and is not recommended (63–65); however, PTH/PTHrP followed by DEN leads to continued BMD increases (66,67) (Figure 4).

Recommendations for treatment when GC are discontinued (Figure 5)

For adults taking OP therapy and discontinuing GC therapy, with no new fragility fracture and a current BMD *t*-score ≥–2.5, we strongly recommended stopping current OP therapy and continuing calcium and vitamin D. However, sequential therapy is strongly recommended after stopping DEN, PTH/PTHrP, and ROM (Figure 5**) (PICO 11.1, 13.1–13.4).**

This recommendation is based on low-certainty evidence and on the balance of benefits and harms of continued treatment with OP medication. BP and RAL can be discontinued without need for sequential therapy. DEN, PTH/PTHrP, and ROM should be transitioned to anti-resorptive therapy, but the best formulation and duration of treatment is unclear at this time (68–70). Discontinuation of DEN can be associated with vertebral fractures that may be averted if a BP is started 6 to 7 months after the last DEN administration (41,42). Significant bone loss may occur after discontinuation of PTH/PTHrP, although anti-fracture efficacy may persist for 18 months; therefore, anti-resorptive therapy is recommended. ROM can be followed by DEN or BP (71).

For adults ≥40 years discontinuing GC therapy and continuing to be at high risk of fracture (BMD *t*-score ≤–2.5, or history of a fragility fracture occurring after ≥12 months of therapy), we conditionally recommend continuing current OP therapy or switching to another class of OP medication (PICO 13.5–13.6).

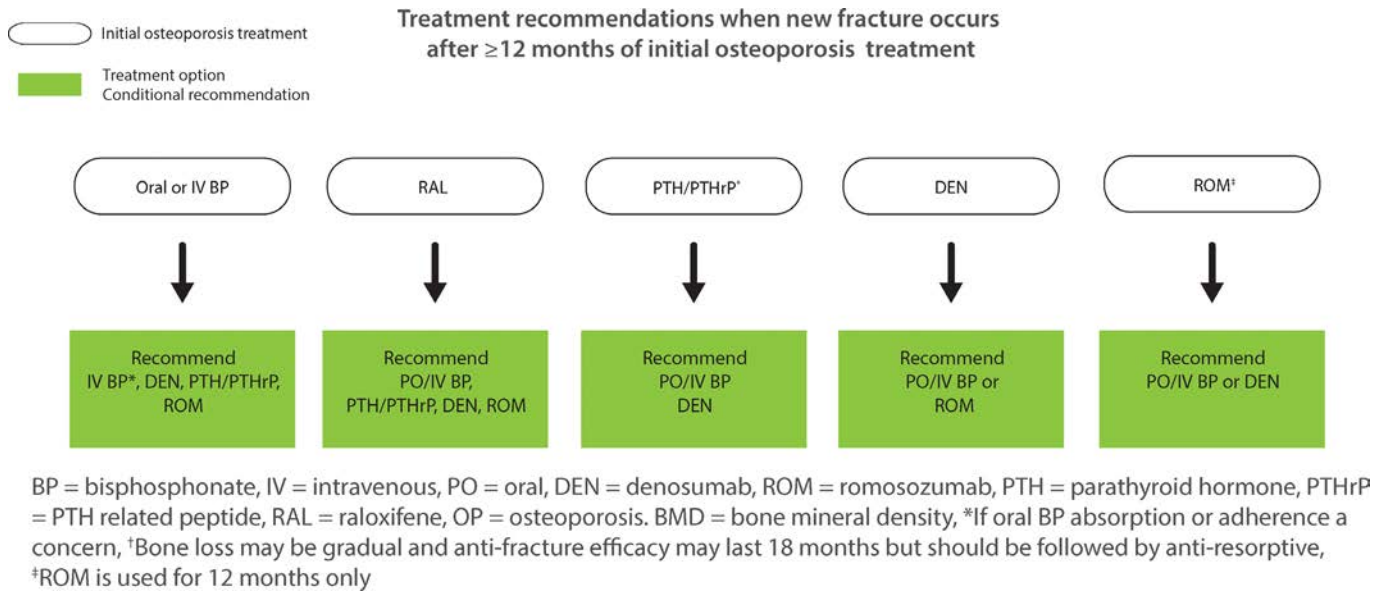


Figure 4. Treatment recommendations when new fracture occurs after ≥12 months of initial osteoporosis treatment. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25240/abstract.

For adults ≥40 years continuing chronic GC who discontinue DEN, we strongly recommend starting an anti-resorptive over not starting OP medication (PICO 13.1, 13.3, 13.5).

DEN remains effective at longer than 10 years in patients with postmenopausal OP. However, discontinuation of DEN after two or more doses has been associated with rapid loss of BMD and development of new vertebral compression fractures as soon as 7 to 9 months after the last DEN dose. As such, 6 to 9 months after the last dose of DEN, BP or ROM therapy is recommended (41,42). The precise timing, dose, and duration of BP or ROM use after DEN cessation is still under study, but treatment for at least 1 year with an oral BP or 1 to 2 years of IV BP seems prudent, until additional research is available

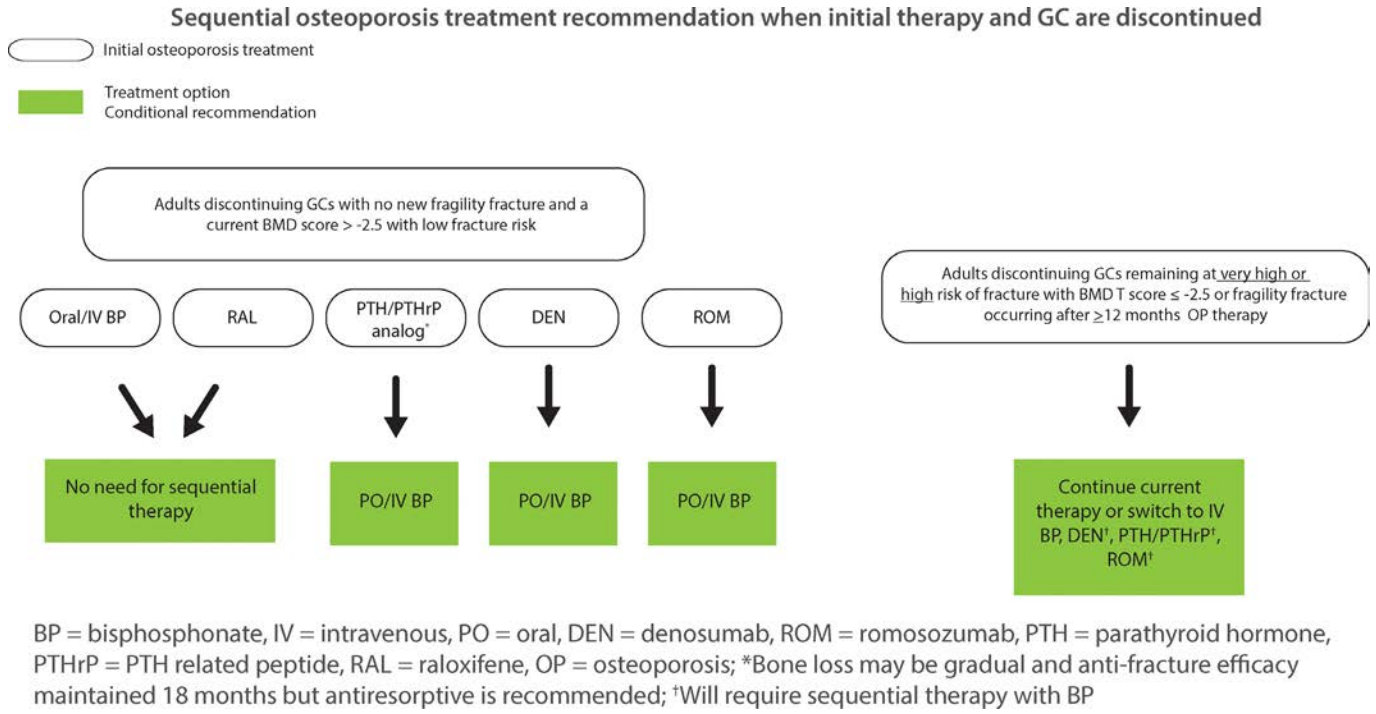


Figure 5. Sequential osteoporosis treatment recommendation when initial therapy and glucocorticoids (GCs) are discontinued. BMD = bone mineral density. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25240/abstract.

(69,71,72). If ROM is used after DEN, then it must be followed with a course of BP.

For adults ≥40 years discontinuing chronic GC treatment who have completed a course of a PTH/PTHrP, we conditionally recommend starting BP over not starting an OP medication (PICO 13.4, 13.6).

Discontinuation of PTH/PTHrP medication may lead to gradual loss of bone gained over 12 to 18 months, which can be prevented by treatment with BP or DEN (73). If DEN is used sequentially after discontinuation of PTH/PTHrP, then a BP should be started at the completion of DEN therapy (Figure 4). Therefore, BP therapy is recommended after discontinuation of PTH/PTHrP.

DISCUSSION

The objective of this updated ACR guideline for the prevention and treatment of GIOP (25) is to aid clinicians who prescribe GC, across all specialties, to best identify GC-treated patients who would benefit from prevention and treatment of GIOP. The overall goal is to reduce the number of fractures and their adverse consequences in this patient population, while minimizing harm due to medications. Fractures, especially hip and vertebral fractures, are associated with increased mortality, and patients frequently do not return to their baseline mobility (2,4,74). This guideline now addresses several new areas compared to the 2017 guideline: 1) Previously only fracture data were considered; with this guideline, both fracture reduction and BMD outcomes were considered because most GIOP studies are not powered for fracture outcomes (however, if fracture outcomes were not available, BMD data were evaluated and evidence downgraded to very low certainty); 2) a very high fracture risk category was added; 3) a preference for anabolic agent as initial OP therapy in very high fracture risk was made; 4) a need for sequential therapy after DEN, ROM, and PTH/PTHrP was made; and 5) we recommended the choice of therapies be based on clinician and patient preferences and comorbidities, rather than rank ordering the available OP therapies.

We risk stratified patients as low, moderate, high or very high risk of fracture based on FRAX 10-year probability and DXA *t*- or z-scores (Table 1). Similar to other organizational postmenopausal OP guidelines (AACE, SBEM, UK, and National Osteoporosis Foundation (NOF)) (33–35), we have now included a very high fracture risk category (prior OP fracture(s) or BMD *t*-score ≤–3.5 or FRAX (GC-Adjusted) 10-year risk of MOF ≥30% or hip ≥4.5% or high GC ≥30 mg/day for >30 days or cumulative doses ≥5 g/year) (Figures 2, 3, and5). These cut points were used to stratify PICO questions and weigh potential benefits versus harms of OP therapy. For prednisone-equivalent doses >7.5 mg/day, a FRAX GC correction is recommended and is achieved by multiplying the risk of MOF by 1.15 and the risk of hip fractures by 1.2. Fracture risk is considered highest for patients treated with very high (≥30 mg/day) or large cumulative GC doses (≥5 g/year) (75).

Risk assessment in children, youths, and adults <40 years is not as clear because these populations have substantially lower fracture risk than older adults. BP treatment for children was recommended only after a diagnosis of pediatric OP, which requires a clinically significant history of vertebral or long bone fractures. For children with a GC-associated fracture who continue to take high-dose GC therapy (>0.1 mg/kg/day), BP therapy is warranted.

For adults ≥40 years, the panel voted to give clinicians the ability to select an OP therapy based on the patient's specific comorbidities and preferences, BMD values, fracture history, and other characteristics, rather than rank ordering the medication recommendations. Fracture prevention data in GIOP is currently limited to oral BP and PTH. Anabolic agents may be the preferred initial therapy for those at very high risk for fracture based on BMD and vertebral fracture prevention superiority compared to anti-resorptives in patients with very high risk postmenopausal OP. Of note, abaloparatide and ROM are not approved in GIOP, and we recognize it may be difficult to access these medications for GIOP.

The panel specifically noted that the potential harms of RAL (venous thromboembolism [VTE] and fatal stroke) and ROM (major myocardial infarction, stroke, and death) would often favor the other available options when possible.

The panel emphasized the need for shared decision-making with patients to ensure they understand that some OP therapies (DEN, PTH/PTHrP, ROM) require another course of anti-resorptive OP therapy to prevent rapid bone loss and vertebral fractures (76,77). Discontinuation of DEN without the addition of anti-resorptive therapy is associated with vertebral fractures occurring as soon as 7 to 9 months after the last dose (76,77). Until the optimal therapy strategy is determined, many experts favor starting BP therapy 6 to 7 months after discontinuation of DEN for at least 1 year (78). Although the use of PTH after DEN causes transient loss of hip BMD, these drugs have been successfully cycled with increases in BMD (41,67). It is important that clinicians, patients, and/or their care partners understand and discuss the need for additional OP therapy after completing DEN, PTH/PTHrP, or ROM therapy.

The use of OP medications in patients after kidney transplant and with CKD was addressed in this guideline. When eGFR <35 mL/min, the risk of renal osteodystrophy is significantly increased, including adynamic bone disease, osteomalacia, osteitis fibrosa cystica, and mixed uremic osteodystrophy. As such, MBD expert evaluation for CKD-MBD is conditionally recommended to exclude these conditions. Once excluded, no dose adjustment is needed when prescribing DEN, PTH/PTHrP, or ROM, but BP should be avoided. Use of DEN in this group may lead to prolonged and more severe hypocalcemia (60).

A limitation of this guideline is the lack of fracture data in GIOP-specific clinical trials and population studies. As such,

general OP population clinical trials data were reviewed when GIOP data were not available. This introduced indirectness into the certainty of the evidence and imprecision in the estimate of benefits for treatment in the GIOP population. Because of these limitations, most of the recommendations in this guideline are conditional.

Future studies in the treatment of GIOP should be powered to assess fracture risk reduction. Studies should focus on children and patients with CKD stage 4 and 5. As part of risk assessment, studies should explore the use of quantitative computed tomography (CT), bone finite element analysis from CT scans, and BMD measurements from CT colonography. It would be helpful to have validated fracture prediction scores for patients aged <40 years. More studies are needed to better identify the patient populations that might benefit from combination therapy and sequential therapy in GIOP. Additional studies are required to determine the best treatment options and duration of therapy after discontinuation of DEN. In conclusion, GIOP remains a common and challenging clinical scenario that is frequently unrecognized and undertreated. By systematically synthesizing the current knowledge and available clinical trials, we have provided an updated guideline to help clinicians best care for patients requiring long-term GC use.

ACKNOWLEDGMENTS

We thank Dr. Madhumathi Rao for her review of the evidence and participation in the round 1 voting. We thank the patient who (along with authors Sara Gore and Charles Womack) participated in the Patient Panel meeting: Meredith Cathryn Rainey. We thank the ACR staff, including Regina Parker for assistance in coordinating the administrative aspects of the project and Cindy Force for assistance with manuscript preparation. We thank Janet Waters for her 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. All authors approved of the final version to be published. Drs. Humphrey and Russell had full access to all the data in the study and take full responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Humphrey, Russell, Danila, Fink, Uhl, Guyatt, Turner

Acquisition of data. Uhl, Abdulhadi, Charles, Cheah, Chou, Goyal, Haseltine, Jackson, Mirza, Moledina, Punni, Rinden, Turgunbaev, Wysham

Analysis and interpretation of data. Humphrey, Russell, Danila, Fink, Uhl, Guyatt, Turner, Uhl, Abdulhadi, Charles, Cheah, Chou, Goyal, Haseltine, Jackson, Mirza, Moledina, Punni, Rinden, Turgunbaev, Wysham, Cannon, Caplan, Grossman, Hansen, Lane, Ma, Magrey, McAlindon, Robinson, Saha, Gore, Womack

REFERENCES

1. Johnston CB, Dagar M. Osteoporosis in older adults. Med Clin North Am 2020;104:873–84.
2. Katsoulis M, Benetou V, Karapetyan T, et al. Excess mortality after hip fracture in elderly persons from Europe and the USA: the CHANCES project. J Intern Med 2017;281:300–10.
3. Sattui SE, Saag KG. Fracture mortality: associations with epidemiology and osteoporosis treatment [review]. Nat Rev Endocrinol 2014; 10:592–602.
4. Schousboe JT. Epidemiology of Vertebral Fractures. J Clin Densitom. 2016;19(1):8–22.
5. Fardet L, Petersen I, Nazareth I. Monitoring of patients on long-term glucocorticoid therapy: a population-based cohort study. Medicine (Baltimore) 2015;94:e647.
6. Van Staa TP, Leufkens HG, Abenhaim L, et al. Oral corticosteroids and fracture risk: relationship to daily and cumulative doses. Rheumatology (Oxford) 2000;39:1383–9.
7. Saag KG, Koehnke R, Caldwell JR, et al. Low dose long-term corticosteroid therapy in rheumatoid arthritis: an analysis of serious adverse events. Am J Med 1994;96:115–23.
8. Angeli A, Guglielmi G, Dovio A, et al. High prevalence of asymptomatic vertebral fractures in post-menopausal women receiving chronic glucocorticoid therapy: a cross-sectional outpatient study. Bone 2006; 39:253–9.
9. Curtis JR, Westfall AO, Allison J, et al. Population-based assessment of adverse events associated with long-term glucocorticoid use. Arthritis Rheum 2006;55:420–6.
10. De Vries F, Bracke M, Leufkens HG, et al. Fracture risk with intermittent high-dose oral glucocorticoid therapy. Arthritis Rheum 2007;56: 208–14.
11. Van Staa TP, Cooper C, Leufkens HG, et al. Children and the risk of fractures caused by oral corticosteroids. J Bone Miner Res 2003;18: 913–8.
12. Canalis E, Mazziotti G, Giustina A, et al. Glucocorticoid-induced osteoporosis: pathophysiology and therapy. Osteoporos Int 2007;18: 1319–28.
13. Hansen KE, Kleker B, Safdar N, et al. A systematic review and meta-analysis of glucocorticoid-induced osteoporosis in children. Semin Arthritis Rheum 2014;44:47–54.
14. LeBlanc CM, Ma J, Taljaard M, et al. Incident vertebral fractures and risk factors in the first three years following glucocorticoid initiation among pediatric patients with rheumatic disorders. J Bone Miner Res 2015;30:1667–75.
15. Rodd C, Lang B, Ramsay T, et al. Incident vertebral fractures among children with rheumatic disorders 12 months after glucocorticoid initiation: a national observational study. Arthritis Care Res (Hoboken) 2012;64:122–31.
16. Ward LM, Ma J, Robinson ME, et al. Osteoporotic fractures and vertebral body reshaping in children with glucocorticoid-treated rheumatic disorders. J Clin Endocrinol Metab 2021;106:e5195–207.
17. Laan RF, van Riel PL, van de Putte LB, et al. Low-dose prednisone induces rapid reversible axial bone loss in patients with rheumatoid arthritis: a randomized, controlled study. Ann Intern Med 1993;119: 963–8.
18. Feldstein AC, Elmer PJ, Nichols GA, et al. Practice patterns in patients at risk for glucocorticoid-induced osteoporosis. Osteoporosis Int 2005;16:2168–74.
19. Solomon DH, Katz JN, Jacobs JP, et al. Management of glucocorticoid-induced osteoporosis in patients with rheumatoid arthritis: rates and predictors of care in an academic rheumatology practice. Arthritis Rheum 2002;46:3136–42.
20. Ettinger B. A personal perspective on fracture risk assessment tools. Menopause 2008;15:1023–6.
21. Kanis JA, Johansson H, Oden A, et al. Guidance for the adjustment of FRAX according to the dose of glucocorticoids. Osteoporosis Int 2011;22:809–16.

22. Van Staa TP, Geusens P, Pols HA, et al. A simple score for estimating the long-term risk of fracture in patients using oral glucocorticoids. QJM 2005;98:191–8.
23. American College of Rheumatology Task Force on Osteoporosis Guidelines. Recommendations for the prevention and treatment of glucocorticoid-induced osteoporosis: Arthritis Rheum 1996;39: 1791–801.
24. American College of Rheumatology Ad Hoc Committee on Glucocorticoid-Induced Osteoporosis. Recommendations for the prevention and treatment of glucocorticoid-induced osteoporosis: 2001 update. Arthritis Rheum 2001;44:1496–503.
25. Buckley L, Guyatt G, Fink HA, et al. 2017 American College of Rheumatology Guideline for the Prevention and Treatment of Glucocorticoid-Induced Osteoporosis. Arthritis Rheumatol 2017;69: 1521–37.
26. Grossman JM, Gordon R, Ranganath VK, et al. American College of Rheumatology 2010 recommendations for the prevention and treatment of glucocorticoid-induced osteoporosis. Arthritis Care Res (Hoboken) 2010;62:1515–26.
27. Andrews JC, Schunemann HJ, Oxman AD, 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.
28. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924–6.
29. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ 2010;182:E839–42.
30. Leib ES, Saag KG, Adachi JD, et al. Official Positions for FRAX(®) clinical regarding glucocorticoids: the impact of the use of glucocorticoids on the estimate by FRAX(®) of the 10 year risk of fracture from Joint Official Positions Development Conference of the International Society for Clinical Densitometry and International Osteoporosis Foundation on FRAX(®). J Clin Densitom 2011;14:212–9.
31. Saag KG, Agnusdei D, Hans D, et al. Trabecular bone score in patients with chronic glucocorticoid therapy-induced osteoporosis treated with alendronate or teriparatide. Arthritis Rheumatol 2016; 68:2122–8.
32. Huber AM, Gaboury I, Cabral DA, et al. Prevalent vertebral fractures among children initiating glucocorticoid therapy for the treatment of rheumatic disorders. Arthritis Care Res (Hoboken) 2010;62:516–26.
33. Gregson CL, Armstrong DJ, Bowden J, et al. UK clinical guideline for the prevention and treatment of osteoporosis. Arch Osteoporos 2022;17:58.
34. Silva BC, Madeira M, d'Alva CB, et al. Definition and management of very high fracture risk in women with postmenopausal osteoporosis: a position statement from the Brazilian Society of Endocrinology and Metabolism (SBEM) and the Brazilian Association of Bone Assessment and Metabolism (ABRASSO). Arch Endocrinol Metab 2022;66: 591–603.
35. Watts NB, Camacho PM, Lewiecki EM, et al. American Association of Clinical Endocrinologists/American College of Endocrinology Clinical Practice Guidelines for the Diagnosis and Treatment of Postmenopausal Osteoporosis—2020 update. Endocr Pract 2021;27:379–80.
36. Ross AC, Manson JE, Abrams SA, et al. The 2011 report on dietary reference intakes for calcium and vitamin D from the Institute of Medicine: what clinicians need to know. J Clin Endocrinol Metab 2011;96: 53–8.
37. Saag KG, Shane E, Boonen S, et al. Teriparatide or alendronate in glucocorticoid-induced osteoporosis. N Engl J Med 2007;357: 2028–39.
38. Saag KG, Zanchetta JR, Devogelaer JP, et al. Effects of teriparatide versus alendronate for treating glucocorticoid-induced osteoporosis: thirty-six-month results of a randomized, double-blind, controlled trial. Arthritis Rheum 2009;60:3346–55.
39. Allen CS, Yeung JH, Vandermeer B, et al. Bisphosphonates for steroid-induced osteoporosis [review]. Cochrane Database Syst Rev 2016;10:Cd001347.
40. Black DM, Rosen CJ. Clinical practice: postmenopausal osteoporosis. N Engl J Med 2016;374:254–62.
41. Eastell R, Rosen CJ, Black DM, et al. Pharmacological management of osteoporosis in postmenopausal women: an Endocrine Society* clinical practice guideline. J Clin Endocrinol Metab 2019;104: 1595–622.
42. Lamy O, Gonzalez-Rodriguez E, Stoll D, et al. Severe rebound-associated vertebral fractures after denosumab discontinuation: 9 clinical cases report. J Clin Endocrinol Metab 2017;102:354–8.
43. Reid DM, Devogelaer JP, Saag K, et al. Zoledronic acid and risedronate in the prevention and treatment of glucocorticoid-induced osteoporosis (HORIZON): a multicentre, double-blind, double-dummy, randomised controlled trial. Lancet 2009;373:1253–63.
44. Small RE. Uses and limitations of bone mineral density measurements in the management of osteoporosis. MedGenMed 2005;7:3.
45. Everts-Graber J, Lehmann D, Burkard JP, et al. Risk of osteonecrosis of the jaw under denosumab compared to bisphosphonates in patients with osteoporosis. J Bone Miner Res 2022;37:340–8.
46. Barrett-Connor E, Mosca L, Collins P, et al. Effects of raloxifene on cardiovascular events and breast cancer in postmenopausal women. N Engl J Med 2006;355:125–37.
47. Fixen C, Tunoa J. Romosozumab: a review of efficacy, safety, and cardiovascular risk. Current Osteoporos Rep 2021;19:15–22.
48. Lyu H, Zhao SS, Yoshida K, et al. Comparison of teriparatide and denosumab in patients switching from long-term bisphosphonate use. J Clin Endocrinol Metab 2019;104:5611–20.
49. Cosman F, Eriksen EF, Recknor C, et al. Effects of intravenous zoledronic acid plus subcutaneous teriparatide [rhPTH(1-34)] in postmenopausal osteoporosis. J Bone Miner Res 2011;26:503–11.
50. Deal C, Omizo M, Schwartz EN, et al. Combination teriparatide and raloxifene therapy for postmenopausal osteoporosis: results from a 6-month double-blind placebo-controlled trial. J Bone Miner Res 2005;20:1905–11.
51. Sun Y, Li Y, Li J, et al. Efficacy of the combination of teriparatide and denosumab in the treatment of postmenopausal osteoporosis: a meta-analysis. Front Pharmacol 2022;13:888208.
52. Brown JP, Engelke K, Keaveny TM, et al. Romosozumab improves lumbar spine bone mass and bone strength parameters relative to alendronate in postmenopausal women: results from the Active-Controlled Fracture Study in Postmenopausal Women With Osteoporosis at High Risk (ARCH) trial. J Bone Miner Res 2021;36:2139–52.
53. Geusens P, Oates M, Miyauchi A, et al. The effect of 1 year of romosozumab on the incidence of clinical vertebral fractures in postmenopausal women with osteoporosis: results from the FRAME study. JBMR Plus 2019;3:e10211.
54. Miyauchi A, Dinavahi RV, Crittenden DB, et al. Increased bone mineral density for 1 year of romosozumab, vs placebo, followed by 2 years of denosumab in the Japanese subgroup of the pivotal FRAME trial and extension. Arch Osteoporos 2019;14:59.
55. Kvist AV, Faruque J, Vallejo-Yagüe E, et al. Cardiovascular safety profile of romosozumab: a pharmacovigilance analysis of the US Food and Drug Administration Adverse Event Reporting System (FAERS). J Clin Med 2021;10:1660.
56. Adomaityte J, Farooq M, Qayyum R. Effect of raloxifene therapy on venous thromboembolism in postmenopausal women: a meta-analysis. Thromb Haemost 2008;99:338–42.
57. Van Staa TP, Leufkens HG, Abenhaim L, et al. Use of oral corticosteroids and risk of fractures. J Bone Miner Res 2000;15:993–1000.

58. Patlas N, Golomb G, Yaffe P, et al. Transplacental effects of bisphosphonates on fetal skeletal ossification and mineralization in rats. Teratology 1999;60:68–73.
59. Moe S, Drueke T, Cunningham J, et al. Definition, evaluation, and classification of renal osteodystrophy: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney Int 2006; 69:1945–53.
60. Hiramatsu R, Ubara Y, Sawa N, et al. Hypocalcemia and bone mineral changes in hemodialysis patients with low bone mass treated with denosumab: a 2-year observational study. Nephrol Dial Transplant 2021;36:1900–7.
61. Shuhart CR, Yeap SS, Anderson PA, et al. Executive summary of the 2019 ISCD Position Development Conference on monitoring treatment, DXA cross-calibration and least significant change, spinal cord injury, peri-prosthetic and orthopedic bone health, transgender medicine, and pediatrics. J Clin Densitom 2019;22:453–71.
62. Galindo-Zavala R, Bou-Torrent R, Magallares-López B, et al. Expert panel consensus recommendations for diagnosis and treatment of secondary osteoporosis in children. Pediatr Rheumatol Online J 2020;18:20.
63. Boonen S, Marin F, Obermayer-Pietsch B, et al. Effects of previous antiresorptive therapy on the bone mineral density response to two years of teriparatide treatment in postmenopausal women with osteoporosis. J Clin Endocrinol Metab 2008;93:852–60.
64. Cosman F, Keaveny TM, Kopperdahl D, et al. Hip and spine strength effects of adding versus switching to teriparatide in postmenopausal women with osteoporosis treated with prior alendronate or raloxifene. J Bone Miner Res 2013;28:1328–36.
65. Ettinger B, San Martin J, Crans G, et al. Differential effects of teriparatide on BMD after treatment with raloxifene or alendronate. J Bone Miner Res 2004;19:745–51.
66. Cosman F. Anabolic and antiresorptive therapy for osteoporosis: combination and sequential approaches. Curr Osteoporos Rep 2014;12:385–95.
67. Leder BZ, Tsai JN, Uihlein AV, et al. Denosumab and teriparatide transitions in postmenopausal osteoporosis (the DATA-Switch study): extension of a randomised controlled trial. Lancet 2015;386:1147–55.
68. Cosman F, Huang S, McDermott M, et al. Multiple vertebral fractures after denosumab discontinuation: FREEDOM and FREEDOM extension trials additional post hoc analyses. J Bone Miner Res 2022;37:2112–20.
69. Kendler D, Chines A, Clark P, et al. Bone mineral density after transitioning from denosumab to alendronate. J Clin Endocrinol Metab 2020;105:e255–64.
70. Tutaworn T, Nieves JW, Wang Z, et al. Bone loss after denosumab discontinuation is prevented by alendronate and zoledronic acid but not risedronate: a retrospective study. Osteoporos Int 2023;34:573–84.
71. Cosman F, Kendler DL, Langdahl BL, et al. Romosozumab and antiresorptive treatment: the importance of treatment sequence. Osteoporos Int 2022;33:1243–56.
72. Solling AS, Harslof T, Langdahl B. Treatment with zoledronate subsequent to denosumab in osteoporosis: a 2-year randomized study. J Bone Miner Res 2021;36:1245–54.
73. Napoli N, Langdahl BL, Ljunggren O, et al. Effects of teriparatide in patients with osteoporosis in clinical practice: 42-month results during and after discontinuation of treatment from the European Extended Forsteo(R) Observational Study (ExFOS). Calcif Tissue Int 2018;103: 359–71.
74. Johnell O, Cauley JA, Kulkarni PM, et al. Raloxifene reduces risk of vertebral fractures [corrected] in postmenopausal women regardless of prior hormone therapy. J Fam Pract 2004;53:789–96.
75. Amiche MA, Abtahi S, Driessen JH, et al. Impact of cumulative exposure to high-dose oral glucocorticoids on fracture risk in Denmark: a population-based case-control study. Arch Osteoporos 2018;13:30.
76. Lyu H, Yoshida K, Zhao SS, et al. Delayed denosumab injections and fracture risk among patients with osteoporosis: a population-based cohort study. Ann Intern Med 2020;173:516–26.
77. Cummings SR, Ferrari S, Eastell R, et al. Vertebral fractures after discontinuation of denosumab: a post hoc analysis of the randomized placebo-controlled FREEDOM trial and its extension. J Bone Miner Res 2018;33:190–8.
78. Tsourdi E, Zillikens MC, Meier C, et al. Fracture risk and management of discontinuation of denosumab therapy: a systematic review and position statement by ECTS. J Clin Endocrinol Metab. 2021;106: 264–81.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2420–2422
DOI 10.1002/acr.25195
© 2023 American College of Rheumatology.

EDITORIAL

Leveraging Quality Measurement to Achieve Best Practice Rheumatology Care: Can Pediatrics Lead Us?

Shraddha Jatwani[1] and Lisa G. Suter[2]

As adult rheumatologists, we often overlook what is happening in pediatric rheumatology, unless we are studying for our boards or brushing up for renewing our board certification. In many ways, this is a reasonable triage step because there are ever-increasing medical advances to track and less time available for peripheral topics. Furthermore, most pediatric rheumatic diseases are relatively rarer than their adult counterparts, and children are a protected research population (1), making them harder to study and develop robust evidence-based clinical practice guidelines for (2). Additionally, pediatric diseases are historically underfunded and underresearched compared with adult diseases (3). In adult rheumatology, we benefit from enhanced research resources, allowing for larger clinical trials, a more extensive research base, and increasing payer focus because of abundant costly biologic and small molecule therapeutics. And yet, despite adult rheumatology's large, high-quality (2) evidence pool, there remain large gaps between evidence-based best practice and real-world clinical care (4–7).

In this issue of *Arthritis Care & Research*, Bingham et al published the article "Pediatric Rheumatology Care and Outcomes Improvement Network's quality measure set to improve care of children with juvenile idiopathic arthritis" (8). This article reminds us how valuable it is to zoom out from our often insular view of rheumatology practice and refocus on advances we might otherwise consider tangential. The authors describe the creation and evolution of a true national learning health network (LHN), the Pediatric Rheumatology Care & Outcomes Improvement Network (PR-COIN). LHNs or learning health systems have been around for decades. Many examples of successful LHNs exist (9), ranging from coordinated care models to practice-based research. But nearly 20 years after the National Academy of Medicine's Learning Health Series (10) began, there are few national examples of LHNs. The thoughtful, patient-centered, data-driven approach to improving juvenile idiopathic arthritis (JIA) care outlined in this issue offers an important model for advancing rheumatologic care for all patients and their families.

Bingham and her co-authors summarize the ongoing success of PR-COIN, a multiyear, national effort to build the infrastructure, data repository, leadership framework and committees, and processes for a national collaborative pediatric rheumatology research, family engagement, and quality improvement (QI) system—a true functioning and evolving LHN. They use a continuously iterative approach that emphasizes patient and family input and is responsive to real-world experiences. They created a complete measure set that reflects foundational process measures to establish best practices, followed by meaningful, patient-reported outcomes measures chosen to balance both patient and provider priorities (11). As part of continuous QI efforts, they removed "topped out" measures (those with uniformly perfect or near-perfect performance across sites) and replaced them with new measures to support evolving QI targets. They included measures evaluating data quality and patient capture rates to ensure their findings were scientifically valid and representative of diverse populations. Although PR-COIN was developed as a QI LHN and is not used for accountability purposes, the designers recognized the potential for unintended consequences of measurement and included a balancing measure to monitor for adverse effects of measurement and QI efforts. Because the process and outcome measures focus on control of symptoms and disease activity, overtreatment remains a concern. Therefore, the authors included a population health measure to monitor the time interval between hospitalizations for infection, reflecting a severe negative health outcome of aggressive immunosuppressive therapy.

The authors reflect that their work was modeled on a similar successful approach to pediatric inflammatory bowel disease management (12). There are many factors that advantage pediatric subspecialities in creating LHNs. Pediatric subspecialists,

[1]Shraddha Jatwani, MD: Dignity Health Mercy Medical Group, Citrus Heights, California, and Tuoro University, Vallejo, California; [2]Lisa G. Suter, MD: Yale School of Medicine, New Haven, Connecticut, and Veterans Affairs Medical Center, West Haven, Connecticut.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25195.

Address correspondence via email to Lisa G. Suter, MD, at lisa.suter@yale.edu.

Submitted for publication June 13, 2023; accepted in revised form July 13, 2023.

including pediatric rheumatologists, tend to be clustered near academic medical centers (13,14). The authors acknowledge this, noting most PR-COIN sites are academic medical centers. This reduces, but does not eliminate, technology as a barrier to measurement and data collection. There are fewer electronic health record vendors serving academic medical centers compared to ambulatory clinical practices, and they are more likely to participate in voluntary certification programs, like the Office of the National Coordinator for Health Information Technology Health Information Technology Certification Program (15), which establishes federal standards for data interoperability. This helps streamline data capture and exchange. It also provides a unique opportunity to measure population outcomes as well as granular disease activity and patient-reported outcomes, something currently impossible on a large-scale level in adult rheumatology. The centrality of academic medical centers in PR-COIN also ensures that participation in and dissemination of research is rewarded, helping to keep these an important added focus of the LHN.

Pediatric rheumatology also benefits from the long-standing culture of a multidisciplinary approach to care that includes physicians, nurses, social workers, and other professionals such as physical and occupational therapists, as demonstrated by PR-COIN site teams. Because of the complexity of managing chronic illness in children, pediatrics also has a strong history of both patient and parental engagement and a deep commitment to advocacy. Together, these provide a strong foundation for human-centered care and large-scale change management.

The American College of Rheumatology (ACR) provides many of the same infrastructure and leadership advantages demonstrated by PR-COIN: the ACR's Rheumatology Informatics System for Effectiveness registry has over 1,000 participating rheumatology clinicians and 3.1 million patients, and the ACR has a parallel leadership core with (sub)committees attending to registries and health information technology, quality of care, quality measures, guidelines, research and publications, and ad hoc cross-committee task forces and work groups. Despite these assets, the comparatively large numbers of adult rheumatologists and patients, our dispersed nature and variable care settings make reproducing PR-COIN's model in adult rheumatology challenging. Furthermore, costly biologic and other advanced therapeutics have made adult rheumatology a focus of federal health care payment reform; a rheumatology-focused merit-based incentive payment system value pathway (MVP) (16) was one of the first such MVPs the Centers for Medicare and Medicaid Services adopted in its ongoing drive towards value-based payment. This has forced the ACR to focus on high-stakes quality measures over rapid cycle QI models like PR-COIN.

The authors acknowledge the need for continued improvement efforts and research. They specifically identify the need for increased attention to health care disparities, a planned future focus of PR-COIN. Given the success of the current work in JIA, PR-COIN and Bingham et al.'s article provides a critical roadmap for addressing health care and disparities across the age spectrum (17–19). Margaret Wheatley once said that "leadership is a series of behaviors rather than a role for heroes" (20). This statement rings true to all of us who have attempted to enact meaningful change in our clinical practices, institutions, communities, and personal lives. Bingham and colleagues provide us with a clear set of steps for creating and sustaining a national rheumatology-focused LHN. As health care continues to evolve to a more distributed care delivery model, as technology allows greater data interoperability and real-time patient-reported outcome collection, and as payers increasingly move to value-based payment, the successful LHN presented by Bingham and colleagues serves as an important example of what quality measurement implemented with insight, patient engagement, and broad collaboration can achieve.

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. Suter 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. Jatwani, Suter.
Acquisition of data. Jatwani, Suter.
Analysis and interpretation of data. Jatwani, Suter.

REFERENCES

1. US Department of Health and Human Services. HHS protection of human subjects 45CFR 46: Subpart D–Additional protections for children involved as subjects in research. 2018. URL: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/common-rule-subpart-d/index.html.
2. Martinez-Castaldi C, Silverstein M, Bauchner H. Child versus adult research: the gap in high-quality study design. Pediatrics 2008;122:52–7.
3. Bourgeois FT, Murthy S, Pinto C, et al. Pediatric versus adult drug trials for conditions with high pediatric disease burden. Pediatrics 2012;130:285–92.
4. Meisters R, Putrik P, Ramiro S, et al. EULAR/eumusc.net standards of care for rheumatoid arthritis: cross-sectional analyses of importance, level of implementation and care gaps experienced by patients and rheumatologists across 35 European countries. Ann Rheum Dis 2020;79:1423–31.
5. Perez-Ruiz F, Carmona L, Yébenes MJ, et al. An audit of the variability of diagnosis and management of gout in the rheumatology setting: the gout evaluation and management study. J Clin Rheumatol 2011;17:349–55.
6. Duarte-García A, Zamore R, Wong JB. The evidence basis for the American College of Rheumatology practice guidelines. JAMA Intern Med 2018;178:146–8.
7. Gvozdenović E, Allaart CF, van der Heijde D, et al. When rheumatologists report that they agree with a guideline, does this mean that they practise the guideline in clinical practice? Results of the International Recommendation Implementation study (IRIS). RMD Open 2016;2:e000221.

8. Bingham A, Harris JG, Qui T, et al. Pediatric Rheumatology Care and Outcomes Improvement Network's quality measure set to improve care of children with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2023;75(12):2442–52.
9. Cutler DM. Harvard Business Review. Searching for health care's entrepreneurial spirit. 2011. URL: https://hbr.org/2011/03/searching-for-health-cares-ent.
10. National Academy of Medicine. The learning health system series. URL: https://nam.edu/programs/value-science-driven-health-care/learning-health-system-series/.
11. Bermas BL, Feldman CH. Assessing systemic lupus erythematosus care: myopic evaluations and the elephant in the room. Arthritis Care Res (Hoboken) 2023;75:1194–5.
12. Crandall WV, Margolis PA, Kappelman MD, et al. Improved outcomes in a quality improvement collaborative for pediatric inflammatory bowel disease. Pediatrics 2012;129:e1030–41.
13. Mayer ML. Are we there yet? distance to care and relative supply among pediatric medical subspecialties. Pediatrics 2006;118:2313–21.
14. Mayer ML, Skinner AC. Influence of changes in supply on the distribution of pediatric subspecialty care. Arch Pediatr Adolesc Med 2009; 163:1087–91.
15. The Office of the National Coordinator for Health Information Technology (ONC). Health IT certification program. URL: https://www.healthit.gov/topic/certification-ehrs/about-onc-health-it-certification-program.
16. Quality Payment Program. Advancing rheumatology patient care (MVP ID: G0053) URL: https://qpp.cms.gov/mips/explore-mips-value-pathways/G0053.
17. Akinsete AM, Woo JM, Rubinstein TB. Disparities in pediatric rheumatic diseases [review]. Rheum Dis Clin North Am 2022:48:183–98.
18. Sandhu VK, Hojjati M, Blanco I. Healthcare disparities in rheumatology: the role of education at a global level [review]. Clin Rheumatol 2020;39:659–66.
19. Ciofoaia EI, Pillarisetty A, Constantinescu F. Health disparities in rheumatoid arthritis [review]. Ther Adv Musculoskelet Dis 2022;14: 1759720X221137127.
20. Wheatley, Margaret. Daily Good. The unplanned organization: learning from nature's emergent creativity. Excerpted from Noetic Sciences Review #37. 1996. URL: https://www.dailygood.org/story/1929/the-unplanned-organization-learning-from-nature-s-emergent-creativity-margaret-wheatley/.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2423–2427
DOI 10.1002/acr.25162
© 2023 American College of Rheumatology.

American College of Rheumatology
Empowering Rheumatology Professionals

CLINICOPATHOLOGIC CONFERENCE

Sarcoidosis and Worsening Lymphadenopathy: Sarcoidosis Flare? Think Again!

Manush Sondhi, Anusha Vuppala, and Madiha Tariq

CASE PRESENTATION

Chief symptoms

A 62-year-old African American man with a past medical history of long-standing sarcoidosis, open-angle glaucoma, and benign prostatic hyperplasia (BPH) after transurethral resection of the prostate (TURP) in 2015 presented with worsening abdominal distention and pain for the last 3 weeks prior to presentation.

History of present illness

Abdominal pain was located in the left quadrant and was constant, moderate, and dull in nature, without overt radiation exacerbated after meals. He had only a liquid diet since then. There was a history of diarrhea without blood for the last 2 days. He stated that he lost some weight over a month but could not quantify it. The patient denied any appetite change, fever with chills or rigors, congestion, sneezing, chest pain, palpitations, hematuria, hematochezia, melena, dysuria, rash, pallor, dizziness, or headaches.

The patient developed a nonresolving left thigh nodule, which he stated resembled "the size of a grapefruit" at age 37 years, which upon biopsy showed noncaseating granulomas. He was asymptomatic and managed conservatively without medications until age 57, when he presented with symptoms of cough and dyspnea with computed tomography (CT) of the chest showing pulmonary nodules. The patient was started on mycophenolate mofetil but later escalated to rituximab (received 2 cycles—within the last 3 months before admission) due to worsening symptoms. He reported improvement on mycophenolate mofetil and rituximab infusions and was never started on steroids due to a history of open-angle glaucoma, which could potentially worsen it.

About 4 years into the treatment, he presented at a different hospital with worsening symptoms of shortness of breath and fever. CT of the chest showed worsening hilar mediastinal adenopathy and ground-glass opacities bilaterally. Bronchoscopy was done with preliminary reports consistent with benign bronchial mucosa and pulmonary parenchyma, noncaseating granulomas, and negative microbiologic stains. He was treated with antibiotics, on which he reported transient improvement, after which he was discharged with 2 liters of oxygen. He was also discharged with tapering doses of prednisone for a week (30 mg for 2 days, 20 mg for the next 2 days, and 10 mg for the next 3 days) and antibiotics for possible sarcoidosis flare with pneumonia. However, the patient presented to our hospital with the above symptoms a month after discharge. All the previous pulmonary lung functions before the last hospitalization were normal, with no evidence of obstruction or restriction.

Past medical history

The patient's medical history is notable for long-standing sarcoidosis, open-angle glaucoma, and BPH (after TURP in 2015).

Social and family history

The patient is from Texas but has spent most of his time in Arkansas. He had a history of occasional alcohol use and had never consumed tobacco or any recreational drugs. The family history was noncontributory.

Review of systems

The patient noted fatigue, weight loss, shortness of breath, occasional cough, abdominal distention, pain, and diarrhea. The review of systems was otherwise negative.

Manush Sondhi, MD, Anusha Vuppala, MD, Madiha Tariq, MD: Overton Brooks VA Medical Center, Shreveport, Louisiana

Author disclosures and a graphical abstract are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25162.

Address correspondence via email to Manush Sondhi, MD, at manushsondhi@gmail.com.

Submitted for publication February 8, 2023; accepted in revised form May 18, 2023.

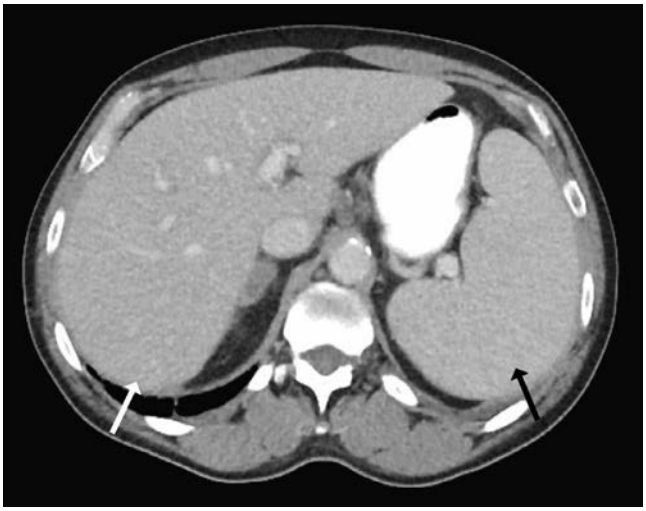


Figure 1. Computed tomography of the abdomen revealing hepatomegaly (**white arrow**) and splenomegaly (**black arrow**). The craniocaudal length of the right hepatic lobe is 21.6 cm. The spleen is enlarged, with dimensions of 12.8 × 6.1 × 6.9 cm.

Physical examination

On physical examination, the patient's blood pressure reading was 112/67 mm Hg, with a heart rate of 86 beats per minute, respiratory rate of 16 per minute, temperature of 36.5°C, and oxygen saturation of 98% on 2 liters of oxygen. He had a body mass index of 24.6 kg/m^2. The patient demonstrated intact extraocular movements and normal conjunctiva/sclera. Pupils were equal, round, and reactive to light. Lungs were clear to auscultation without wheezing, rhonchi, or rails. Cardiovascular examination showed normal findings for S1 and S2, and no murmurs, rubs, or gallops were heard. The skin was warm and dry, with no erythema or rash. Abdomen was distended and mild tenderness was present on palpating all areas (left greater than right). Normal bowel movements were heard. Radial pulses and dorsalis pedis were 2+ bilaterally. The patient had normal range of motion, and no leg edema was present. The patient was alert and oriented to person, place, and time, with no cranial nerve deficit, sensory deficit, or coordination abnormalities. The mood, thought content, and judgment of the patient was normal.

Laboratory studies

A complete blood count demonstrated a hemoglobin count of 9.3 g/dl (normal range 14–16 g/dl), white blood cell count of 1.9 thousand/m^3 (normal range 4.5 thousand–10.5 thousand/m^3), granulocyte count of 54.8% (normal range 50–80%), lymphocyte count of 20.8% (normal range 20–44%), monocyte count of 21.6% (normal range 2–15%), eosinophil count of 1.7% (normal range 0–8%), basophil count of 1.1% (normal range 0–5%), and platelet counts of 112,000/m^3 (normal range 150,000–450,000/m^3). The comprehensive metabolic panel showed sodium 137 mmoles/liter (normal range 136–145 mmoles/liter), potassium 4.2 mmoles/liter (normal range 3.5–5.1 mmoles/liter), chloride 102 mmoles/liter (normal range 98–109 mmoles/liter), calcium 9.1 mg/dl (normal range 8.9–10.3 mg/dl), albumin 2.9 g/dl (normal range 3.4–5.0 g/dl), total protein level 5.5 g/dl (normal range 6.5–8.2 g/dl), total bilirubin 2.3 mg/dl (normal range 0.2–1.2 mg/dl), aspartate aminotransferase 86 units/liter (normal range 15–40 units/liter), alanine aminotransferase 59 units/liter (normal range 14–63 units/liter), alkaline phosphatase 846 units/liter (normal range 38–126 units/liter), lipase 53 units/liter (normal range 10–140 units/liter), creatinine level 1.19 mg/dl (normal range 0.6–1.3 mg/dl), glomerular filtration rate 75 ml/minute/1.73 m2, blood urea nitrogen (BUN) 14.5 mg/dl (normal range 6–20 mg/dl), and BUN/creatinine ratio 12.2 (normal range 12–20). The results of an iron profile demonstrated iron saturation of 9% (normal range 20–40%), transferrin level 150 mg/dl (180–329 mg/dl), total iron binding capacity 210 μg/dl (normal range 261–450 μg/dl), and iron level 19 μg/dl (normal range 49–175 μg/dl). The findings of a urinalysis showed urine protein 1+, negative leukocyte esterase, negative nitrite, red blood cell count of 2/high power field (HPF) (normal range 0–2), white blood cell count 2/HPF (normal range 0–5), and pH 6 (normal range 5–8). The coagulation profile showed a prothrombin time of 13 seconds (normal range 10–13) and internalized normalized ratio of 1.

CLINICAL COURSE

The patient was admitted to the hospital. CT of the chest, abdomen and pelvis revealed a 3-cm consolidation in the upper lobe of the left lung, subcentimeter pulmonary nodularity in the

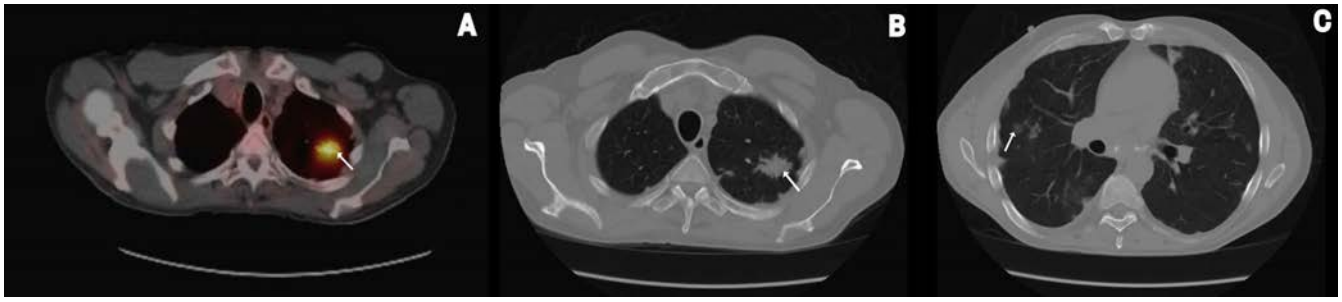


Figure 2. **A** and **B,** Interval development of a 3-cm left upper lobe consolidation in positron emission tomography scan (**arrow** in **A**) and computed tomography (CT) of the chest (**arrow** in **B**). **C,** Subcentimeter pulmonary nodularity in the right lung on CT of the chest (**arrow**).

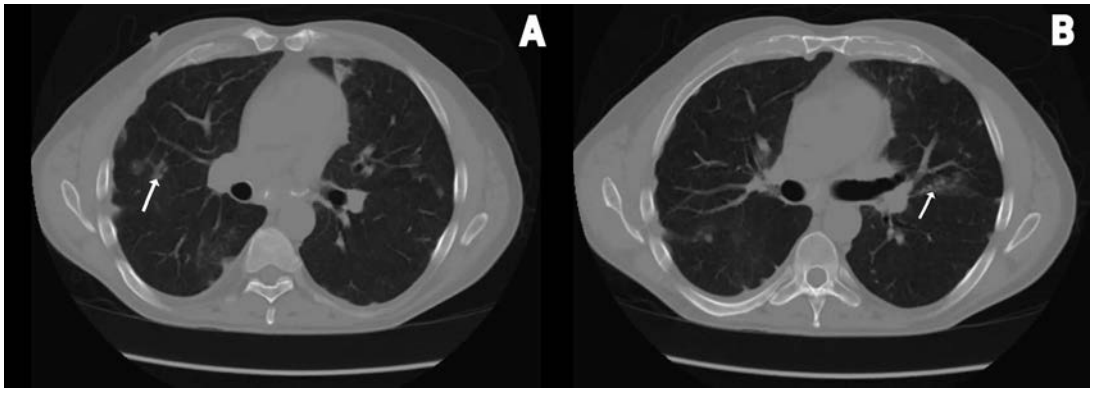


Figure 3. **A** and **B,** Bilateral patchy ground-glass opacities and scarring on computed tomography of the chest (**arrows**).

right lung, bibasilar patchy ground-glass opacities, fibrotic changes in the lung bases, hepatosplenomegaly, and bilateral adrenal lesions (See Figures 1–3). The HIV test and tuberculosis TB QuantiFeron were negative. Due to concerns of malignancy, a positron emission tomography (PET) scan was conducted that was consistent with the findings of the CT of the chest and abdomen and revealed multiple enlarged retroperitoneal and anterior mediastinal lymph nodes (see Figures 4–6). Hematology and oncology were consulted and recommended bone marrow biopsy to rule out lymphoproliferative disorders. The bone marrow biopsy showed hypercellular marrow with multiple noncaseating granulomas, and special stain grocott methenamine silver (GMS) showed fungal spores with rare budding forms consistent with *Histoplasma* species. Flow cytometry of bone marrow was negative for evidence of a clonal population of B cells or an abnormal population of T cells. Cultures from a bronchoscopy done a month ago at an outside hospital also resulted in histoplasmosis. All the immunosuppressants were stopped. Infectious disease was consulted and recommended administering amphotericin B, followed by oral itraconazole, after which the patient reported improvement.

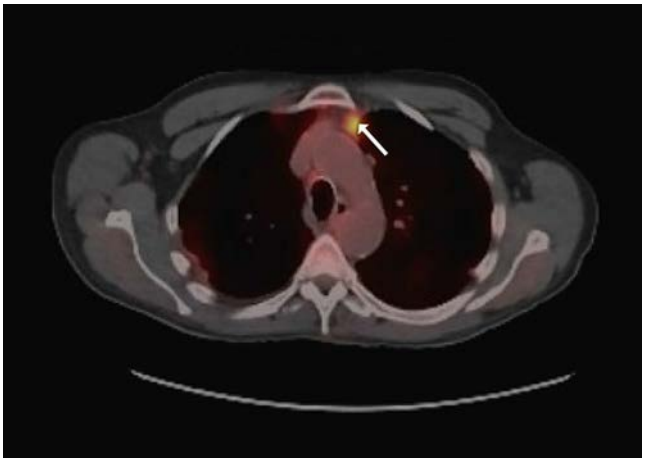


Figure 4. Anterior mediastinal fluorodeoxyglucose-avid lymph nodes, the largest measuring 1.5 cm in length on positron emission tomography (**arrow**). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25162/abstract.

Case study

A 62-year-old African American man with a past medical history of long-standing sarcoidosis presented after hospital admission with shortness of breath after a month with worsening abdominal distention and pain for the last 3 weeks. CT of the chest, abdomen, and pelvis revealed hepatosplenomegaly, bilateral adrenal lesions, and fibrotic changes in the lung bases. Due to CT findings and pancytopenia, a bone marrow biopsy was performed to rule out lymphoproliferative disorders, which showed hypercellular marrow with multiple noncaseating granulomas, and special stain GMS showed fungal spores with rare budding forms consistent with *Histoplasma* species. Amphotericin B was started, followed by oral itraconazole.

DIFFERENTIAL DIAGNOSES

Differential diagnoses for this patient include sarcoidosis flare-up, hematologic malignancy, disseminated histoplasmosis, TB, HIV, primary lung cancer, and metastasis.

DISCUSSION

Sarcoidosis is an inflammatory disease characterized by the formation of noncaseating granulomas in multiple organs and tissues. Patients are predisposed to infections due to immunosuppression. Histoplasmosis is the most prevalent endemic mycosis in the US and is mainly found in Ohio and Mississippi river valleys as well as in southeastern states (1). Sarcoidosis and histoplasmosis can present very similarly, involving uveitis, arthralgias, skin ulcers, erythema nodosum, splenomegaly, and hepatitis, with similar radiologic findings of hilar lymphadenopathy and reticulonodular opacities (2). The most common findings in disseminated histoplasmosis include

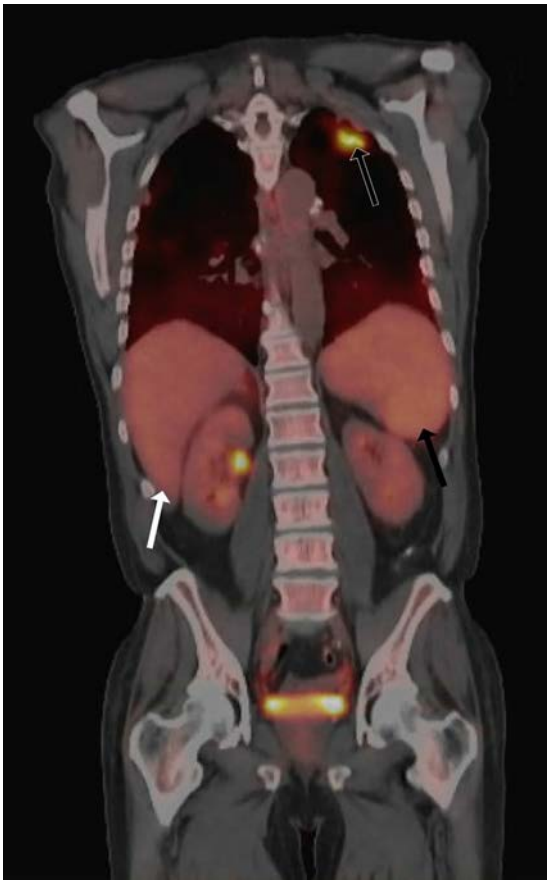


Figure 5. Coronal view of a positron emission tomography scan revealing hepatomegaly (**white arrow**), splenomegaly (**black arrow**), and consolidation in the upper lobe of the left lung (**black arrow with white outline**). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25162/abstract.

splenomegaly (72.0%), hepatomegaly (68.1%), and lymphadenopathy (41.2%) (3). Involvement in sarcoidosis constitutes mediastinal lymphadenopathy (90%), peripheral lymphadenopathy (40%), cutaneous (25%), ocular (25%), hepatomegaly (20%), splenomegaly (6%), and musculoskeletal, including polyarthralgia (10%) (4).

Histoplasmosis can present clinically in varied forms, ranging from asymptomatic infection, acute or chronic pulmonary infection, mediastinal fibrosis or granulomas, or as chronic disseminated histoplasmosis. Patients with disseminated histoplasmosis have high levels of antigenuria and antigenemia. It provides rapid and accurate diagnosis with a sensitivity and specificity of >90% (5). The standard methodology for serology is immunodiffusion, which is more specific (100%) than sensitive (70%) (5). Histopathology is frequently used to diagnose histoplasmosis but lacks sensitivity (40%) and facilitates diagnosis primarily in patients with acute, chronic disseminated infection or severe pulmonary infection (5). Definitive diagnosis is still based on isolating and identifying histoplasma, and the culture has a sensitivity of almost 85% in disseminated histoplasmosis (5). The treatment includes amphotericin B, and once there is improvement in the symptoms, it can be transitioned to itraconazole after 1 or 2 weeks, which is continued for at least 1 year to reduce the risk of relapses. Once started on itraconazole, histoplasma antigens are checked in urine and serum for the first few months and then at 3-month intervals until treatment is finished (6).

Immunosuppressed patients, e.g., transplant recipients, patients with AIDS, and those receiving glucocorticoids, cannot develop a proper cell-mediated immunity against the pathogen, which leads to the dissemination of the fungi (7). Patients with active sarcoidosis already have suppressed cell-mediated immunity; doses of steroids and immunosuppressive medications further exacerbate this immunosuppression, increasing the risk for atypical/opportunistic infections. Our patient presented with abdominal pain and distention and was found to have hepatosplenomegaly and lymphadenopathy, which are common symptoms of both sarcoidosis and disseminated histoplasmosis. This underscores the importance of conducting a thorough

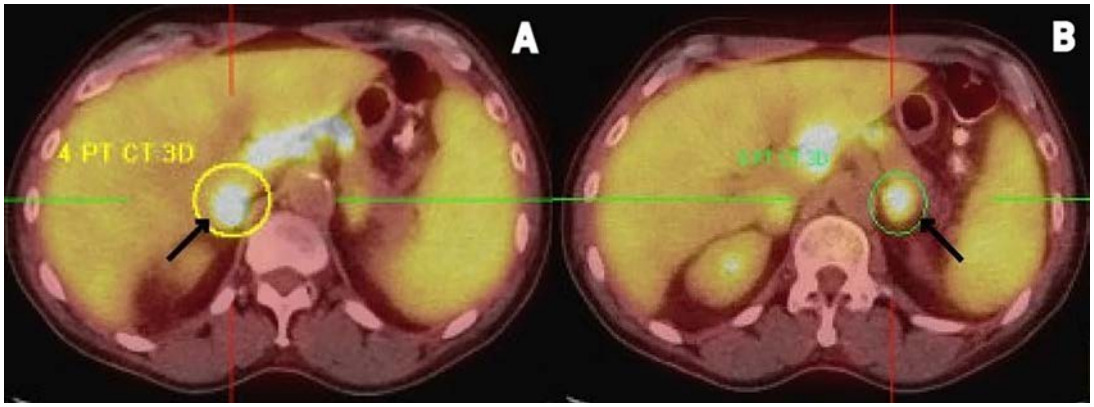


Figure 6. **A** and **B,** Intensely fluorodeoxyglucose-avid right adrenal mass measuring 3.5 cm in length (**arrow** in **A**) and left adrenal mass measuring 2.4 cm in length (**arrow** in **B**), along with retroperitoneal nodes. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25162/abstract.

evaluation, including biopsy, culture collection, serology, and radiography, to rule out any infectious causes before diagnosing a sarcoidosis flare. Improper treatment can lead to significant morbidity and even mortality.

In summary, there is considerable overlap in the clinical, pathologic, and radiologic presentations of histoplasmosis and sarcoidosis. Therefore, physicians must remain vigilant and consider the possibility of opportunistic infections when patients with sarcoidosis experience worsening symptoms. Prompt and accurate diagnosis can make a significant difference in patient outcomes.

FINAL DIAGNOSIS

Disseminated histoplasmosis.

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. Sondhi 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. Sondhi, Vuppala, Tariq.

Acquisition of data. Sondhi, Tariq.

Analysis and interpretation of data. Sondhi, Tariq.

REFERENCES

1. Akram SM, Koirala J. Histoplasmosis. In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK448185/.
2. Bansal A, Drewek R: Sarcoidosis and histoplasmosis: is one a consequence of the other? A case report and review of the literature. Case Rep Rheumatol 2015;2015:108459.
3. Xiong XF, Fan LL, Kang M, et al. Disseminated histoplasmosis: a rare clinical phenotype with difficult diagnosis. Respirol Case Rep 2017;5:e00220.
4. King TE Jr, Flaherty KR, Dieffenbach P. UptoDate: extrapulmonary manifestations of sarcoidosis. URL: https://www.uptodate.com/contents/extrapulmonary-manifestations-of-sarcoidosis.
5. Guimarães AJ, Nosanchuk JD, Zancopé-Oliveira RM. Diagnosis of histoplasmosis. Braz J Microbiol 2006;37:1–13.
6. Hage CA, Kirsch EJ, Stump TE, et al. Histoplasma antigen clearance during treatment of histoplasmosis in patients with AIDS determined by a quantitative antigen enzyme immunoassay. Clin Vaccine Immunol 2011;18:661–6.
7. Kauffman CA. Histoplasmosis: a clinical and laboratory update [review]. Clin Microbiol Rev 2007;20:115–32.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2428–2434
DOI 10.1002/acr.25165
© 2023 American College of Rheumatology.

BRIEF REPORT

Incorporating Telemedicine in Rheumatology Fellowship Training Programs: Needs Assessment, Curricular Intervention, and Evaluation

Marcy B. Bolster,[1] Jason Kolfenbach,[2] Alexandra Poeschla,[1] Lisa Criscione-Schreiber,[3] Faye Hant,[4] Rumey Ishizawar,[5] Beth Jonas,[5] David Leverenz,[3] Kenneth S. O'Rourke,[6] Rachel M. Wolfe,[7] and Lisa Zickuhr[8]

Objective. To increase the confidence of rheumatology fellows in training (FITs) in delivering virtual care (VC) and prepare them for independent practice, we developed educational materials addressing gaps in their skills.

Methods. We identified gaps in telemedicine skills based on FIT performance in a virtual rheumatology objective structured clinical examination (vROSCE) station on VC delivery using video teleconference technology and survey (survey 1) responses. We created educational materials including videos of "mediocre" and "excellent" VC examples, discussion/reflection questions, and a document summarizing key practices. We measured change in the confidence levels of FITs for delivering VC with a post-intervention survey (survey 2).

Results. Thirty-seven FITs (19 first-year, 18 second- plus third-year fellows) from 7 rheumatology fellowship training programs participated in a vROSCE and demonstrated gaps in skills mapping to several Rheumatology Telehealth Competency domains. Confidence levels of FITs improved significantly from survey 1 to survey 2 for 22 of 34 (65%) questions. All participating FITs found the educational materials helpful for learning and reflecting on their own VC practice; 18 FITs (64%) qualified usefulness as "moderately" or "a lot." Through surveying, 17 FITs (61%) reported implementing skills from the instructional videos into VC visits.

Conclusion. Continually assessing our learners' needs and creating educational materials addressing gaps in training are requisite. Using a vROSCE station, needs assessments, and targeted learning with videos and discussion-guidance materials enhanced the confidence level of FITs in VC delivery. It is imperative to incorporate VC delivery into fellowship training program curricula to ensure breadth in skills, attitudes, and knowledge of new entrants into the rheumatology workforce.

INTRODUCTION

Telemedicine, the delivery of synchronous and asynchronous patient care using technology platforms, was gaining momentum when the onset of the COVID-19 pandemic occurred in February 2020. The pandemic subsequently prompted widespread clinical and educational integration of virtual care (VC), a synchronous form of health care delivery through the use of audio-only or video teleconference technology (1–5). A unique skillset is necessary for patients and clinicians to effectively engage in VC, including successfully interacting with technology, addressing barriers to equitable access to VC, performing a virtual physical examination, and developing "webside" manner (6). Notably, 2 surveys of rheumatology fellows-in-training (FITs) in 2020 highlighted the interruption of clinical training alongside inadequacies of teaching and supervision during VC encounters (7,8). FITs perceived a reduced quality of clinical teaching during VC encounters, with 70% of FITs reporting a negative or slightly negative impact of VC on teaching quality (7), highlighting the need for improved curricular content and clinical training.

Supported by the Rheumatology Research Foundation (Clinician Scholar Educator Award to Drs. Bolster, Kolfenbach, and Zickuhr).

[1]Marcy B. Bolster, MD, Alexandra Poeschla: Massachusetts General Hospital, Boston; [2]Jason Kolfenbach, MD: University of Colorado, Aurora; [3]Lisa Criscione-Schreiber, MD, MEd, David Leverenz, MD: Duke University, Durham, North Carolina; [4]Faye Hant, MD: Medical University of South Carolina, Charleston; [5]Rumey Ishizawar, MD, Beth Jonas, MD: University of North Carolina at Chapel Hill; [6]Kenneth S. O'Rourke, MD: Maine Medical Center, Portland; [7]Rachel M. Wolfe, MD: Wake Forest School of Medicine, Winston-Salem, North Carolina; [8]Lisa Zickuhr, MD, MHPE: Washington University, St. Louis, Missouri.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25165.

Address correspondence via email to Marcy B. Bolster, MD, at mbolster@mgh.harvard.edu.

Submitted for publication December 23, 2022; accepted in revised form May 23, 2023.

SIGNIFICANCE & INNOVATIONS

- The targeted design of educational materials to address gaps in virtual care (VC) skills enhances the confidence levels of rheumatology fellows in training (FITs) in VC delivery.
- Content conveyed through instructional videos and accompanying educational materials demonstrating nuanced differences between "mediocre" and "excellent" VC practices is highly transferrable to the VC clinical practice of rheumatology FITs.
- Resource sharing within a broad rheumatology education community will enrich the opportunities for learning as well as strengthen the preparedness of our FITs, the new entrants into our workforce.

The rheumatology specialty has expanded its educational footprint in VC to address the gaps identified in the 2020 FIT surveys (7,8). The American College of Rheumatology (ACR) Core Curriculum now includes telehealth topics. Additionally, based on the Association of American Medical College's Telehealth Competencies (9), the ACR Committee on Training and Workforce developed the Telehealth Competencies outlining the core domains and skills necessary for VC in rheumatology (10). Despite these initiatives, gaps in fellowship training persist and must be addressed to keep pace with the continued implementation and advancement of VC.

In this study, we present an updated needs assessment for VC training, nearly 2 years after the 2020 FIT surveys (7,8), utilizing input from trainees and faculty from geographically diverse programs, as well as direct observation of FIT clinical performance during a simulated VC encounter. Second, guided by the needs assessment, we implemented an educational initiative and assessed its impact on FIT confidence in providing VC.

MATERIALS AND METHODS

Participants. FITs from 2 collaborative groups, representing 7 rheumatology fellowship training programs, participated in this study. The first group comprised a well-established collaborative between 5 fellowship training programs, the Carolinas Fellows Collaborative-Massachusetts General Hospital (CFC-MGH) group: Duke University; University of North Carolina; Wake Forest University; Medical University of South Carolina; and Massachusetts General Hospital. The second group comprised Washington University in St. Louis (WUSL) and the University of Colorado (CU) rheumatology fellowship training programs; these programs were added based on ongoing collaboration between study authors (MBB, JK, and LZ) with a specific focus on fellowship training program curriculum design in telemedicine.

Assessment of baseline VC encounter skills. Both collaboration groups sponsored a virtual rheumatology objective structured clinical examination (vROSCE) (February or March 2022) that included one station simulating a video teleconference patient encounter in the rheumatology outpatient setting. This station was created (by MBB, reviewed, edited, and approved by all coauthors) to evaluate FITs providing VC, and assessment was based on the Accreditation Council of Graduate Medical Education (ACGME) Core Competencies (11) and ACR Telehealth Competencies (10).

The VC station (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25165) was 10 minutes in duration followed by 2 minutes for feedback to the trainee. The FIT received instructions prior to starting the vROSCE station. The simulated video teleconference encounter featured a patient actor with stable lupus who was planning a pregnancy. There were 4 patient actors, as each of the vROSCEs (CFC-MGH and WUSL-CU) consisted of 2 simultaneous circuits to accommodate all FITs. Each patient actor was provided with written instructions for the station and received pre-ROSCE training for their roles by the lead faculty preceptor. The faculty preceptor and patient actor used standardized checklists to assess fellows' performance of VC delivery and medical knowledge of reproductive health considerations for patients with lupus. Faculty vROSCE checklist items specific to VC mapped to the Rheumatology Telehealth Competencies domains (10) of patient safety and appropriate use, communication skills, data acquisition and assessment, and systems-based practice, while patient vROSCE checklist items mapped to the 2 domains of patient safety and appropriate use and communications skills.

Needs assessment survey. In the absence of a previously validated survey that fit our study aims, we created a 34-question survey (see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25165) assessing FIT confidence levels regarding VC skill development, mapping survey 1 questions to domains within the Rheumatology Telehealth Competencies (10).

Links and QR codes to survey 1 were distributed to FITs via email after participation in the vROSCE. Survey participation was voluntary and anonymous, and FITs provided consent for participation before submitting responses in RedCap. Survey development occurred at MGH, and the MGH Institutional Review Board approved its distribution.

Educational materials. The results of survey 1 and FIT performance on the vROSCE station were used to develop VC educational materials, addressing skills identified as areas of low confidence by FITs and low performance by vROSCE preceptors. Educational materials (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25165) included 2 videos of "mediocre" and "excellent" staged outpatient visits to guide FITs in

distinguishing nuances in VC delivery skillsets. The videos depicted a virtual follow-up encounter with a patient actor (AP) portraying a person recently diagnosed with rheumatoid arthritis and a clinician (MBB). FITs also received a table outlining the differentiating characteristics between the "mediocre" and "excellent" videos, including features related to communication skills (the clinician's camera positioning, use of pointed versus open-ended questioning for obtaining the history, "webside manner"), data collection and assessment (virtual physical examination), patient safety and appropriate use (characteristics aligning with suitability for virtual versus in-person care) and systems-based practice (post-visit implementation of the patient care plan). Last, FITs received self-reflection questions prompting self-analysis of their abilities to deliver VC and areas for potential development. Self-reflection data were not collected. The educational materials were distributed to the FITs (March through June 2022) after their participation in the vROSCE station. FITs voluntarily interacted with the educational materials, either individually or as a small-group learning activity within their training programs. We did not collect validity evidence in the use and distribution of the educational materials because the materials were designed with the intent of providing flexibility for implementation at various programs; each training program was allowed to implement the educational materials as best fit their setting and needs.

Evaluation of educational materials and skills progression survey. The second survey (see Supplementary Appendix C, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25165) was distributed to FITs who completed survey 1, and it reassessed FIT confidence in performing skills requisite to VC, comprising the initial 34 questions (survey 1) with 8 additional questions about learning and change in practice related to the educational materials. Survey 2 was distributed to FITs via email link and QR code for completion in RedCap, 4 months after completion of survey 1.

Statistical analysis. Statistical analyses were performed using SPSS. Data are presented as the entire group of FITs as well as a subgroup analysis of first-year FITs and second-plus-year (second- and third-year) FITs. Chi-square analysis and Mann-Whitney U tests determined demographic differences between first- and second-plus-year FITs.

We calculated the total scores awarded by faculty and patient actors on vROSCE checklists and the total confidence scores submitted by FITs on survey 1 in each Rheumatology Telehealth Competency domain. Aggregate scores for each domain were converted to total percentages and were compared across FITs, faculty, and patient-generated data (limited to 2 domains scored by the patient actors) using chi-square analysis, thereby triangulating data sources to identify areas in which FITs could benefit from additional training. Wilcoxon's signed rank test was used to compare pre- and post-survey responses for the entire group, for first-year FITs, and for second-plus-year FITs.

RESULTS

Participants. Thirty-seven FITs (19 first-year fellows and 18 second-plus-year fellows) from 7 rheumatology fellowship training programs participated in a vROSCE and were then invited to complete survey 1. Thirty (30 of 37, 81%) FITs completed survey 1 (Table 1). Data from vROSCE for 35 FITs were available; technology failures prevented 2 FITs from completing the vROSCE station. Most fellows surveyed had participated in VC for ≥18 months, although a significant proportion of first-year FITs had <12 months of experience compared to the second-plus-year group ($P < 0.001$).

Needs assessment: survey 1 and vROSCE. Faculty and patient actor assessment of FIT vROSCE performance revealed insufficiencies triangulating with fellows' self-reported confidence in VC skills (Table 2). Faculty identified a gap in FIT skills within the Rheumatology Telehealth Competency domain of patient safety and appropriate-use at a level commensurate with fellow self-reported confidence, while patients scored FITs skills in patient safety and appropriate-use more favorably ($\chi^2 = 6.423$, $P = 0.04$). FITs and faculty similarly identified opportunities for improvement in systems-based practice ($\chi^2 = 3.227$, $P = 0.072$), while FITs self-reported gaps in data acquisition (e.g., gathering patient-generated data, conducting the virtual physical examination) and assessment at a significantly greater degree than faculty observed ($\chi^2 = 6.737$, $P = 0.009$). FITs self-reported significantly lower VC encounter communications skills than were observed by faculty and patients during the vROSCE ($\chi^2 = 9.233$, $P = 0.01$).

Survey 1 contained questions mapping to content beyond the scope of the vROSCE, further revealing that at least one-fourth of participating FITs lacked confidence in developing rapport during audio-only encounters (45%), in engaging patients' care partners during audio-only (90%) and video teleconference

Table 1. Demographic characteristics of fellows in training participants*

	Survey 1 (n = 30)	Survey 2 (n = 28)
Male sex	16 (53.5)	14 (50.0)
Year in fellowship		
First	13 (43.3)	12 (42.9)
Second-plus†	17 (56.6)	16 (57.2)
Participated in ≥18 months of virtual care		
First year, no./total no. (%)	7/13 (54)	–
Second-plus-year, no./total no. (%)	16/17 (94)	–

* Values are the number (%) unless indicated otherwise.
† Second-plus-year: second- and third-year fellows were combined due to small numbers of third-year fellows.

Table 2. Comparison of mean scores from Survey1 (fellows in training [FITs] self-assessment of confidence levels) and virtual rheumatology observed structured clinical examination (vROSCE) checklist items (faculty and patient actor assessment of FITs) across rheumatology telehealth competency domains*

Rheumatology TH competency domains	FIT self-assessment	Faculty assessment of FIT	Patient actor assessment of FIT	χ^2 for TH domain	*P*
Patient safety and appropriate use†	68.7 (33.3–100)	57.4 (25–100)	86.9 (20–100)	6.42	0.04
Communication skills‡	62.9 (50–100)	96.5 (86.5–100)	98.2 (90–100)	9.23	0.01
Data collection and assessment§	59.8 (33.3–100)	92.0 (66.6–100)	–	6.74	0.009
Systems-based practice¶	63.6 (33.3–100)	85.8 (62.5–100)	–	3.23	0.072

* Values are the mean score % (% range) unless indicated otherwise. TH = telehealth.
† Included items related to teaching patients about the use and application of the video teleconference platform and determining if a patient is appropriate for a virtual care visit based on patient- and disease-specific parameters.
‡ Included items related to establishing rapport during a video teleconference visit, speaking respectfully to patients, and informing patients about post-visit care.
§ Included items related to gathering a history, examining the patient, and collecting other clinical data during a video teleconference visit; patient actors did not assess this domain during the vROSCE.
¶ Systems-based practice included items related to establishing workflow expectations with the patient and acknowledging the limitations of virtual care; patient actors did not assess this domain during the vROSCE.

visits (45%), in obtaining comprehensive histories during audio-only encounters (25%), in incorporating remote patient data (35%), and in knowing the legal limitations of VC (60%).

FIT evaluation of educational materials and self-assessment of skill progression (survey 2). Survey 2 was completed by 28 of 30 FITs who completed survey 1 (93%), and 27 (94%) indicated educational materials utilization. FITs engaged with the educational materials independently (54%) and in groups (46%). All participating FITs deemed the educational materials useful for considering their own VC practice, and 64% rated helpfulness as "moderately" or "a lot." Seventeen FITs (61%) reported implementing skills from the educational materials into their own VC visits.

A significant improvement in confidence levels was demonstrated from survey 1 to survey 2 for 22 of 34 (65%) questions (Table 3). Confidence in VC skills with most significant improvement included recognizing and mitigating one's own unconscious biases, developing rapport in a video teleconference visit, engaging the patient's care partner in an audio-only visit, informing the patient about pre-visit logistics, performing virtual physical examinations, and knowledge of legal limitations. Areas in which confidence levels were high in both surveys, thus not changing significantly among the entire group of FITs, included the ability of FITs to use and instruct a patient in VC technology, engaging the patient's care partner in a video teleconference visit, performing a comprehensive history, and instructing the patient in post-visit care coordination.

In a subgroup analysis, first-year FITs demonstrated significant improvement in level of confidence in only 4 survey items: developing rapport within a video teleconference visit; conducting a virtual musculoskeletal examination; being able to use the virtual platform for a video teleconference visit; and acknowledging legal limitations.

Second-plus-year FITs demonstrated significant improvement in level of confidence in 16 items including: determining appropriateness for a virtual care visit based on patient-and disease-specific parameters; recognizing and mitigating unconscious biases; establishing rapport in an audio-only and video teleconference visit; engaging the care partner, pre- and post-visit care and logistics; legal aspects of virtual care; shared decision-making; and use of remote patient data.

Additionally, when asked on survey 2 to compare confidence levels in VC abilities as compared to these abilities during in-person visits, the second-plus-year FITs reported increased confidence in performing valuable virtual musculoskeletal, skin, and neurologic physical examinations, participating in shared decision-making, and pre- and post-visit coordination of care within the VC setting in the time period between survey 1 and survey 2.

DISCUSSION

VC is now well integrated into the care of patients with rheumatic diseases, offering options for care delivery as well as potentially improving access to care in underserved areas. In recognizing the ongoing impact of the rheumatology workforce shortage (12), in showing awareness of a maldistribution of rheumatologists across the US that provides large regions with limited access to rheumatology clinicians (13), and in acknowledging mobility limitations of many of our patients, virtual care is an asset to enhancing access to rheumatology care. As such, it is imperative that fellowship training programs incorporate VC delivery into curricula (14) to ensure breadth in the skills, attitudes, and knowledge of new entrants into the rheumatology workforce. VC curricula must integrate tenets of the ACGME Core Competencies (11) and Rheumatology Telehealth Competencies (10), while remaining flexible to address evolving areas of need.

Two years into the pandemic, this study's data highlight ongoing gaps in training. Surveys of FITs early in the pandemic

Table 3. Change in levels of confidence among the entire group of fellows in training (FITs)*

Question	Survey1	Survey2†	*P*
I can teach my patient about the use and application of the audio-only telemedicine platform	2.17 ± 0.592	2.43 ± 0.634	0.097
I can teach my patient about the use and application of the video telemedicine platform	2.13 ± 0.629	2.36 ± 0.621	0.058
I can determine if my patient is appropriate for a virtual care visit based on patient-specific parameters	2.07 ± 0.521	2.39 ± 0.567	0.007
I can determine if my patient is appropriate for a virtual care visit based on disease-specific parameters	2.03 ± 0.615	2.43 ± 0.504	0.002
I am able to recognize my own unconscious biases and mitigate the impact on patient care	1.60 ± 0.563	2.25 ± 0.585	<0.001
I can address my patient's ability to access the telemedicine platform	1.77 ± 0.568	2.11 ± 0.567	0.007
I can select a patient for a telehealth visit such that it will enhance the care I am providing	2.03 ± 0.490	2.25 ± 0.518	0.058
I am able to develop rapport with my patient in an audio- only virtual care visit	1.67 ± 0.547	2.00 ± 0.544	0.013
I am able to develop rapport with my patient in a video virtual care visit	1.30 ± 0.535	1.86 ± 0.591	0.001
I am able to engage my patient's care partner during an audio-only virtual care visit	1.17 ± 0.461	1.68 ± 0.670	0.003
I am able to engage my patient's care partner during a video virtual care visit	1.60 ± 0.563	1.86 ± 0.591	0.071
I am able to inform my patient about pre-visit preparation for the telehealth visit	1.57 ± 0.626	2.04 ± 0.508	0.003
I am able to inform my patient about post-visit care (includes labs, x-rays, follow-up scheduling)	2.13 ± 0.507	2.25 ± 0.585	0.317
Compared to in-person visits, I am able to inform my patient about post-visit care (includes labs, x-rays, follow- up scheduling)	1.90 ± 0.481	2.11 ± 0.416	0.034
I can participate in shared decision making with my patient via telehealth visit	2.0 ± 0.371	2.25 ± 0.441	0.020
Compared to in-person visits, I can participate in shared decision making with my patient via telehealth visit	1.87 ± 0.434	2.04 ± 0.331	0.025
Compared to in-person visits, I can obtain a comprehensive history via a telehealth encounter	1.87 ± 0.507	2.04 ± 0.429	0.046
I can obtain a comprehensive history via an audio-only telehealth encounter	2.0 ± 0.587	2.21 ± 0.499	0.132
I can obtain a comprehensive history via a video telehealth encounter	2.13 ± 0.507	2.29 ± 0.460	0.317
I can conduct a valuable MSK physical examination via telehealth	1.37 ± 0.615	1.93 ± 0.604	<0.001
I can conduct a valuable skin physical examination via telehealth	1.40 ± 0.563	1.89 ± 0.629	0.003
I can conduct a valuable neuro physical examination via telehealth	1.20 ± 0.484	1.79 ± 0.686	0.001
Compared to in-person visits, I can conduct a valuable MSK physical examination via telehealth	1.10 ± 0.403	1.50 ± 0.694	0.005
Compared to in-person visits, I can conduct a valuable skin physical examination via telehealth	1.17 ± 0.431	1.46 ± 0.693	0.021
Compared to in-person visits, I can conduct a valuable neuro physical examination via telehealth	1.07 ± 0.365	1.46 ± 0.693	0.005
I can incorporate remote patient data collection into my telehealth visit (such as a pre-visit Rapid 3)	1.73 ± 0.583	2.07 ± 0.604	0.033
Compared to in-person visits, I can incorporate remote patient data collection into my telehealth visit (such as a pre-visit Rapid 3)	1.67 ± 0.547	2.00 ± 0.609	0.013
I can use the virtual platform to conduct an audio-only telehealth visit	2.17 ± 0.531	2.29 ± 0.600	0.564
I can use the virtual platform to conduct a video telehealth visit	2.10 ± 0.607	2.36 ± 0.559	0.109
I know the legal limitations for conducting a telehealth visit	1.43 ± 0.626	2.07 ± 0.716	0.001
I am able to provide HIPAA-compliant care via a telehealth visit	1.87 ± 0.629	2.32 ± 0.548	0.005
Compared to in-person visits, I am able to provide HIPAA- compliant care via a telehealth visit	1.80 ± 0.551	2.00 ± 0.609	0.132
I can establish workflow expectations with my patient in a virtual care visit	1.90 ± 0.607	2.11 ± 0.567	0.096
Compared to in-person visits, I can establish workflow expectations with my patient in a virtual care visit	1.77 ± 0.626	1.96 ± 0.508	0.096

* Values are the mean ± SD unless indicated otherwise. Survey1 and survey2 asked FITs to indicate their level of confidence as 1) "Lack confidence," 2) "Feel confident," or 3) "Confident in role modeling and teaching this to others." When comparing level of confidence for a virtual care encounter to an in-person encounter, FITs indicated their level of confidence as 1) "Have less confidence," 2) "Feel equally confident," or 3) "Feel more confident." HIPPA = Health Insurance Portability and Accountability Act; MSK = musculoskeletal.
† Survey2 was conducted 4 months after survey1.

identified deficiencies in supervision, inadequate teaching of the virtual physical examination, concerns regarding accurate clinical assessments in the VC environment, and enhancing the VC visit experience for patients and clinicians as areas of concern (7,8). While some aspects of the VC visit, faculty supervision, and teaching have likely improved since the time of the initial FIT

surveys (7,8), our study demonstrates areas warranting continued enrichment. Current lack of standardization and definition of best practices represent challenges for faculty development in the area of teaching in the VC setting. The virtual physical examination (part of the Rheumatology Telehealth Competencies domain of Data Acquisition and Assessment) is highlighted as a gap in training in our study. We identified low levels of fellows' confidence in additional areas such as optimizing the VC experience for patients, developing rapport, engaging the care partner, and outlining pre-visit preparation for VC encounters for the patient. These features are integral for the development of "webside manner," an area uniquely deserving of attention in the training environment. FITs expressed low confidence levels in areas not identified as gaps by faculty or patients, and it is important to address these and include this content in curricular initiatives to build trainees' confidence in VC practice.

Based on survey data and evaluations from a vROSCE station, we created educational materials in the form of instructional videos, comparison tables outlining the distinguishing characteristics between the "mediocre" and "excellent" videos, and discussion questions that were well received. The vROSCE station itself can be used as a formative assessment tool to provide direct observation and immediate feedback to FITs on skills in the VC setting. This study describes the piloted implementation of these curricular elements in 7 fellowship training programs, thereby enhancing the generalizability for successful implementation within the wider rheumatology education community.

Survey 2 was administered 3–4 months after survey 1, as it was expected that FITs would gain confidence in many aspects of rheumatology ambulatory care by virtue of accomplishing another 25–33% of a year of training. Additionally, targeted educational materials addressing areas of low confidence levels and vROSCE station-identified gaps in VC skills by FITs were provided. Interestingly, the second-year FITs noted a significant improvement in level of confidence on more survey questions (16 questions) compared to the first-year FITs (4 questions), and this is likely multifactorial. Most second-plus-year FITs had had >18 months of VC experience, and additionally, toward the end of the second year of training, FITs have greater skills in general rheumatology patient care and thus may be more apt to focus on distinctive areas of professional growth, such as VC delivery. First-year FITs may still be focused on areas of more foundational learning, thus the expansion of VC skills may come later, such as during years 2 and 3 of training. The educational videos similarly may have provided more nuanced educational value for second-plus-year compared to first-year FITs.

Limitations of this study include its small size and anonymity, permitting assessment of data in aggregate, and that it precluded assessment of an individual's improvement. VC skills were assessed with direct observation (a strength); however, the vROSCE station had a 10-minute-imposed time limit, which may contribute to erroneously identified skillset gaps based on time limitations. While reassessment of the confidence levels of FITs was collected (survey 2), a follow-up vROSCE VC station was not conducted to reassess the skills acquisition of FITs over time. The vROSCE station was limited to a video teleconference assessment activity; audio-only VC visits, although assessed in survey 2 for FIT confidence levels, were not part of the formative assessment of the vROSCE station. Not all trainees found high value in the VC educational materials, and solicitation of additional feedback from this group would be useful to aid in the future development of materials with both appeal and effectiveness for our trainees. Additionally, improvements in FIT confidence between survey 1 and survey 2 may not have been attributable solely to the educational materials, but rather have stemmed from instruction in their individual training programs as well as the growth that occurs with greater time in training.

The strengths of this study include a targeted approach to addressing gaps in FIT education as identified by a combination of a needs assessment survey, FIT performance in a simulated VC setting, and published areas of identified concern (7,8). Further, this study broadly represents differing program sizes and geographic locations. There was a high response rate on both surveys, and the surveys identified many areas of significant improvement with the implemented curricular additions. Additionally, the curricular tools of the vROSCE station and the educational videos do not require in-person participation; these could feasibly be shared with any rheumatology fellowship training program.

Published surveys of FITs (7,8), integrated with foundational tools such as the ACGME Core Competencies (11) and Rheumatology Telehealth Competencies (10), contribute to directing curricular design in a time of an evolving landscape of VC, fulfilling the educational best practice of continually reassessing learners' needs and addressing gaps in training. Our educational materials can be adapted to other training programs, including pediatric rheumatology and other specialties, where different skillsets can be addressed. Resource sharing within a broad rheumatology education community will enrich the opportunities for learning, as well as strengthen the preparedness of our FITs, the new entrants into our workforce.

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. Bolster 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. Bolster, Kolfenbach, Poeschla, Criscione-Schreiber, Hant, Ishizawar, Jonas, Leverenz, O'Rourke, Wolfe, Zickuhr.

Acquisition of data. Bolster, Kolfenbach, Poeschla, Criscione-Schreiber, Hant, Ishizawar, Jonas, Leverenz, O'Rourke, Wolfe, Zickuhr.

Analysis and interpretation of data. Bolster, Kolfenbach, Poeschla, Zickuhr.

REFERENCES

1. Ferucci ED, Holick P, Day, GM, et al. Factors associated with use of telemedicine for follow-up of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2020;72:1404–9.
2. Hilty DM, Crawford A, Teshima J, et al. A framework for telepsychiartric training and e-health: competencey-based education, evaluation and implications [review]. Int Rev Psychiatry 2015;27:569–92.
3. McDougall JA, Ferucci ED, Glover J, et al. Telerheumatology: a systematic review [review]. Arthritis Care Res (Hoboken) 2017;69: 1546–57.
4. Papanagnou D, Stone D, Chandra S, et al. Integrating telehealth emergency department follow-up visits into residency training. Cureus 2018;10:e2433.
5. Young JD, Abdel-Massih R, Herchline T, et al. Infectious Diseases Society of America position statement on telehealth and telemedicine as applied to the practice of infectious diseases. Clin Infect Dis 2019; 68:1437–43.
6. Lockwood MM, Wallwork RS, Lima K, et al. Telemedicine in adult rheumatology: in practice and in training. Arthritis Care Res (Hoboken) 2022;74:1227–33.
7. Yeoh SA, Young K, Putman M, et al. Rapid adoption of telemedicine in rheumatology care during the COVID-19 pandemic highlights training and supervision concerns among rheumatology trainees. ACR Open Rheumatol 2022;4:128–33.
8. ACR Fellows-in-Training Subcommittee, Saygin D, Lin J, et al. Report of the ACR Fellows-in-Training Subcommittee: experiences of rheumatology fellows early in the COVID-19 pandemic. Arthritis Rheumatol 2022;74:1457–8.
9. Association of American Medical Colleges. Telehealth competencies across the learning continuum. Washington (DC): AAMC; 2021. URL: https://www.aamc.org/data-reports/report/telehealth-competencies.
10. Zickuhr L, Albert DA, Herndon C, et al. Addressing competency in rheumatology telehealth care delivery. Arthritis Care Res (Hoboken) 2023;75:1213–19.
11. Liebowitz JE, Torralba KD, Kolfenbach J, et al. Rheumatology Milestones 2.0: a roadmap for competency-based medical training of rheumatology fellows in the 21st century. Arthritis Care Res (Hoboken) 2022;74:675–85.
12. Battafarano DF, Ditmyer M, Bolster MB, et al. 2015 American College of Rheumatology workforce study: supply and demand projections of adult rheumatology workforce, 2015–2030. Arthritis Care Res (Hoboken) 2018;70:617–6.
13. American College of Rheumatology Committee on Rheumatology Training and Workforce Issues, FitzGerald JD, Battistone M, et al. Regional distribution of adult rheumatologists. Arthritis Rheum 2013; 65:3017–25.
14. Bolster MB, Chandra S, Demaerschalk BM, et al. Crossing the virtual chasm: practical considerations for rethinking curriculum, competency, and culture in the virtual care era. Acad Med 2022; 97:839–46.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2435–2441
DOI 10.1002/acr.25170
© 2023 The Authors. *Arthritis Care & Research* published by Wiley Periodicals LLC on behalf of American College of Rheumatology.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Virtual Learning and Assessment in Rheumatology Fellowship Training: Objective Structured Clinical Examination Revisited

Rachel M. Wolfe,[1] Faye N. Hant,[2] Rumey C. Ishizawar,[3] Lisa G. Criscione-Schreiber,[4] Beth L. Jonas,[3] Kenneth S. O'Rourke,[5] and Marcy B. Bolster[6]

Objective. With the onset of the COVID-19 pandemic, an annual multi-institutional face-to-face rheumatology objective structured clinical examination (ROSCE) was transformed into a virtual format. The educational goals of the virtual ROSCE (vROSCE) were to reproduce the educational value of the previous in-person ROSCE, providing a valuable formative assessment of rheumatology training activities encompassing the 6 Accreditation Council for Graduate Medical Education (ACGME) core competencies for fellows-in-training (FITs). This article describes the novel design, feasibility, and stakeholder value of a vROSCE.

Methods. Through an established collaboration of 5 rheumatology fellowship training programs, in February 2021, a vROSCE was created and conducted using a Zoom platform. Station development included learning objectives, FIT instructions, faculty proctor instructions, and a checklist by which to provide structured formative feedback. An anonymous, optional web-based survey was sent to FIT participants to evaluate the experience.

Results. Twenty-three rheumatology FITs from 5 institutions successfully rotated through 6 stations in the vROSCE. Immediate feedback was given to each FIT using standardized rubrics structured around ACGME core competencies. A total of 65% of FITs (15 of 23) responded to the survey, and 93% of survey respondents agreed or strongly agreed that the vROSCE was a helpful educational activity and identified individualized opportunities for improvement.

Conclusion. A vROSCE is an innovative, feasible, valuable, and well-received educational technology tool. The vROSCE enriched rheumatology FITs' education and offered collaborative learning experiences across institutions.

INTRODUCTION

The objective structured clinical examination (OSCE) is a well-described tool used in medical education, providing both formative and summative assessments (1–4). Since 2006, the Carolinas Fellows Collaborative, a multi-institutional collaboration composed of the fellows-in-training (FITs) and program directors from Duke University, Medical University of South Carolina (MUSC), University of North Carolina (UNC) at Chapel Hill, and Wake Forest (WF) School of Medicine, with the addition of Massachusetts General Hospital (MGH) in 2014, has conducted an annual rheumatology OSCE (ROSCE) (5). This activity provides assessments to individual FITs to help identify knowledge gaps and to the program directors to inform additional training and educational activity development within and between fellowship training programs. The annual ROSCE has included 6–8 stations, assessing FITs' performance through station-specific checklists, with FITs receiving immediate feedback following direct observation by faculty and/or simulated patients. The composite scope of the stations provides assessment across all 6 Accreditation Council for Graduate Medical Education (ACGME) core competencies, which include patient care, medical knowledge, interpersonal and communication skills, professionalism, problem-based learning and improvement, and systems-based practice (6).

To maintain delivery of a yearly ROSCE in the 2020–2021 academic year, a year enveloped by the COVID-19 pandemic, adapting the ROSCE to the virtual setting was necessary to comply with ongoing public health recommendations regarding

[1]Rachel M. Wolfe, MD: Wake Forest School of Medicine, Winston-Salem, North Carolina; [2]Faye N. Hant, DO, MSCR: Medical University of South Carolina, Charleston; [3]Rumey C. Ishizawar, MD, PhD, Beth L. Jonas, MD: University of North Carolina at Chapel Hill, Chapel Hill; [4]Lisa G. Criscione-Schreiber, MD, MEd: Duke University School of Medicine, Durham, North Carolina; [5]Kenneth S. O'Rourke, MD: Maine Medical Center, Portland; [6]Marcy B. Bolster, MD: Harvard Medical School and Massachusetts General Hospital, Boston.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25170.

Address correspondence via email to Rachel M. Wolfe, MD, at rwolfe@wakehealth.edu.

Submitted for publication December 7, 2022; accepted in revised form June 6, 2023.

SIGNIFICANCE & INNOVATIONS

- This study describes the design, feasibility, and stakeholder value of a virtual rheumatology objective structured clinical examination (vROSCE).
- Fellows-in-training (FITs) from 5 institutions successfully completed a vROSCE and received immediate formative feedback using standardized rubrics structured around Accreditation Council for Graduate Medical Education core competencies.
- A novel telemedicine station incorporated a patient actor and case for which learners were evaluated on their virtual patient care delivery and communication skills.
- The vROSCE was rated highly by FITs and identified individualized opportunities for improvement.
- A vROSCE provides an opportunity to leverage multi-institution expertise and cross collaboration across training programs.

restrictions on travel and group meetings. Building on the experience gained by others using virtual formats to perform OSCEs (7,8), we created and implemented a virtual ROSCE (vROSCE). Here we describe the transition of a successful annual multi-institutional in-person ROSCE to a virtual educational and assessment activity with enriched station content for the virtual landscape, learner, and program evaluations, and we identify strengths and limitations of a vROSCE.

SUBJECTS AND METHODS

The vROSCE planning committee (the authors), included a collective 92 years of experience as program directors and comprised 5 current program directors, 1 division director, and 1 community practice rheumatologist. The vROSCE participants included first- and second- year FITs from 5 rheumatology fellowship training programs. All FITs were required to participate unless pre-excused by their respective program director. There were 23 FITs who completed the vROSCE, with the following complement of FITs participating from each institution: 7 from Duke, 5 from MUSC, 4 from UNC, 3 from WF, and 4 from MGH. In total, there were 12 first-year FITs and 11 second-year FITs. The educational aim of the vROSCE was to provide, in a virtual setting, assessments for the 6 ACGME core competencies for FITs as previously provided through an annual in-person ROSCE. The vROSCE was administered using the Zoom virtual platform.

vROSCE design. The vROSCE stations were modeled after ROSCE stations previously implemented at in-person Carolinas Fellows Collaborative ROSCEs and were designed to address fundamental skills for rheumatology FITs, as outlined in the ACGME Program Requirements for Graduate Medical Education (GME) in Rheumatology (9). The vROSCE included 2 parallel circuits of 6 stations run concurrently, with 12 faculty proctors from the participating institutions. Planning faculty developed instruction forms for each station, including station content, learning objectives, and expectations. Station authors provided training and scripting for station proctors to ensure accurate station information, highlight important details and clarifications, and provide guidance on the assessment tools. A unique virtual care telemedicine station, simulating a virtual ambulatory visit, included 2 patient actors; both are actual patients of the telemedicine station proctors. Patient actors were provided similar instruction forms to guide the virtual encounter vROSCE station. Similar to the previous in-person ROSCE, all FITs received instructions and an overview of each vROSCE station scenario and expectations via email prior to participation.

Assessment tools. Each vROSCE station included a checklist tool for FIT performance. Most checklist tools were previously used for similar in-person ROSCE stations, several of which had been previously validated (10). These tools were adapted as indicated for the specific content of the vROSCE. Each station's checklist delineated the ACGME core competencies being assessed; the telehealth vROSCE station additionally incorporated the Association of American Medical Colleges telehealth competencies into the assessment tool (11). These checklist forms provided the basis for the feedback provided to each FIT immediately upon completion of each station during the allotted feedback time.

Station design and competency assessment. The vROSCE stations (Table 1) were designed to assess skills spanning the ACGME core competencies and topics crucial to rheumatology practice, with the intent of providing immediate feedback to FITs to identify areas for growth. Interactive stations were precepted by a faculty proctor and included 1 simulated virtual patient encounter, 2 simulations involving giving advice via telephone to a colleague, and 3 oral examination-style stations. Some stations from the prior in-person ROSCE were not suitable for the virtual format, such as an interactive rehabilitation station with ambulatory aids or braces and a procedural station; thus these were not integrated into the vROSCE. The faculty proctor observed FIT performance with the video turned off to better replicate the true clinical scenario in 2 vROSCE stations: New patient virtual care visit and PCP phone call.

Most performance assessment forms evaluated skills on a 4-point scale: 1) incorrect or not addressed/discussed, 2) incomplete and/or partially correct, 3) generally correct and/or complete, and 4) complete and/or correct. Forms also included global ratings for medical knowledge. For several stations, competencies including professionalism, systems-based practice, and interpersonal communication skills were rated using a 4-point rubric from 1 = not done to 4 = completed fully. Space was included for narrative comments from the proctor. A sample

Table 1. vROSCE station descriptions and learning objectives*

Station	Description of the task	Learning objectives assessed	ACGME competencies evaluated
New patient virtual care visit	Evaluation of a patient with psoriasis presenting with hand pain for a video teleconference visit	AAMC telehealth competencies Evaluation and workup of patient with hand pain	Medical knowledge Patient care and procedural skills Interpersonal and communication skills Professionalism
Radiographs	Evaluation of plain radiographs of 6 patients with arthritis	Interpret plain radiographs Discuss rheumatic disease diagnoses based on plain film findings	Medical knowledge
Bone densitometry	Discuss evaluation and management of a patient with osteoporosis	Interpret bone densitometry scan in a case-based scenario Counsel and manage osteopenia/osteoporosis	Medical knowledge Patient care and procedural skills
PCP phone call	Advise a referring physician calling the consult line with concern for a patient with giant cell arteritis	Counsel a referring primary care physician over the phone regarding a patient with possible giant cell arteritis	Medical knowledge Professionalism Interpersonal and communication skills Systems-based practice
Medication counseling	Counsel physician with a newly pregnant patient with systemic lupus erythematosus by phone	Counsel a local rheumatologist regarding a mutual patient who has lupus and is recently pregnant Discuss use of medications in the setting of pregnancy in a patient with SLE	Medical knowledge Patient care and procedural skills Systems-based practice Interpersonal and communication skills Professionalism
Pathology	Interpret renal pathology slides of a patient with lupus nephritis	Discuss indications for kidney biopsy Interpret pathology findings Discuss treatment regimen options and recommendations in a case-based scenario	Medical knowledge Patient care and procedural skills

* AAMC = Association of American Medical Colleges; ACGME = Accreditation Council for Graduate Medical Education; PCP = primary care physician; SLE = systemic lupus erythematosus; vROSCE = virtual rheumatology objective structured clinical examination.

assessment form can be found in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25170.

Faculty proctors. All 12 faculty proctors are currently or were previously affiliated with academic institutions and are involved in the education and training of FITs. Most of the proctors had prior experience with participating in the traditional in-person ROSCE. During the vROSCE some faculty assessed their own trainees, as the FITs' rotations were randomly assigned.

Scheduling logistics. To accommodate 23 FITs, the group was split in 2 groups (group A and group B). During 90 minutes, group A completed the vROSCE. For logistic purposes, using a separate Zoom link, an additional faculty proctor led group B through the self-administered insurance letter appeal writing exercise outside the vROSCE, permitting some time for a break from on-camera Zoom.

The vROSCE stations were duplicated in parallel. Each station lasted 10 minutes, with 2 additional minutes for feedback, signaled by a Zoom message sent by the program coordinator. After the feedback, the program coordinator shifted FITs to the next assigned station within the 1-minute time allotted for this task. Faculty proctors remained in their virtual breakout rooms throughout the duration of the vROSCE. A sample schedule is shown in Table 2.

At the start of each station, instructions and pertinent clinical information were provided to the FITs via a PowerPoint slide using the shared screen function in each Zoom virtual room. Once the FITs had completed the station or the 2-minute warning was received (whichever happened first), the faculty proctor (and patient actor, if applicable) gave the FITs immediate feedback, reviewing the structured evaluation assessment forms. Upon completion, evaluation forms were sent to the respective program director of each of the FITs for review and use in aggregate. After 90 minutes, the groups were exchanged. Group B rotated to complete the vROSCE circuit in parallel and group A participated in the educational session.

Participant feedback. FITs' feedback was solicited immediately following the vROSCE through an anonymous, web-based survey. Participants rated each station and their overall vROSCE experience using 5-point Likert scales (strongly disagree = 1, disagree = 2, neutral = 3, agree = 4, and strongly agree = 5). For each station, the FITs were asked to evaluate whether audiovisual materials enhanced the presentation, whether material was presented in an interesting and effective manner, and whether the presentation met their educational needs. Additionally, FITs responded as to whether the vROSCE station helped them identify areas for improvement/further study and whether the station was deemed a helpful educational activity. Space was also provided for FITs' narrative comments. Although a formal survey of faculty was not performed, most of the faculty proctors participated in a debriefing following the vROSCE to discuss experiences, challenges, lessons learned, and constructive feedback on the process.

RESULTS

In February 2021, 23 first- and second-year rheumatology FITs from the 4 Carolinas Fellows Collaborative institutions and MGH participated in the ROSCE, which was successfully conducted in a virtual format. There were no technical malfunctions, and the vROSCE was completed in the allotted time, with all FITs completing each of the stations.

Table 2. Sample rotation of fellows*

Start–end times	Room 1: new patient virtual visit	Room 2: radiographs	Room 3: medication counseling	Room 4: pathology	Room 5: PCP phone call	Room 6: bone densitometry
First group						
9:00–9:14 AM	Fellow 1	Fellow 2	Fellow 3	Fellow 4	Fellow 5	Fellow 6
9:14–9:28 AM	Fellow 2	Fellow 3	Fellow 4	Fellow 5	Fellow 6	Fellow 1
9:28–9:42 AM	Fellow 3	Fellow 4	Fellow 5	Fellow 6	Fellow 1	Fellow 2
9:42–9:56 AM	Fellow 4	Fellow 5	Fellow 6	Fellow 1	Fellow 2	Fellow 3
9:56–10:10 AM	Fellow 5	Fellow 6	Fellow 1	Fellow 2	Fellow 3	Fellow 4
10:10–10:24 AM	Fellow 6	Fellow 1	Fellow 2	Fellow 3	Fellow 4	Fellow 5
10:24–10:40 AM	BREAK	–	–	–	–	–
Second group						
10:40–10:54 AM	Fellow 13	Fellow 14	Fellow 15	Fellow 16	Fellow 17	Fellow 18
10:54–11:08 AM	Fellow 14	Fellow 15	Fellow 16	Fellow 17	Fellow 18	Fellow 13
11:08–11:22 AM	Fellow 15	Fellow 16	Fellow 17	Fellow 18	Fellow 13	Fellow 14
11:22–11:36 AM	Fellow 16	Fellow 17	Fellow 18	Fellow 13	Fellow 14	Fellow 15
11:36–11:50 AM	Fellow 17	Fellow 18	Fellow 13	Fellow 14	Fellow 15	Fellow 16
11:50 AM–12:04 PM	Fellow 18	Fellow 13	Fellow 14	Fellow 15	Fellow 16	Fellow 17

* Duplicated in a second circuit for rooms 7–12 (fellows 7–12 and 19–23). PCP = primary care physician.

Participant feedback results. Fifteen of 23 FITs (65%) completed an electronic survey evaluation of the vROSCE educational activity. Respondents included 8 first-year and 5 second-year FITs; 2 FITs did not indicate their training level. To help keep anonymity, the survey did not ask the FITs to identify their institutions, given the small size of each training program. A total of 93% of respondents agreed or strongly agreed that the vROSCE was a helpful educational activity and identified individual opportunities for improvement. For 4 stations, 93% of respondents agreed or strongly agreed that the learning activity was presented in an interesting and effective manner, and 93% agreed or strongly agreed that the station met their educational needs (Table 3). For the remaining 2 stations, new patient virtual care visit and pathology, 73% and 80% of respondents, respectively, agreed or strongly agreed that the material was presented in an interesting and effective manner. For the new patient virtual care visit station, 60% of FITs who responded agreed or strongly agreed that it met their educational needs. In all, 73% agreed or strongly agreed that the pathology station met their educational needs.

The vROSCE performed well compared to the in-person ROSCE evaluations from 2020. The evaluation format (Likert scale and questions) was the same in both years although administered on paper on 2020 and online in 2021. In 2020, 24 of the 26 participants (92%) in the ROSCE completed the evaluation forms. While the stations' content was not identical between the 2 years, we compared 3 stations that evaluated similar educational objectives both years: 1) bone densitometry, 2) radiographs and 3) primary care physician phone call. Although the numbers are too small for meaningful statistical evaluation, the evaluations of the stations were similar in both formats.

Narrative comments from the vROSCE and the traditional ROSCE shared similar themes. FITs noted that both ROSCE formats were valuable to their education and learning. A sample FIT comment from the in-person ROSCE, "Although nerve racking, the ROSCE activity was very stimulating and helpful, revealing areas of knowledge gaps and future areas of learning." From the vROSCE, a FIT commented "Excellently structured [vROSCE]. Very helpful for my learning." In both the ROSCE and vROSCE, some FITs found completing the assigned task in the allotted time difficult on some of the stations, with heterogeneity of the stations mentioned, as the survey question was free-response.

DISCUSSION

Since the COVID-19 pandemic and the ensuing transition to a virtual environment, graduate medical education curricular development has benefitted from a focus on optimizing virtual learning (12). In our case, the vROSCE serves as an innovative form of educational technology that is a feasible way to continue to assess and provide feedback to rheumatology FITs. Educational technology in medicine should be valued for how well the technologic process informs, aids learning, and preserves clinical expertise (13). Although the literature is still limited, virtual OSCEs across specialties and health professions maintain reliability in evaluation despite the transition to a virtual platform (14,15). Equally important, stakeholder evaluations in published studies of virtual OSCEs have been similarly positive, as shown in our study.

Using the virtual platform to perform a ROSCE expands on traditional teaching methods, prompting FITs to interact with educational information in different ways and tapping into educational theory as well. As a teaching strategy that aligns with the best principles of cognitive learning theory and memory formation, the OSCE is an effective method to reduce extraneous load, the effort required to process new information due to the way it is presented (16). In the clinical setting, learners (or FITs) are presented not only with novel presentations of cases but also with many

Table 3. Evaluations of the ROSCE and vROSCE*

	ROSCE 2020 (n = 24)		vROSCE 2021 (n = 15)	
Station	Interesting and effective manner	Met educational need	Interesting and effective manner	Met educational need
New patient virtual care visit	–	–	4.1 (73)	3.9 (60)
Radiographs	4.4 (95)	4.3 (79)	4.6 (93)	4.6 (93)
Bone densitometry	4.1 (62)	4.2 (83)	4.5 (93)	4.5 (93)
PCP phone call	4.6 (88)	4.6 (92)	4.5 (93)	4.5 (93)
Medication counseling	–	–	4.5 (93)	4.5 (93)
Pathology	–	–	4.3 (80)	4.3 (73)
Infusion reaction	4.8 (95)	4.6 (92)	–	–
Physical rehabilitation	4.8 (95)	4.7 (92)	–	–
Pediatric transition station	4.2 (92)	4.3 (79)	–	–

* Values are the number (%). PCP = primary care physician; ROSCE = rheumatology objective structured clinical examination; vROSCE = virtual ROSCE

unpredictable complexities. This load can be decreased by having learners practice whole tasks in increasingly realistic settings (16,17). In our vROSCE, the new patient virtual care visit was truly a telehealth visit, as the FIT, the patient, and even the faculty proctor were not in the same physical location. An OSCE as an instructional technique also represents a form of retrieval practice. Memory/learning is well established to be enhanced and made durable through retrieval practice, such as occurs in an OSCE, as well as with formal testing (18). With these additional facets included, virtual technology allows for more robust pursuit of competency within rheumatology (12,19,20). Additionally, while our study did not aggregate competency data, there is a growing body of evidence to suggest that OSCEs can help identify gaps in competency that may otherwise be missed by standard assessments (21). For our study, each FIT's assessments from the vROSCE were collated and given to their respective program director; subsequently, program directors reviewed with each FIT individually and used program-specific results to adjust scheduled FIT educational content, if need.

Expanding on new educational opportunities that did not exist in an in-person iteration, the vROSCE allowed for participation of faculty preceptors from several institutions. FITs interacted with a larger group of faculty members across institutions, each with unique expertise and teaching skills to impart to FITs. This aspect is particularly important for smaller training programs with limited faculty numbers, commonly occurring in rheumatology, to enrich the experience and exposure of their FITs. Additionally, the vROSCE could allow for faculty expertise across geographically diverse training programs (when previous geographic logistics prohibited participation in the in-person ROSCE). An expanding collective of participating faculty, who help to develop the content, would likely enrich an enduring collection of materials for both future virtual and in-person ROSCEs.

Novel ROSCE stations are also possible in the virtual space. For example, as mentioned above, a new patient virtual care visit was created to directly observe FITs performing this more recently used, and now routine, clinical activity. Given the expanding telemedicine presence, an acknowledged need has arisen to incorporate formal telemedicine curricular elements and assessment of competency in graduate medical education (12,19,22,23). As these skills have not historically been routinely taught, applied, or evaluated, educational technologies like the vROSCE emerge as a useful tool for assessment, adding to the educational armamentarium in this area.

Limitations for this vROSCE included the inability to incorporate hands-on and procedural stations in the virtual setting. FIT feedback raised the concern over sufficient time to complete each station during the vROSCE, which had similarly occurred in the ROSCE. Performing a ROSCE, whether in-person or virtual, requires well-orchestrated timing and a predetermined workflow. Stations within an OSCE must be of the same duration for transitions to successfully occur. Although lengthening ROSCE/vROSCE station times is a consideration, this extension can lead to a longer event, with fatigue for both the FITs and proctors, which has been cited in other OSCEs, specifically virtual OSCEs (8,24). Faculty station developers incorporate this feedback to adjust station design for subsequent ROSCEs and vROSCEs.

It is important to note that the low total numbers of FIT responses (15 total) may skew the results in terms of stakeholder acceptability, despite a high response rate (65%). Aligned with others' experiences (8), more detailed questions in the survey would likely yield more data to help inform future virtual events. Key areas to improve would be to assess whether FITs found the vROSCE station content, as a simulated experience, to be reflective of clinical practice. Additionally, including a formalized faculty survey of the process would help to identify areas for honing and future development.

It is also important to recognize that virtual educational tools, including the vROSCE, decrease direct nonverbal interaction and peer support among FITs. To address these concerns, face-to-face group interactions among FITs during the parallel educational session as well as face-to-face feedback from faculty during the vROSCE was of paramount importance. In developing a vROSCE, considering station duration, participant (FIT and faculty) fatigue, and other opportunities is similarly important, to optimize the educational experience and value.

In conclusion, as many medical education and training activities have transitioned into the virtual environment, educational innovation absolutely must include more than adaptation of didactic learning to the virtual environment. The virtual environment provides the opportunity to look beyond training programs functioning independently with limited cross-collaboration and to better leverage expertise to enhance training. This multi-institutional vROSCE provides an educational technology template for formative assessment of rheumatology FITs.

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. Wolfe 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. Wolfe, Hant, Ishizawar, Criscione-Schreiber, Jonas, O'Rourke, Bolster.

Acquisition of data. Wolfe, Hant, Ishizawar, Criscione-Schreiber, Jonas, O'Rourke, Bolster.

Analysis and interpretation of data. Wolfe, Hant, Ishizawar, Criscione-Schreiber, Jonas, O'Rourke, Bolster.

REFERENCES

1. Criscione-Schreiber L. Turning objective structured clinical examinations into reality [review]. Rheum Dis Clin North Am 2020;46:21–35.
2. Harden RM, Stevenson M, Downie WW, et al. Assessment of clinical competence using objective structured examination. Br Med J 1975; 1:447–51.

3. Kissin EY, Grayson PC, Cannella AC, et al. Musculoskeletal ultrasound objective structured clinical examination: an assessment of the test. Arthritis Care Res (Hoboken) 2014;66:2–6.
4. Berman JR, Aizer J, Bass AR, et al. Fellow use of medical jargon correlates inversely with patient and observer perceptions of professionalism: results of a rheumatology OSCE (ROSCE) using challenging patient scenarios. Clin Rheumatol 2016;35:2093–9.
5. Jonas BL, Criscione-Schreiber LG, Bolster MB, et al. The Carolinas Fellows Collaborative (CFC): an educational model for rheumatology training programs. Arthritis Rheum 2009;60 Suppl:1360.
6. Batalden P, Leach D, Swing S, et al. General competencies and accreditation in graduate education. Health Aff (Millwood) 2002;21:103–11.
7. Blythe J, Patel NS, Spiring W, et al. Undertaking a high stakes virtual OSCE ("VOSCE") during Covid-19. BMC Med Educ 2021;21:221.
8. Hopwood J, Myers G, Sturrock A. Twelve tips for conducting a virtual OSCE. Med Teach 2021;43:633–6.
9. American Board of Internal Medicine. Rheumatology: certification examination blueprint. URL: https://www.abim.org/Media/khzogbsv/rheumatology.pdf.
10. Criscione-Schreiber LG, Sloane RJ, Hawley J, et al. Expert panel consensus on assessment checklists for a rheumatology objective structured clinical examination. Arthritis Care Res (Hoboken) 2015;67:898–904.
11. Association of American Medical Colleges. New and emerging areas in medicine series: telehealth competencies. URL: https://www.aamc.org/data-reports/report/telehealth-competencies.
12. Bolster MB, Chandra S, Demaerschalk BM, et al. Crossing the virtual chasm: practical considerations for rethinking curriculum, competency and culture in the virtual care era. Acad Med 2022;97:839–46.
13. Han H, Resch DS, Kovach RA. Educational technology in medical education [review]. Teach Learn Med 2013;25 Suppl 1:S39–43.
14. Gortney JS, Fava JP, Berti AD, et al. Comparison of student pharmacists' performance on in-person vs. virtual OSCEs in a pre-APPE capstone course. Curr Pharm Teach Learn 2022;14:1116–21.
15. Felthun JZ, Taylor S, Shulruf B, et al. Assessment methods and the validity and reliability of measurement tools in online objective structured clinical examinations: a systematic scoping review. J Educ Eval Health Prof 2021;18:11.
16. Gooding HC, Mann K, Armstrong E. Twelve tips for applying the science of learning to health professions education. Med Teach 2017;39:26–31.
17. Young JQ, Van Merrienboer J, Durning S, et al. Cognitive load theory: implications for medical education: AMEE Guide No. 86. Med Teach 2014;36:371–84.
18. Weinstein Y, Madan CR, Sumeracki MA. Teaching the science of learning [review]. Cogn Res Princ Implic 2018;3:2.
19. Zickuhr L, Albert DA, Herndon C, et al. Addressing competency in rheumatology telehealth care delivery. Arthritis Care Res (Hoboken) 2023;75:1213–9.
20. Liebowitz JE, Torralba KD, Kolfenbach J, et al. Rheumatology Milestones 2.0: a roadmap for competency-based medical training of rheumatology fellows in the 21st century. Arthritis Care Res (Hoboken) 2022;74:675–85.
21. Lebdai S, Bouvard B, Martin L, et al. Objective structured clinical examination versus traditional written examinations: a prospective observational study. BMC Med Educ 2023;23:69.
22. Sartori DJ, Olsen S, Weinshel E, et al. Preparing trainees for telemedicine: a virtual OSCE pilot. Med Ed 2019;53:517–8.
23. Charon R. The patient-physician relationship. Narrative medicine: a model for empathy, reflection, profession, and trust. JAMA 2001;286:1897–902.
24. Guled M, Islam J, Qureshi H. Response to: twelve tips for conducting a virtual OSCE. Med Teach 2022;44:101.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2442–2452
DOI 10.1002/acr.25168
© 2023 The Authors. *Arthritis Care & Research* published by Wiley Periodicals LLC on behalf of American College of Rheumatology.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

American College of Rheumatology
Empowering Rheumatology Professionals

Pediatric Rheumatology Care and Outcomes Improvement Network's Quality Measure Set to Improve Care of Children With Juvenile Idiopathic Arthritis

Catherine A. Bingham,[1] Julia G. Harris,[2] Tingting Qiu,[3] Mileka Gilbert,[4] Sheetal S. Vora,[5] Cagri Yildirim-Toruner,[6] Kerry Ferraro,[7] Daniel J. Lovell,[3] Janalee Taylor,[3] Melissa L. Mannion,[8] Jennifer E. Weiss,[9] Ronald M. Laxer,[10] Michael Shishov,[11] Edward J. Oberle,[12] Beth S. Gottlieb,[13] Tzielan C. Lee,[14] Nancy Pan,[15] Jon M. Burnham,[7] Danielle C. Fair,[16] Michelle Batthish,[17] Melissa M. Hazen,[18] Charles H. Spencer,[19] and Esi M. Morgan,[20] for the Pediatric Rheumatology Care and Outcomes Improvement Network

Objective. To describe the selection, development, and implementation of quality measures (QMs) for juvenile idiopathic arthritis (JIA) by the Pediatric Rheumatology Care and Outcomes Improvement Network (PR-COIN), a multihospital learning health network using quality improvement methods and leveraging QMs to drive improved outcomes across a JIA population since 2011.

Methods. An American College of Rheumatology–endorsed multistakeholder process previously selected initial process QMs. Clinicians in PR-COIN and parents of children with JIA collaboratively selected outcome QMs. A committee of rheumatologists and data analysts developed operational definitions. QMs were programmed and validated using patient data. Measures are populated by registry data, and performance is displayed on automated statistical process control charts. PR-COIN centers use rapid-cycle quality improvement approaches to improve performance metrics. The QMs are revised for usefulness, to reflect best practices, and to support network initiatives.

Results. The initial QM set included 13 process measures concerning standardized measurement of disease activity, collection of patient-reported outcome assessments, and clinical performance measures. Initial outcome measures were clinical inactive disease, low pain score, and optimal physical functioning. The revised QM set has 20 measures and includes additional measures of disease activity, data quality, and a balancing measure.

Conclusion. PR-COIN has developed and tested JIA QMs to assess clinical performance and patient outcomes. The implementation of robust QMs is important to improve quality of care. PR-COIN's set of JIA QMs is the first comprehensive set of QMs used at the point-of-care for a large cohort of JIA patients in a variety of pediatric rheumatology practice settings.

INTRODUCTION

Launched in 2011, the Pediatric Rheumatology Care and Outcomes Improvement Network (PR-COIN) is a learning health network to improve implementation of evidence-based care for pediatric rheumatic diseases using quality improvement science (1). Leveraging the Institute for Health Care Improvement Breakthrough Series approach and the model for improvement, PR-

Supported by an anonymous private donor, the Cincinnati Children's Research Foundation Place Outcomes Research Award, the Arthritis Foundation (grant AHRQ U19-HS-021114), the Pfizer Independent Grant for Learning and Change, and the Patient Centered Outcomes Research Institute (grant RI-PCC-2017).

[1]Catherine A. Bingham, MD: Penn State Children's Hospital and Penn State College of Medicine, Hershey, Pennsylvania; [2]Julia G. Harris, MD: Children's Mercy Hospitals and Clinics, Kansas City, Missouri; [3]Tingting Qiu, MPH, Daniel J. Lovell, MD, MPH, Janalee Taylor, APRN, CNP-BC: Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; [4]Mileka Gilbert, MD, PhD: Medical University of South Carolina, Charleston; [5]Sheetal S. Vora, MD: Levine Children's Hospital and Atrium Health, Charlotte, North Carolina; [6]Cagri Yildirim-Toruner, MD: Baylor College of Medicine and Texas Children's Hospital, Houston; [7]Kerry Ferraro, MISM, Jon M. Burnham, MD, MS: Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; [8]Melissa L. Mannion, MD: University of Alabama at Birmingham; [9]Jennifer E. Weiss, MD: Hackensack University Medical Center and Hackensack Meridian Health, Hackensack, New Jersey; [10]Ronald M. Laxer, MDCM, FRCPC: The Hospital for Sick Children and the University of Toronto, Toronto, Ontario, Canada; [11]Michael Shishov, MD, MPH: Phoenix Children's Hospital, Phoenix, Arizona; [12]Edward J. Oberle, MD: Nationwide Children's Hospital and The Ohio State University, Columbus; [13]Beth S. Gottlieb, MD, MS: Cohen Children's Medical Center of New York and Zucker School of Medicine at Hofstra/Northwell, Queens, New York; [14]Tzielan C. Lee, MD: Stanford Medicine Children's Health, Stanford University, Stanford, California; [15]Nancy Pan, MD: Hospital for Special Surgery and Cornell University, New York, New York; [16]Danielle C. Fair, MD: Medical College of Wisconsin and Children's Wisconsin, Milwaukee; [17]Michelle Batthish, MD, MSc: McMaster Children's Hospital and McMaster University, Hamilton, Ontario, Canada; [18]Melissa M. Hazen, MD, PhD: Boston Children's Hospital, Boston, Massachusetts; [19]Charles H. Spencer, MD: University of Mississippi Medical Center, Jackson; [20]Esi M. Morgan, MD, MSCE: Seattle Children's Hospital and the University of Washington, Seattle.

SIGNIFICANCE & INNOVATIONS

- Our article describes the selection and development of a quality measure (QM) set for the Pediatric Rheumatology Care and Outcomes Improvement Network (PR-COIN).
- We illustrate how QMs are used to monitor processes of care and disease outcomes to optimize the care provided to children with juvenile idiopathic arthritis (JIA).
- PR-COIN's set of JIA QMs is the first comprehensive set of QMs used at the point-of-care for a large cohort of JIA patients in a variety of pediatric rheumatology practice settings.

COIN emulated ImproveCareNow, a learning health network for pediatric inflammatory bowel disease, whose work has increased remission rates (2–4). Learning health networks are an effective organizational infrastructure to enact quality improvement in health care (5,6). PR-COIN is a growing voluntary network of 23 academic pediatric rheumatology centers in the US and Canada. Its mission is "to build a thriving and inclusive community of patients, families, clinical teams, and researchers that uses quality improvement science to deliver exceptional and equitable health care to children with rheumatic disease and to bring research discovery to patients faster." PR-COIN initially focused its efforts on improving care of children and adolescents with juvenile idiopathic arthritis (JIA) (7).

JIA, which affects approximately 1 per 1,000 children, is a chronic autoimmune disease complicated by significant morbidity and disability (8–12). Large pediatric rheumatology research organizations, such as the Pediatric Rheumatology Collaborative Study Group, the Pediatric Rheumatology International Trials Organization, and the Childhood Arthritis and Rheumatology Research Alliance, have facilitated clinical trials and comparative effectiveness studies to overcome barriers to studying rare pediatric rheumatic diseases and to improve outcomes (13,14). PR-COIN team members work to improve patient outcomes by facilitating the rapid adoption of best practices, quicker implementation of new research findings, and elimination of unintended variation in care.

A learning health network needs to develop and track quality measures (QMs) to assess whether changes/interventions are resulting in improvement. Process measures of clinical performance track patients' receipt of care per best practices. Outcome QMs are crucial to assess response to interventions on patient outcomes. A recent systematic review of QMs for inflammatory arthritis found that the vast majority of QM sets were for adult rheumatoid arthritis (RA) (15). Lovell et al published a proposed set of process QMs for JIA in 2011 with the intention that they would be tested and used by the newly forming PR-COIN (16). There are no previous studies reporting the performance of these process QMs for patients with JIA. Additionally, there is a gap in the literature with respect to JIA outcome QMs. Although there are several different standardized outcome assessments for JIA, QMs with specific, well-defined operational definitions are lacking (17). Similarly, a recent review of QMs for inflammatory arthritis found that only 3% of published QMs for RA assessed outcomes (15). Therefore, we describe the selection, development, and implementation of JIA QMs in PR-COIN and how they are used to monitor not only processes of care but also disease outcomes to optimize patient care.

PATIENTS AND METHODS

PR-COIN membership and structure. PR-COIN was launched in 2011, and the initial cohort included 12 pediatric rheumatology member centers. A coordinating center provides quality improvement specialist consultation, quality improvement education and maintenance of certification programs, data management, data analysis, legal and regulatory oversight, program management and administration, and supports the network development. The Anderson Center for Health Systems Excellence at Cincinnati Children's Hospital created a Learning Network Core (led by Carole Lannon, MD, MPH) that fostered a staff experienced with QM development and coordinating learning health networks (5,6). PR-COIN operationalizes activities within a committee structure, including Quality Measures, Outcomes, Informatics, Family Engagement (Parent Working Group, Patient Advocacy Team), Scientific Development and Oversight (Research), and Finance and External Partnership Committees. Committee leads form an Executive Committee that with the Improvement Advisor oversee network improvement activities consistent with the network's mission/vision. A volunteer Steering Committee reviews and approves network activities. Participating sites of PR-COIN are shown in Appendix A.

A shared patient registry platform was designed to aggregate structured JIA patient data collected at the point-of-care to enable monitoring of QM performance. Patient or parent consent was obtained for participation in PR-COIN, or an Institutional Review Board waiver of need for consent was granted at the PR-COIN sites. Two sites had to leave the network due to lack of resources. Thirteen additional pediatric rheumatology centers joined PR-COIN over time, so the current number of participating PR-COIN sites is 23.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25168.

Address correspondence via email to Catherine A. Bingham, MD, at cbingham@pennstatehealth.psu.edu.

Submitted for publication December 20, 2022; accepted in revised form May 25, 2023.

PR-COIN member centers are typically rheumatology practices based at an academic medical center. Each center forms a local quality improvement team composed of pediatric rheumatologists (clinicians and clinical researchers), nurses, medical assistants, social workers, occupational and physical therapists, and research coordinators. Some centers have local quality improvement specialist support. Because the voices of patients and families are invaluable to inform the challenges to care and impact of disease, patient/parent representatives are also included. Members conduct quality improvement projects using rapid-cycle improvement (e.g., plan-do-study-act cycles) at their respective clinics, collaborate on network-wide initiatives, and share best practices via monthly webinars and at semi-annual conferences (7). Supportive tools for QM use include change packages that guide PR-COIN sites with implementation of quality improvement interventions.

QM selection process. Figure 1 illustrates the QM development process. PR-COIN's initial JIA QM set was developed by a pilot working group in 2011. The pilot working group included pediatric rheumatologists, a data analyst, a data manager, a programmer, and a quality improvement specialist. A formal PR-COIN Measures Committee was assembled in 2013 and was composed of the above members as well as volunteers from center teams and JIA parent representatives who served as a liaison with the PR-COIN Parent Working Group to share valuable parent/patient feedback.

PR-COIN representatives had previously participated in American College of Rheumatology (ACR) JIA QM development

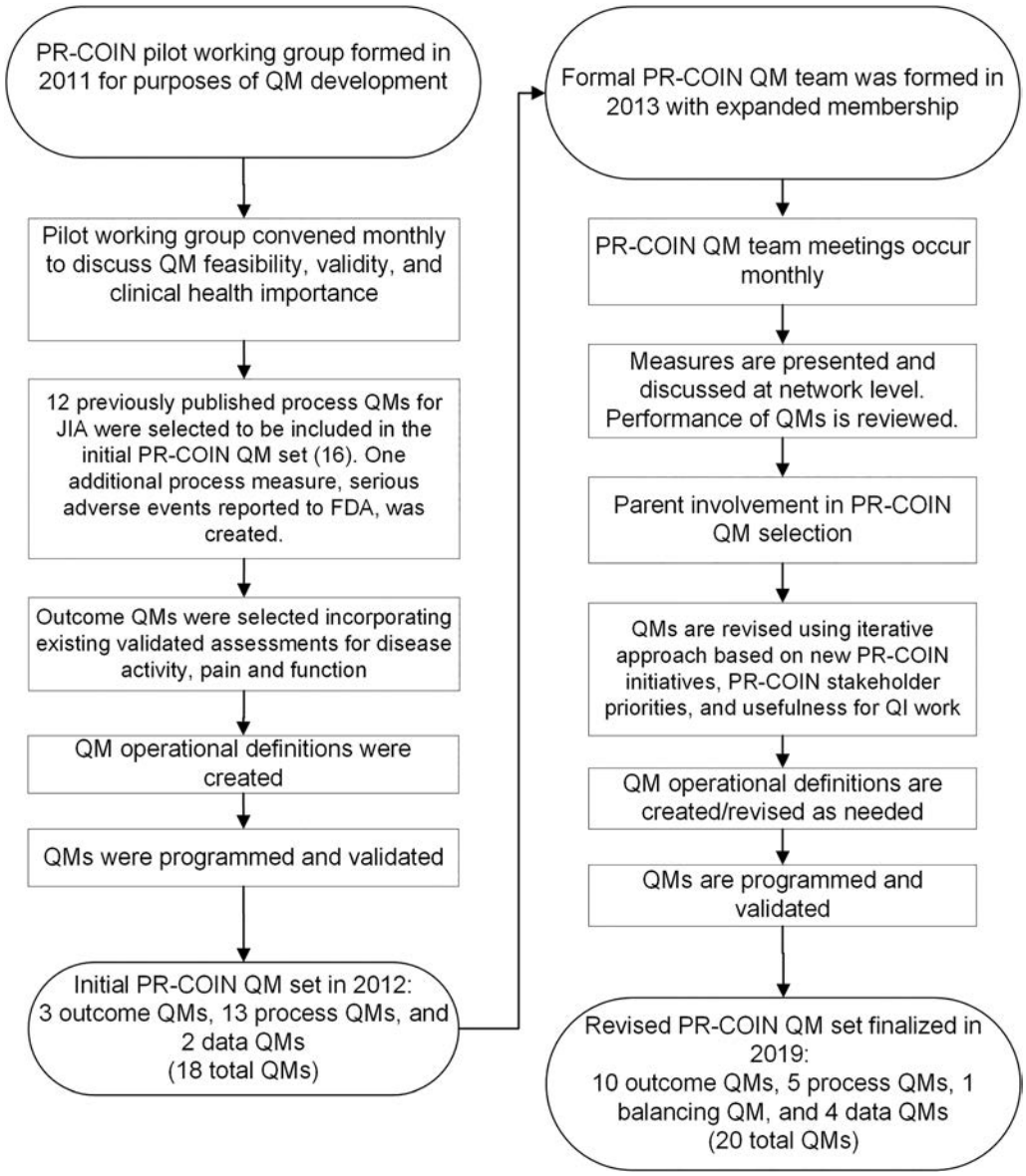


Figure 1. Flow diagram. Pediatric Rheumatology Care and Outcomes Improvement Network (PR-COIN) quality measure (QM) development. FDA = Food and Drug Administration; JIA = juvenile idiopathic arthritis; QI = quality improvement.

published by Lovell et al, which was informed by surveys completed by patients and parents, so these published JIA process QMs were used as a starting point (16). Given the patient-centered focus of PR-COIN, patient-reported outcome measures were prioritized along with provider-reported outcomes. Based on clinical experience and prior focus group work with parents that prioritized relief of pain and physical function as outcomes, the pilot working group chose outcome measures that incorporated existing published validated JIA outcome assessments in those domains (18–20).

Measurement of disease activity was also emphasized. Outcome measures assessed included clinical inactive disease, which requires the patient to meet all of the following: active joint count = 0, physician global assessment (PhGA) of disease activity score = 0, morning stiffness duration ≤15 minutes, no active uveitis, no elevation of serum inflammatory marker (if measured), and no active systemic features for children with systemic JIA (18). The Juvenile Arthritis Disease Activity Score (JADAS) outcome measures were later adopted. The clinical JADAS10 (cJADAS10) version was selected as the scoring system, as this instrument was most feasible at the point-of-care. The cJADAS10 is the sum of patient (or parent) global assessment (PtGA) of overall well-being, number of joints with active arthritis based on a complete joint examination (up to a maximum value of 10), and PhGA (21). Exclusion of an inflammatory marker (which is incorporated as part of the regular JADAS score) increased the simplicity and feasibility of calculating the cJADAS10 score at the point-of-care so that data could be used in clinical decision-making during the visit. We used published cutoffs for cJADAS10 scores that designate states of inactive disease and minimal, moderate, or high disease activity for certain JIA subtypes (22).

PR-COIN quality improvement initiatives and priorities helped shape the PR-COIN QM set over time. A network-wide self-management support initiative led to the development of a self-management support QM. In addition, when a treat-to-target strategy was adopted by PR-COIN as a network-wide initiative due to evidence suggesting this intervention would help improve outcomes, process QMs for treat-to-target were devised and added to the measurement set to track performance on PR-COIN's treat-to-target project (23–25). In this manner, new research and best practices informed iterative development and selection of QMs. PR-COIN's treat-to-target intervention is described elsewhere (23).

Factors considered in QM selection included the importance of the measure to stakeholders, feasibility of data collection at PR-COIN centers, ability to influence performance on a measure, sensitivity to change, and usefulness for quality improvement work. Incorporating stakeholder input, an iterative approach was taken to finetune the list of QMs. Consensus on QMs selected was reached via committee discussions. The team developed an operational definition for each process and outcome measure that included numerator, denominator, and inclusion and exclusion criteria. Once a measure was programmed, sample clinical data were used to test each measure for validation.

QMs were created to monitor the quality of PR-COIN registry data. PR-COIN created a data QM that tracks the percentage of visits with completeness of all critical data elements recorded. The critical data elements are the data needed to be able to calculate performance on outcome measures. Another data QM monitors the timeliness of enrollment into the PR-COIN registry after JIA diagnosis, to ensure inclusion of newly diagnosed patients, who typically have higher disease activity. A third data QM looks at whether data are entered into the registry soon after clinic visits. This measure helps ensure that there is timely data entry into the registry, as this will impact the usefulness of the data to improve care and the validity of QMs as a reflection of the state of clinical practice. Finally, the percentage of eligible patients with JIA at PR-COIN sites who are enrolled in PR-COIN is reported monthly (based on the total number of JIA patients seen at a given site, assessed annually), to ensure that the sample is representative, inclusive, and generalizable to the entire clinic population.

In 2018, PR-COIN reconsidered the entire QM set due to movement to a new registry platform. Some QMs were revised and updated, and others were discontinued if no longer needed for quality improvement work. Operational definitions for all the QMs were reviewed, clarified, and revalidated in the new registry. A new revised set of QMs was finalized in 2019, with programming completed in 2020. Transition to another registry platform in 2022 has seen repetition of this process of measure reprogramming and validation of the 2019 measure set.

Statistical methods for QM performance display. Statistical process control is used for analysis and interpretation of data in PR-COIN (26). Statistical process control is the accepted methodology used in quality improvement, as it is better suited to consider random variation in measurement and allows for more timely analysis of performance than traditional statistical analysis methods (26). Statistical process control charts are used to display performance on the QMs for both site-specific and aggregate data (Figure 2). Data are displayed with center line (mean) and upper and lower control limits (±3 SE of the mean) to monitor for statistically significant change in performance, known as special cause variation. Probability-based rules are used to analyze control charts to detect evidence of change (26,27). Center lines can be shifted to indicate statistical change in performance when 8 successive points occur on the same side of the center line, when 6 successive points increase or decrease, or if a data point occurs outside the control limits. PR-COIN team members monitor site and aggregate control charts regularly to see whether change in performance on the QMs has occurred.

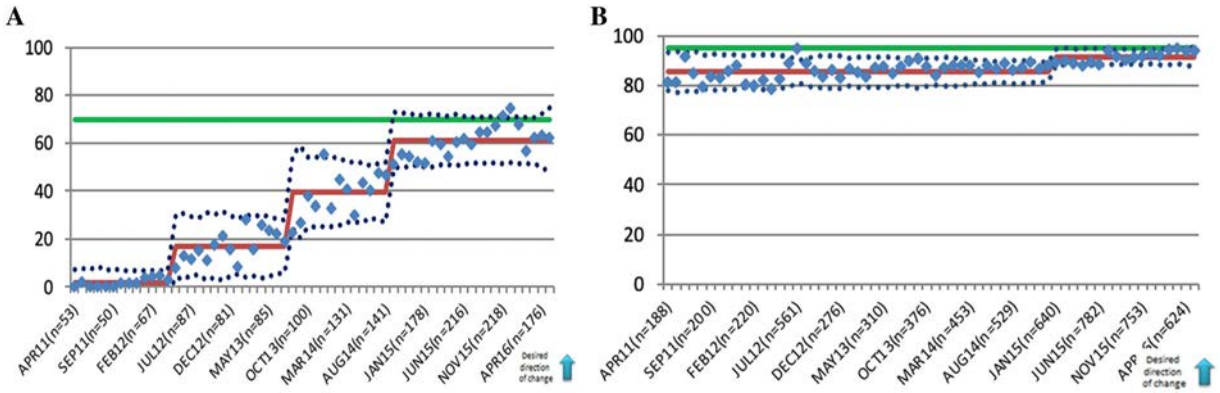


Figure 2. Pediatric Rheumatology Care and Outcomes Improvement Network juvenile idiopathic arthritis process measure examples. **A**, Percentage of patients on disease-modifying antirheumatic drugs who had visits in the month and had medication counseling within the past 12 months. **B**, Percentage of visits where physician's global assessment of disease activity was performed. Y axis represents the percent who achieved the measure; X axis represents time; red line = mean; blue diamond = aggregate data each month; dotted lines = control limits; green line = goal.

RESULTS

Thirteen process measures, 3 outcome measures, and 2 data QMs were adopted by PR-COIN as the initial set of JIA QMs (Table 1). Twelve process measures were selected using the 2011 published proposed set of process QMs for JIA (16). These measures assess the collection of patient- and provider-reported outcome scores, uveitis screening, and medication safety laboratory monitoring. One additional process measure, serious adverse event reporting to the US Food and Drug Administration, was created to focus on safety of medications used to treat JIA. The PR-COIN pilot working group created 3 outcome QMs, including the percentage of patients with clinical inactive disease, arthritis pain score ≤3, and optimal physical function.

Figure 2 shows graphic examples of PR-COIN performance on 2 initial process measures. Some sites found that they were not performing and/or not documenting certain aspects of recommended care (such as medication counseling or performing a joint count). Auditing allowed sites to see which items they were not doing reliably and to take steps to improve. For some process QMs, including collection of arthritis-related pain and PhGA, PR-COIN achieved high performance that was sustained, successfully completing this measure >90% of the time. For example, PR-COIN sites collected complete joint counts in 99% of patients from 2012 through 2016. PR-COIN sites that were not collecting/documenting patient-reported outcome measures and physician-reported measurements, i.e., PhGA scores or active joint counts, before joining the network were able to start collecting these data with high reliability. By tracking performance over time, PR-COIN observed sustainability in collection of these components in a variety of practice settings over a 4-year period. Therefore, these process measures were later removed from the revised QM set due to mastery. Measurement of health-related quality of life (HRQoL) was identified as important. However, due to the length of HRQoL questionnaires and associated licensing fees, this measure was not considered broadly feasible for adoption. Tuberculosis screening for patients on biologics was retired from the PR-COIN QM set because screening is now required by insurance companies during the prior authorization process. Aiming to keep the number of measures manageable and desiring to focus on outcome measures, PR-COIN removed other process measures such as behavioral/medication counseling for patients on disease-modifying antirheumatic drugs. New QMs were added to measure performance on prioritized interventions.

The performance on certain measures did not change over time or required a prolonged time to show change. For example, the network was unable to improve uveitis screening, even though many PR-COIN sites focused on quality improvement interventions designed to improve this metric, so this measure was removed from the QM set. Failure to improve this measure may have been due to the measure definition relying on an external party (ophthalmologist) to obtain confirmatory documentation. In addition, the QM "percentage of patients with clinical inactive disease" was slow to change (Figure 3). The state of clinical inactive disease is a hard outcome to achieve as this factor is a binary measure. Patients can move in and out of an inactive disease state with a change in only 1 parameter. Therefore, the need clearly arose for a continuous outcome measure that could detect incremental improvement over time. This need led to adoption of the JADAS outcome measures (21,28). The JADAS QMs permit PR-COIN to track whether patients are moving into lower disease activity states, even if they have not yet reached clinical inactive disease. Figure 3 illustrates 1 JADAS outcome measure adopted by PR-COIN. Describing the extent of JIA outcome improvements, which requires detailed report of the interventions and further analysis, is beyond the scope of this article and will be a subject of future publications.

Table 1. PR-COIN JIA quality measures*

Measure classification, subgroup, and measure title	Original measure	Current measure	Reason code for measure addition or deletion†
Outcome			
Disease control			
JIA patients with oligoarthritis or polyarthritis with inactive disease or low disease activity by cJADAS10	–	X	A
Mean cJADAS10 score for all JIA patients		X	A
JIA patients with clinical inactive disease	X	X	–
Mean active joint count for all JIA patients	–	X	A
JIA patients with oligoarthritis or polyarthritis who achieve inactive or low disease activity by 6 months	–	X	A
Quality of life			
Patients with optimal physical function	X	X	–
Patients with pain score ≤3	X	X	–
Patients with patient global assessment ≤2	–	X	B
Mean patient global assessment of overall well-being	–	X	B
Patients with pain interference T score <60	–	X	B
Process			
Model treatment			
Polyarticular course patients with treatment target set	–	X	C
Visits with provider attestation of disease activity status for T2T	–	X	C
Visits where clinical decision support was used	–	X	C
Patients who received self-management support	–	X	D
Safety monitoring			
Ongoing DMARD toxicity laboratory monitoring	X	X	–
Baseline toxicity laboratory monitoring for patients starting DMARDs	X	–	E
Behavioral counseling for patients starting DMARDs‡	X	–	E
Annual behavioral counseling for patients on DMARDs‡	X	–	E
Tuberculosis screening prior to starting biologic	X	–	E
Annual tuberculosis screening for patients on biologics	X	–	E
Serious adverse events reported to FDA	X	–	E
Disease activity monitoring			
Uveitis screening per Heiligenhaus guidelines (ref. 41)	X	–	F
Complete joint count every 180 days	X	–	G
Physician global assessment of disease activity at every visit	X	–	G
Patient-reported assessment			
Functional assessment every 180 days	X	–	E
Health-related quality of life every 180 days	X	–	H
Arthritis-related pain at every visit	X	–	G
Balancing			
Time between hospitalization for infections for all patients	–	X	C
Data quality			
Complete data for clinical inactive disease	X	–	E
All critical data recorded	–	X	I
Patients with a visit recorded in last 13 months	–	X	I
Patients enrolled in PR-COIN within 90 days of diagnosis	–	X	I
Percentage of JIA population that is enrolled in PR-COIN	X	X	–

* cJADAS10 = clinical Juvenile Arthritis Disease Activity Score 10; DMARD = disease-modifying antirheumatic drug; FDA = Food and Drug Administration; JIA = juvenile idiopathic arthritis; PR-COIN = Pediatric Rheumatology Care and Outcomes Improvement Network.
† A: Measure added because PR-COIN desires additional disease activity quality measures (QMs); B: Measure added due to parent/patient input about what outcome was important to them; C: Measure added due to treat-to-target (T2T) initiative in PR-COIN; D: Measure added due to self-management support initiative in PR-COIN; E: Measure retired due to being less useful for quality improvement work or deemed lower priority; F: Measure retired due to lack of responsiveness to multiple improvement efforts; G: Measure retired due to high performance rate over time; H: Measure abandoned due to lack of feasibility; I: New data QMs selected based on new QMs and data quality areas identified that need improvement.
‡ Behavioral counseling was later changed to the term "medication counseling."

Parent engagement in PR-COIN was instrumental in shaping our new revised QM set. Patients and parents conveyed the importance of the single-item measure PtGA of overall well-being. Therefore, 2 related outcome QMs were created and added to the revised QM set by the PR-COIN Measures Committee. Figure 4 shows 1 PR-COIN patient-reported outcome measure, the percentage of patients who had a PtGA overall well-being score ≤2. The percentage of patients who reported a low pain score and optimal physical functioning were among the original PR-COIN measures and remained in the revised QM set, as they are highly valued by patients/parents and providers. PR-COIN's revised set of QMs was finalized in 2019. Table 1 shows both the original and revised PR-COIN JIA QMs. The rationale for addition or discontinuation of a measure is noted.

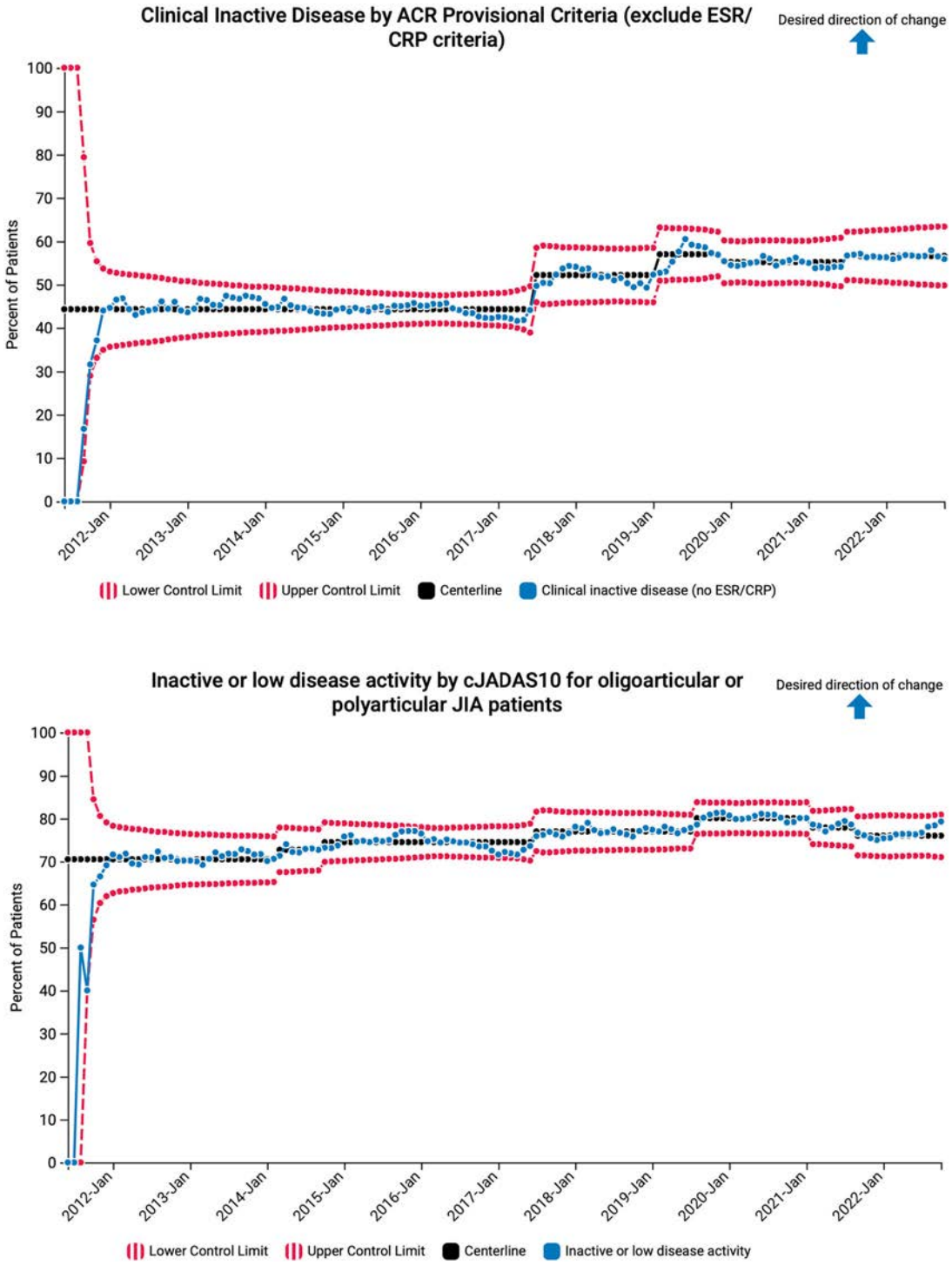


Figure 3. Pediatric Rheumatology Care and Outcomes Improvement Network juvenile idiopathic arthritis (JIA) disease activity outcome measures. ACR = American College of Rheumatology; cJADAS10 = clinical Juvenile Arthritis Disease Activity Score 10; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate.

DISCUSSION

With patient/parent and clinician input, we selected from existing JIA QMs and developed de novo QMs to allow for robust quality improvement for patients with JIA in PR-COIN. Certain attributes should be considered when creating QMs, including clinical and public health importance, scientific validity, acceptability, feasibility, reliability, and sensitivity to change (29,30). A systematic review of QMs for inflammatory arthritis published in 2018, which conducted a quality appraisal of 13 QM sets published in the literature, also emphasized that

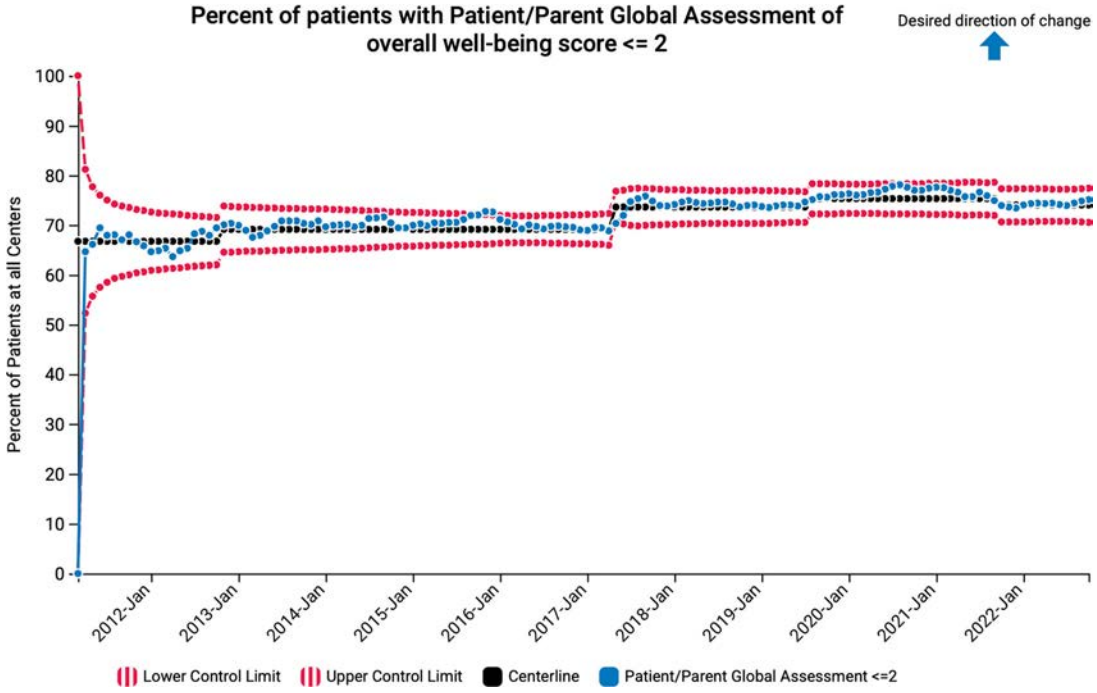


Figure 4. Pediatric Rheumatology Care and Outcomes Improvement Network juvenile idiopathic arthritis patient-reported outcome measure.

measures should be well-defined and incorporate appropriate stakeholder input (15).

Striving for improvement in JIA outcomes is an important, worthwhile task. JIA is the most common pediatric rheumatic disease and was identified as a priority disease category for future ACR QM development (29). PR-COIN JIA QMs, based on evidence, were developed using ACR published recommendations for process QMs in JIA and standardized JIA outcome assessments currently used in pediatric rheumatology (16–18,31). In addition to having clinical importance and being evidence-based, PR-COIN QMs are characterized by several other strengths. Stakeholder input has been paramount. Measures have been selected based on the importance to patients and families. Additionally, rheumatology providers from different practice settings have contributed to selection of QMs, based on usefulness and feasibility in clinical practice.

PR-COIN teams have developed processes and tools to support ease and practicality of QM use across pediatric rheumatology centers with different resources. The centralized PR-COIN patient registry serves as the platform to facilitate data collection for components needed to calculate performance on the QMs. The burden of data collection to calculate the QMs must be considered with respect to feasibility of measurement. One solution in PR-COIN has been the creation of a structured template embedded in the electronic medical record (EMR) to collect discrete data elements during clinic visits, e.g., an Epic SmartForm that is available to any customer (32). These data can then be extracted from the EMR and securely transferred into the PR-COIN registry on an automated basis, reducing the need for double data entry. This process of creating a system of electronic data capture from an EMR to feed into a centralized quality improvement registry has been described elsewhere (33). Other centers rely on standardized data collection forms and web-based data entry. By having access to the unique environments of multiple pediatric rheumatology practices conducting routine clinical care and collecting process and outcome data at each clinic visit, PR-COIN has been able to test JIA QMs for feasibility in everyday practice. The learning health network provides an ideal setting in which to test for both achievability and practicality of QMs.

Elements such as which patients are included, the period at risk, and the reporting period must be precisely delineated in operational definitions (30). PR-COIN measures are well-defined with careful attention to detail in the operational definitions, which are readily accessible to users within the PR-COIN registry platform. Revisions to the operational definitions are made over time, as needed, based on experience with using the measures and the subsequent realization of items needing further standardization or clarification.

Many previously published QM sets for inflammatory arthritis have not been tested in practice (15). By tracking site and aggregate measure performance since 2011, PR-COIN has been afforded a unique opportunity to evaluate how the JIA QMs perform over time. Stakeholders analyze and revise QMs as necessary to make sure the measure set remains useful and to determine whether new measures are needed. For example, the Measures Committee created the cJADAS outcomes measures once a need for an outcome measure that is more sensitive to change became evident. A report of lessons learned during

measure development for inflammatory bowel disease sheds light on the importance of revisiting and revising QMs over time (34). According the ACR White Paper on Quality Measurement, QMs should be reviewed and updated every 3 years (29).

The Institute of Medicine highlights 6 key domains of quality in health care, including effectiveness, safety, patient-centeredness, timeliness, efficiency, and equity (35). PR-COIN's QM set addresses the first 3 of these domains. The PR-COIN JIA QM set clearly addresses effectiveness of care, as there are 5 outcome measures focused on disease activity and 5 outcome measures pertaining to quality of life. These outcome measures incorporate both patient- and provider-reported outcomes. The medication safety laboratory monitoring QM and time between hospitalization for infections QM monitor safety of immunosuppressive medications. Several PR-COIN JIA QMs focus on patient-centeredness, including the measures assessing self-management support, treat-to-target, PtGA overall well-being, arthritis-related pain, and physical function.

Some domains of quality in health care have not been included in the PR-COIN JIA QM set. PR-COIN does not measure access to care, despite recognizing its importance, in part due to difficulty obtaining data from each hospital scheduling center. Patient experience during care delivery, efficiency of care, and equity of care are key domains where there is opportunity for further work. In the review of QMs for inflammatory arthritis, Cooper et al found that only 11% addressed timeliness of patient care, 12% efficiency of care, and 1% equity (15). Assessment of health equity underscores the importance of complete QM data ascertainment across the entire population of JIA patients served.

When collecting patient data for QMs, making sure that the majority of patients eligible for inclusion in the measure are being counted is essential, in order to have accurate representation of performance on that QM for the patient population being studied. PR-COIN sites strive to enroll at least 75% of JIA patients at their site to have a representative population, but optimal enrollment has not yet been achieved at some sites. In addition, patient data elements necessary to calculate performance on the QMs are sometimes not collected or captured in the registry. Due to incomplete data, not all enrolled patients are included in the measure performance. The data QMs assist PR-COIN sites in monitoring missing data. Efforts are underway in PR-COIN to improve data completeness for race, ethnicity, primary language, and insurance status to measure dimensions of health equity.

A weakness of our final QM set is that these measures were not all created using formal consensus methods (e.g., the Delphi process) apart from the process measures in the initial QM set previously published (16). However, the outcome measures (pain, physical function, overall well-being) were subsequently all selected via Delphi and consensus voting to be included in the Outcome Measures in Rheumatology core domain set for JIA, validating their inclusion in PR-COIN (36). Other PR-COIN QMs were developed with specific network-prioritized interventions in mind, e.g., treat-to-target, and were therefore derived by committee recommendations. Other groups have used formal consensus building techniques when creating QMs, but there is variability in techniques used (15,37,38). The ACR has published the process they use to endorse QMs used in rheumatology and has endorsed QMs for several rheumatic diseases, including RA (29,39). Those working on future QM development in PR-COIN and other learning health networks may formalize the process of consensus building using standard methodologies. PR-COIN QMs were developed for purposes of continuous quality improvement efforts by centers who voluntarily assembled into a learning health network and were not developed to be used for systems of reimbursement by third party health care payers. Hence, further analysis of the measures and risk adjustment might be warranted prior to using PR-COIN measures for pay-for-performance programs (40).

In summary, PR-COIN's set of JIA QMs is the first comprehensive set of QMs used at the point-of-care for a large cohort of patients with JIA in a variety of hospital-based pediatric rheumatology practice settings. The PR-COIN JIA QMs demonstrate strong clinical importance, acceptance by stakeholders, and feasibility with performance tracked over 10 years. A need exists for performance outcome measures in rheumatology, and PR-COIN has helped to advance this work, particularly with regard to testing the measures in practice. PR-COIN will continue to use stakeholder expertise and reevaluate the JIA QM set to drive quality improvement work, with the ultimate goal of improving health outcomes in patients with rheumatic diseases.

ACKNOWLEDGMENTS

This project was conducted with the cooperation and assistance of the Pediatric Rheumatology Care and Outcomes Improvement Network (PR-COIN) and the physicians, providers, and families participating in this multicenter quality improvement collaborative (www.pr-coin.org). The authors would like to thank all of the parents, patients, and providers who contribute to the PR-COIN mission. PR-COIN would like to thank the American College of Rheumatology for its support in hosting the PR-COIN Registry (2011–2018). The financial support of participating member institutions defrays infrastructure costs. The authors would like to thank Peter Margolis, MD, PhD, and Carole Lannon, MD, MPH, at Cincinnati Children's Hospital Medical Center (CCHMC) Anderson Center for Health Systems Excellence for their mentorship, vision, and tireless support of the learning health network community in pediatric medicine. We also acknowledge the supporting contributions to this work of the coordinating team at CCHMC Anderson Center (Kendra Wiegand and Drew Warmin), and of the CCHMC Biomedical Informatics (Emily Chesnut and Stephanie Loos). We thank the Seattle Children's Research Institute, which became the PR-COIN Coordinating Center in 2022. This article is dedicated to the memory of Nancy Minnick-Griffin, long-time Quality Improvement Specialist to PR-COIN, a treasured friend and valued colleague who challenged, inspired, and taught us.

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. Bingham 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. Bingham, Harris, Gilbert, Vora, Yildirim-Toruner, Ferraro, Lovell, Taylor, Morgan.

Acquisition of data. Bingham, Harris, Qiu, Gilbert, Vora, Yildirim-Toruner, Lovell, Taylor, Mannion, Weiss, Laxer, Shishov, Oberle, Gottlieb, Lee, Pan, Burnham, Fair, Batthish, Hazen, Spencer, Morgan.

Analysis and interpretation of data. Bingham, Harris, Qiu, Vora, Yildirim-Toruner, Lovell, Taylor, Morgan.

REFERENCES

1. Pediatric Rheumatology Care and Outcomes Improvement Network. URL: http://pr-coin.org.
2. Institute for Healthcare Improvement. The breakthrough series: IHI's collaborative model for achieving breakthrough improvement. 2003. URL: www.IHI.org.
3. Langley GJ, Moen RD, Nolan KM, et al. The improvement guide: a practical approach to enhancing organizational performance. 2nd ed. San Francisco (CA): Jossey-Bass; 2009.
4. Crandall WV, Margolis PA, Kappelman MD, et al. Improved outcomes in a quality improvement collaborative for pediatric inflammatory bowel disease. Pediatrics 2012;129:e1030–41.
5. Lannon CM, Miles PV. Pediatric collaborative improvement networks: bridging quality gaps to improve health outcomes. Pediatrics 2013; 131 Suppl 4:S187–8.
6. Billett AL, Colletti RB, Mandel KE, et al. Exemplar pediatric collaborative improvement networks: achieving results. Pediatrics 2013;131 Suppl 4:S196–203.
7. Harris JG, Bingham CA, Morgan EM. Improving care delivery and outcomes in pediatric rheumatic diseases [review]. Curr Opin Rheumatol 2016;28:110–6.
8. Moorthy LN, Peterson MG, Hassett AL, et al. Burden of childhood-onset arthritis. Pediatr Rheumatol Online J 2010;8:20.
9. Guzman J, Oen K, Tucker LB, 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.
10. Tollisen A, Selvaag AM, Aulie HA, et al. Physical functioning, pain, and health-related quality of life in adults with juvenile idiopathic arthritis: a longitudinal 30-year followup study. Arthritis Care Res (Hoboken) 2018;70:741–9.
11. Minden K, Niewerth M, Listing J, et al. Long-term outcome in patients with juvenile idiopathic arthritis. Arthritis Rheum 2002;46:2392–401.
12. Shoop-Worrall SJ, Verstappen SM, McDonagh JE, et al. Long-term outcomes following achievement of clinically inactive disease in juvenile idiopathic arthritis: the importance of definition. Arthritis Rheumatol 2018;70:1519–29.
13. Ruperto N, Martini A. Networking in paediatrics: the example of the Paediatric Rheumatology International Trials Organisation (PRINTO) [review]. Arch Dis Child 2011;96:596–601.
14. Ringold S, Nigrovic PA, Feldman BM, et al. The Childhood Arthritis and Rheumatology Research Alliance consensus treatment plans: toward comparative effectiveness in the pediatric rheumatic diseases. Arthritis Rheumatol 2018;70:669–78.
15. Cooper M, Rouhi A, Barber CE. A systematic review of quality measures for inflammatory arthritis. J Rheumatol 2018;45:274–83.
16. Lovell DJ, Passo MH, Beukelman T, et al. Measuring process of arthritis care: a proposed set of quality measures for the process of care in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011;63:10–6.
17. Consolaro A, Schiappapietra B, Dalpra S, et al. Optimisation of disease assessments in juvenile idiopathic arthritis [review]. Clin Exp Rheumatol 2014;32 Suppl 85:S126–30.
18. Wallace CA, Giannini EH, Huang B, et al. American College of Rheumatology provisional criteria for defining clinical inactive disease in select categories of juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011;63:929–36.
19. Filocamo G, Davi S, Pistorio A, et al. Evaluation of 21-numbered circle and 10-centimeter horizontal line visual analog scales for physician and parent subjective ratings in juvenile idiopathic arthritis. J Rheumatol 2010;37: 1534–41.
20. Singh G, Athreya BH, Fries JF, et al. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 1994;37:1761–9.
21. Consolaro A, Ruperto N, Bazso A, et al. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. Arthritis Rheum 2009;61:658–66.
22. Trincianti C, Van Dijkhuizen EH, Alongi A, et al. Definition and validation of the American College of Rheumatology 2021 Juvenile Arthritis Disease Activity Score cutoffs for disease activity states in juvenile idiopathic arthritis. Arthritis Rheumatol 2021;73:1966–75.
23. El Tal T, Ryan ME, Feldman BM, et al. Consensus approach to a treat-to-target strategy in juvenile idiopathic arthritis care: report from the 2020 PR-COIN consensus conference. J Rheumatol 2022;49:497–503.
24. Smolen JS, Aletaha D, Bijlsma JW, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010;69:631–7.
25. Ravelli A, Consolaro A, Horneff G, et al. Treating juvenile idiopathic arthritis to target: recommendations of an international task force. Ann Rheum Dis 2018;77:819–28.
26. Benneyan JC, Lloyd RC, Plsek PE. Statistical process control as a tool for research and healthcare improvement [review]. Qual Saf Health Care 2003;12:458–64.
27. Provost LP, Murray SK. The health care data guide: learning from data for improvement. 1st ed. San Francisco (CA): Jossey-Bass; 2011.
28. McErlane F, Beresford MW, Baildam EM, 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.
29. Saag KG, Yazdany J, Alexander C, et al. Defining quality of care in rheumatology: the American College of Rheumatology white paper on quality measurement [review]. Arthritis Care Res (Hoboken) 2011;63:2–9.
30. Suter LG, Barber CE, Herrin J, et al. American College of Rheumatology white paper on performance outcome measures in rheumatology. Arthritis Care Res (Hoboken) 2016;68:1390–401.
31. McErlane F, Beresford MW, Baildam EM, et al. Recent developments in disease activity indices and outcome measures for juvenile idiopathic arthritis [review]. Rheumatology (Oxford) 2013;52:1941–51.
32. Research toolkit: capturing clinical data to advance research on juvenile idiopathic arthritis (JIA). Verona (WI): Epic Systems Corporation; 2019.
33. Marsolo K, Margolis PA, Forrest CB, et al. A digital architecture for a network-based learning health system: integrating chronic care management, quality improvement, and research. EGEMS (Wash DC) 2015;3:1168.
34. Crandall WV, Boyle BM, Colletti RB, et al. Development of process and outcome measures for improvement: lessons learned in a quality improvement collaborative for pediatric inflammatory bowel disease [review]. Inflamm Bowel Dis 2011;17:2184–91.
35. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the quality chasm: a new health system for the 21st century. Washington (DC): National Academy Press; 2001.
36. Morgan EM, Munro JE, Horonjeff J, et al. Establishing an updated core domain set for studies in juvenile idiopathic arthritis: a report from the OMERACT 2018 JIA workshop. J Rheumatol 2019;46:1006–13.
37. Heiligenhaus A, Foeldvari I, Edelsten C, et al. Proposed outcome measures for prospective clinical trials in juvenile idiopathic arthritis-

associated uveitis: a consensus effort from the multinational interdisciplinary working group for uveitis in childhood [review]. Arthritis Care Res (Hoboken) 2012;64:1365–72.
38. MacLean CH, Saag KG, Solomon DH, et al. Measuring quality in arthritis care: methods for developing the Arthritis Foundation's quality indicator set. Arthritis Rheum 2004;51:193–202.
39. Yazdany J, Robbins M, Schmajuk G, et al. Development of the American College of Rheumatology's rheumatoid arthritis electronic clinical quality measures. Arthritis Care Res (Hoboken) 2016;68:1579–90.
40. Lane-Fall MB, Neuman MD. Outcomes measures and risk adjustment. Int Anesthesiol Clin 2013;51:10–21.
41. Heiligenhaus A, Niewerth M, Ganser G, et al. Prevalence and complications of uveitis in juvenile idiopathic arthritis in a population-based nation-wide study in Germany: suggested modification of the current screening guidelines. Rheumatology (Oxford) 2007;46:1015–9.

APPENDIX A: PEDIATRIC RHEUMATOLOGY CARE AND OUTCOMES IMPROVEMENT NETWORK PARTICIPATING SITES

The Pediatric Rheumatology Care and Outcomes Improvement Network participating sites are Boston Children's Hospital, Boston, MA; Children's of Alabama, University of Alabama at Birmingham, Birmingham, AL; Children's Hospital of Philadelphia, Philadelphia, PA; Children's Mercy Kansas City, Kansas City, MO; Cincinnati Children's Hospital Medical Center, Cincinnati, OH (Coordinating Center 2011–2021); Cincinnati Children's Research Informatics Shared Facility, Cincinnati, OH (Registry and Data Coordinating Center 2011–2022); Cohen Children's Medical Center of New York, Queens, NY; Hackensack University Medical Center, Hackensack Meridian Health, Hackensack, NJ; Hospital for Special Surgery and Weill Medical College of Cornell University, New York, NY; Levine Children's Hospital, Atrium Health, Charlotte, NC; London Health Sciences Centre, Western University, London, ON, Canada; McMaster Children's Hospital, McMaster University, Hamilton, ON, Canada; Medical College of Wisconsin–Children's Wisconsin, Milwaukee, WI; Medical University of South Carolina, Charleston, SC; Nationwide Children's Hospital, The Ohio State University, Columbus, OH; Nemours Children's Hospital, Orlando, FL; Penn State Children's Hospital, Penn State College of Medicine, Hershey, PA; Phoenix Children's Hospital, Phoenix, AZ; Seattle Children's Hospital, University of Washington School of Medicine, Seattle, WA (Seattle Children's Research Institute serves as the Coordinating Center 2022–present); Stanford Medicine Children's Health, Stanford University, Stanford, CA; Texas Children's Hospital, Baylor College of Medicine, Houston, TX; The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada; University of Minnesota Masonic Children's Hospital, Minneapolis, MN; and University of Mississippi Medical Center, Jackson, MS.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2453–2462
DOI 10.1002/acr.25171
© 2023 American College of Rheumatology.

American College of Rheumatology
Empowering Rheumatology Professionals

Proposed Response Parameters for Twelve-Month Drug Trial in Juvenile Systemic Sclerosis: Results of the Hamburg International Consensus Meetings

Ivan Foeldvari,[1*] Kathryn S. Torok,[2*] Jordi Anton,[3] Michael Blakley,[4] Tamás Constantin,[5] Megan Curran,[6] Maurizio Cutolo,[7] Christopher Denton,[8] Kim Fligelstone,[9] Francesca Ingegnoli,[10] Suzanne C. Li,[11] Dana Němcová,[12] Catherine Orteu,[13] Clarissa Pilkington,[14] Vanessa Smith,[15] Anne Stevens,[16] Jens Klotsche,[17] Dinesh Khanna,[18] Patrícia Costa-Reis,[19] Francesco Del Galdo,[20] Bernd Hinrichs,[21] Ozgur Kasapcopur,[22] Clare Pain,[23] Nicolino Ruperto,[24] Alison Zheng,[25] and Daniel E. Furst[26]

Objective. Juvenile systemic sclerosis (SSc) is an orphan disease, associated with high morbidity and mortality. New treatment strategies are much needed, but clearly defining appropriate outcomes is necessary if successful therapies are to be developed. Our objective here was to propose such outcomes.

Methods. This proposal is the result of 4 face-to-face consensus meetings with a 27-member multidisciplinary team of pediatric rheumatologists, adult rheumatologists, dermatologists, pediatric cardiologists, pulmonologists, gastroenterologists, a statistician, and patients. Throughout the process, we reviewed the existing adult data in this field, the more limited pediatric literature for juvenile SSc outcomes, and data from 2 juvenile SSc patient cohorts to assist in making informed, data-driven decisions. The use of items for each domain as an outcome measure in an open label 12-month clinical trial of juvenile SSc was voted and agreed upon using a nominal group technique.

Results. After voting, the domains agreed on were global disease activity, skin, Raynaud's phenomenon, digital ulcers, musculoskeletal, cardiac, pulmonary, renal, and gastrointestinal involvement, and quality of life. Fourteen outcome measures had 100% agreement, 1 item had 91% agreement, and 1 item had 86% agreement. The domains of biomarkers and growth/development were moved to the research agenda.

Conclusion. We reached consensus on multiple domains and items that should be assessed in an open label, 12-month clinical juvenile SSc trial as well as a research agenda for future development.

INTRODUCTION

Juvenile systemic sclerosis (SSc) is an orphan disease with an estimated prevalence of 3 in 1,000,000 children, with a high morbidity (1,2). Currently no medications are licensed for juvenile SSc. This proposal is for an open label, 12-month clinical trial in juvenile SSc. To develop such a trial, and for use in any well-done treatment trial in juvenile SSc, clearly defining outcomes and tailoring the outcomes for juvenile SSc is necessary.

The only existing outcome scoring system that has been developed specifically for use in the juvenile SSc population is the Juvenile Systemic Sclerosis Severity Score, which was adapted from an adult SSc severity score, the Medsger Severity

[1]Ivan Foeldvari, MD: Schön Klinik Hamburg Eilbek, Hamburg, Germany; [2]Kathryn S. Torok, MD: University of Pittsburgh and Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania; [3]Jordi Anton, MD: Hospital Sant Joan de Déu and Universitat de Barcelona, Barcelona, Spain; [4]Michael Blakley, MD: Indiana University School of Medicine and Riley Hospital for Children at IU Health, Indianapolis; [5]Tamás Constantin, MD: Semmelweis University, Budapest, Hungary; [6]Megan Curran, MD: University of Colorado School of Medicine and Children's Hospital Colorado, Aurora; [7]Maurizio Cutolo, MD: University of Genoa and IRCCS San Martino Polyclinic Hospital, Genoa, Italy; [8]Christopher Denton, MD: Royal Free London NHS Foundation Trust, London, UK; [9]Kim Fligelstone: Scleroderma & Raynaud's United Kingdom, London, UK; [10]Francesca Ingegnoli, MD: University of Milan, ASST G. Pini, Milan, Italy; [11]Suzanne C. Li, MD: Hackensack University Medical Center, Hackensack, New Jersey; [12]Dana Němcová, MD: Charles University, Prague, Czech Republic; [13]Catherine Orteu, MD: Royal Free London, London, UK; [14]Clarissa Pilkington, MD: Great Ormond Street Hospital, London, UK; [15]Vanessa Smith, MD: Ghent University, Ghent University Hospital, VIB Inflammation Research Center, and ERN ReCONNET, Ghent, Belgium; [16]Anne Stevens, MD, PhD: Children's Hospital Research Institute and University of Washington, Seattle, and Janssen Pharmaceutical Companies of Johnson & Johnson, Spring House, Pennsylvania; [17]Jens Klotsche, PhD: German Rheumatism Research Center, Berlin, Germany; [18]Dinesh Khanna, MD: University of Michigan, Ann Arbor; [19]Patrícia Costa-Reis, MD: Hospital de Santa Maria, Faculdade de Medicina, and Universidade de Lisboa, Lisbon, Portugal; [20]Francesco Del Galdo, MD: University of Leeds and Leeds Teaching Hospital Trust, Leeds, UK; [21]Bernd Hinrichs, MD: Asklepios Klinik Nord–Heidberg, Hamburg, Germany; [22]Ozgur Kasapcopur, MD: Cerrahpasa Medical School and Istanbul University–Cerrahpasa, Istanbul, Turkey; [23]Clare Pain, MD: Alder Hey Children's Foundation

SIGNIFICANCE & INNOVATIONS

- This is the first proposal for outcome measures for a 12-month open label clinical trial for juvenile systemic sclerosis (SSc).
- The proposed outcome measures span the main domains of the organ system involvement in juvenile SSc.
- Patient-reported outcomes are included in the outcomes proposed.

Score (3,4). More recent efforts have been proposed for disease activity indices in SSc (rather than severity/damage scores), which would be more sensitive to clinical change and applicable to clinical trials. Such activity indices for adult SSc include measures like the Composite Response Index in Systemic Sclerosis (CRISS) (5), which is based on a 2-step approach. First, significant disease worsening or new-onset organ damage is defined as nonresponsiveness. In patients who did not fulfill the criteria of part 1, a probability of improvement is calculated for each patient, based on the modified Rodnan Skin Score (mRSS), the percent predicted forced vital capacity (FVC%), the physician global assessment (PhGA), and the patients' Health Assessment Questionnaire disability index (HAQ DI). These efforts have been applied only to adult SSc (5–7), and there is no such disease activity or response index for juvenile SSc. Our goal was to define disease activity outcome parameters in juvenile SSc that would be sensitive to change and useful in an open label, 12-month clinical trial in juvenile SSc.

MATERIALS AND METHODS

These recommendations were developed over a span of years by a dedicated group of multinational pediatric and adult scleroderma experts who are interested in juvenile SSc and outcome measure development, starting with electronic surveys in 2014 and refining juvenile SSc outcome domains and items through an annual face-to-face meeting, through Delphi and nominal group technique processes, hosted at the Hamburg Symposium of Juvenile Scleroderma starting in 2014–2018 (see Figure 1 for details). The final 2018 juvenile SSc consensus meeting is explained in detail here. By consensus, the meeting recommended 12 domains and 22 items for an open-label, 12-month clinical juvenile SSc trial.

In December 2018, international pediatric and adult rheumatology scleroderma experts, dermatologists, pediatric cardiologists, pulmonologists, gastroenterologists, a statistician, and patients met for a 2-day conference. The first day was dedicated to a series of talks and discussions regarding an adult scleroderma expert presentation of the CRISS (by DK) and items included in the CRISS, with possible pediatric performance and adaptations (8), a "lessons learned" talk regarding the response of clinical outcomes from recent clinical trials in adult SSc (by CD), cohort data from the International Juvenile Systemic Sclerosis Inceptions cohort (n = 150) (IF and JK) (1) and from the Childhood Arthritis and Rheumatology Research Alliance (CARRA) juvenile SSc cohort (n = 64) (KST) (9) relevant to the prior voted domains and items of interest, and finally, with pediatric scleroderma workgroup presentations on updates of the various organ systems in juvenile SSc and related outcomes. These discussions provided a background for the second day of the conference, whose goal, using the nominal group technique (by DEF), was to develop consensus recommendations for items to be used in an open, 12-month clinical trial in children with SSc (not clinical practice or general research). The items (n = 22) and domains (n = 12) remaining after the 2017 Hamburg consensus meeting (Figure 1) were reevaluated at the 2018 meeting. Twenty-two of the 27 multidisciplinary members at the 2018 conference voted, with 75% (16 of 22) having been at the preceding 2017 consensus conference (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25171).

General guidelines. Some general guidelines were discussed and agreed upon at the start of the second day of the consensus 2018 meeting, including the following: validated outcomes should have priority; outcomes validated in adults with SSc would be sufficient for application in juvenile SSc; although previously agreed items (from the 2016 and the 2017 consensus meetings) were defined in terms of change, those items will be defined in terms of their absolute value, independent of change per se (e.g., the item "change in mRSS" would now be "mRSS"); and estimation of change from baseline and the significance of change would be examined through statistical analysis. For uniformity and clarity when patients or clinicians used the measures, a scale of 0–10 or 0–100 was to be employed when visual analog scales (VAS) were used as items. The specific length of the scale could be decided on a protocol basis. The minimum clinically important difference (MCID) for any VAS was to be 1.0 for 0–10 and 10.0 for 0–100 scales. This decision was slightly less than that in the literature (7–27 mm, dependent on baseline pain), but participants felt it was easy to use and remember (22 of 22 agreed). Unfortunately, other MCIDs were usually not available

NHS Trust, Liverpool, UK; [24]Nicolino Ruperto, MD: IRCCSG Istituto G. Gaslini, Genoa, Italy; [25]Alison Zheng, MD: Chinese Organization for Scleroderma, Chengdu City, Sichuan Province, China; [26]Daniel E. Furst, MD: University of California, Los Angeles, University of Washington, Seattle, and University of Florence, Florence, Italy.

*Drs. Ivan Foeldvari and Kathryn S. Torok were dual first authors and contributed equally to this work.

Author disclosures are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25171.

Address correspondence via email to Ivan Foeldvari, MD, at sprechstunde@kinderrheumatologie.de and Kathryn S. Torok, MD, at kathryn.torok@chp.edu.

Submitted for publication November 24, 2022; accepted in revised form June 8, 2023.

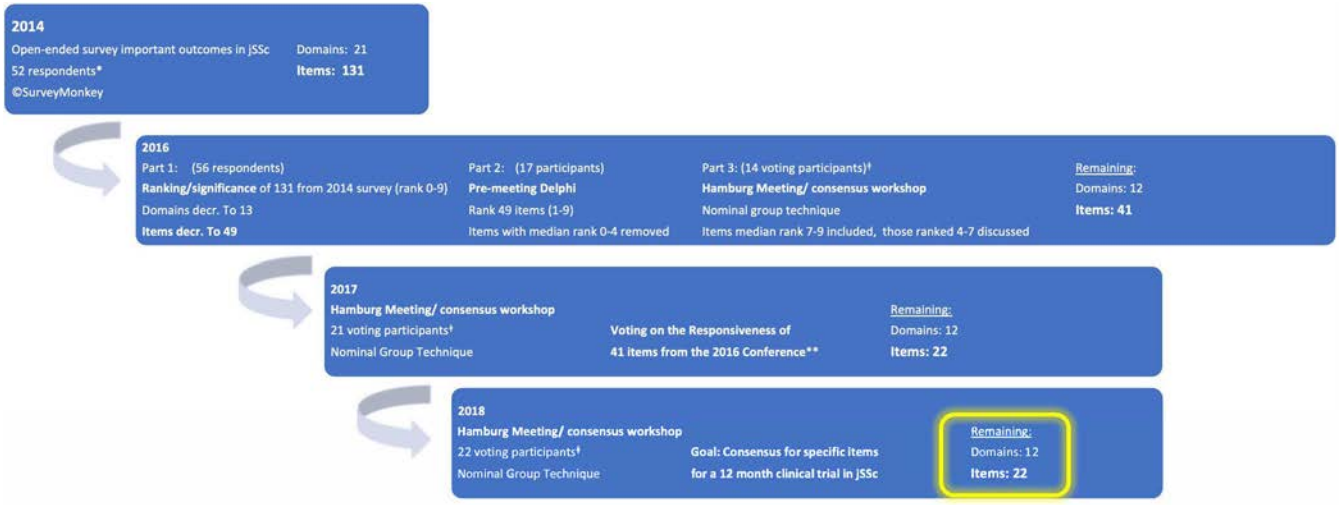


Figure 1. Outcomes important for juvenile systemic sclerosis (SSc) were determined from 2014 to 2016, later defined in context to responsiveness in 2017, and ultimately refined to those appropriate for a 12-month clinical trial in juvenile SSc. The final list includes 22 items within 12 domains through voting at in-person consensus meetings. * = 2014 respondents: all participants of the pediatric rheumatology email board, the members of the members of the Paediatric Rheumatology European Society (PRES) juvenile scleroderma working group and the active participants of the juvenile scleroderma inception cohort project were invited to participate. In total, 70% of the respondents were experienced pediatric rheumatologists (>10 years of experience in the field). The mean number of patients followed-up by respondents was 12.3 juvenile SSc patients. Total number of patients followed-up by all respondents was 574. In all, 95% of respondents work at academic medical hospitals. decr. = decreased; jSSc = juvenile systemic sclerosis; ǂ = moderated by DEF; ** = items also considered in context of the adult Composite Response Index in Systemic Sclerosis, developed by Dinesh Khanna (The American College of Rheumatology Provisional Composite Response Index for Clinical Trials in Early Diffuse Cutaneous Systemic Sclerosis. Arthritis Rheumatol 2016;68:299–311).

for consideration or voting. There was also consensus (22 of 22 agreed) that the time frame for any VAS was to be 7 days unless specifically stated differently. The CRISS, a validated combined measure of response in adult SSc, although discussed at length the day prior, was not voted upon as a composite outcome during the second day consensus meeting, since it comprised multiple important elements that were instead individually voted on within their respective domain.

The consensus process. The process included the following: review of each of the 12 domains and items within each domain, led by the moderator (DEF). Some minimal background was first given for orientation (usually from the leader of the organ working group); during discussion, there was to be 1 speaker at a time, voting (22 members, later 21 members as 1 member had to leave) would close the discussion and consensus (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25171). Voting was not anonymous, and options included agree, disagree, or do not know. Consensus was defined a priori, as agreement by ≥80% of voting members present. If consensus was not reached, more discussion ensued, and ultimately, without consensus, the item was recorded as "no consensus reached," and if applicable, referred to the research agenda.

Three scribes compared notes after the meeting to ensure accuracy. KT merged the notes and DEF reviewed and edited. Consultants were invited to participate in the prior day's meeting and provided some discussion points during the consensus meeting, but they refrained from voting (see Acknowledgments and Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25171). There were 2 SSc patients present (AZ, KF), both currently adults with ages of onset of 8 years and 26 years, who actively participated and were voting members.

RESULTS

Domain 1: patient global assessment (PtGA) and PhGA of disease activity. The PhGA and PtGA VAS (0–100 mm) of disease activity over the previous 7 days have been used in the Juvenile Systemic Sclerosis Inceptions cohort (1), with data in 47 juvenile SSc patients over 12 months demonstrating an MCID of 20 of 100 mm change ($P < 0.001$) in PhGA and a 15 of 100 change in PtGA ($P < 0.001$). Voting was unanimous (22 of 22 for each) to use the PhGA and PtGA of disease activity in juvenile SSc trials (Table 1 and Supplementary Appendices A and B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25171). There was consensus to include knowledge of the patient's previously available clinical data (22 total: 19 agreed, 3 disagreed, 0 did not know). By

Table 1. Domains and items suggested as outcome measures for a 12-month clinical trial in juvenile systemic sclerosis (SSc) from the 2018 International Consensus Meeting*

	Metric, range	Considerations
Physician-measured outcomes		
Physician global assessment of disease activity	VAS: 0–10 or 0–100	Should take into account past 7 days Allowable to consider patients features/conditions compared to prior visit Same physician to assess at study visit for clinical trial
Modified Rodnan Skin Score	Whole number score: 0–51	Physical examination at date of study visit Consider other cutaneous findings in context to scoring children's skin
DUCAS	Number scale 0.5 digit: 0–19.5	Physical examination at date of study visit
Musculoskeletal, total active joint count	Whole number score: 0–75	Physical examination at date of study visit Number of joints that have *either* joint swelling or LOM with pain/tenderness that is considered secondary to juvenile SSc
Cardiac		
Left ventricular ejection fraction	Echocardiogram value, % (30–80)	Echocardiogram closest to date of study visit
New-onset LV failure	Echocardiogram evaluation, yes/no	Echocardiogram closest to date of study visit
New-onset clinically important arrythmia	EKG evaluation, yes/no	EKG closest to study date
Development of pulmonary arterial hypertension	Echocardiogram evaluation, yes/no	Echocardiogram closest to date of study visit
Pulmonary		
Forced vital capacity	Pulmonary function test (PFT) value, % of predicted (20–100)	PFT closest to study date Several demographic variables collected to calculate international standard
DLco	PFT value, % of predicted (20–100)	PFT closest to study date (age eligible) Hemoglobin collected to determine hemoglobin-corrected DLco value
6-minute walk test (6MWT)	Walking test with respiratory therapist, meters (0–700)	6MWT closest to study visit Lowest SpO2 during the test also important to evaluate desaturation Forehead or ear probe preferred over finger probe (Raynaud's)
Development of new scleroderma renal crisis	Clinical phenotype present, yes/no	Blood value abnormalities in setting new hypertension
Body mass index	Measurement for pediatrics using Z scores; Z ≤ –2 is flagged as malnutrition	Weight and height used to calculate
Patient-reported outcomes		
Patient global assessment of disease activity	VAS: 0–10 or 0–100	Should take into account past 7 days Parent of child to fill, depending on age (typically age ≥8 years can self-report) Must be consistent person scoring over the length of the trial
Global health and function		
C-HAQ	Score 0–3 (without any difficulty to unable to do); total score, divided among the 8 domains scores, which are modified if aids or devices are used	Patients age <16 years If child is age <8 years, a parent will fill in this form; for age ≥8 years, if developmentally appropriate, the child will fill this form Timeframe in the past 7 days
Health Assessment Questionnaire	Score 0–3, same scoring system as C-HAQ	Patients age ≥16 years Traditional HAQ, which has been widely validated
Organ systems and general		
VAS captured in the C-SHAQ and SHAQ	VAS: 0–100	Same questions C-SHAQ and SHAQ, since childhood version adapted from adult Patients age ≥16 years fill out SHAQ Timeframe in the past 7 days
Affected by pain because of scleroderma	VAS: 0–100	General, global health
Intestinal problems interfered with daily activities	VAS: 0–100	Gastrointestinal domain

(Continued)

Table 1. (*Cont'd*)

	Metric, range	Considerations
Breathing problems interfered	VAS: 0–100	Pulmonary domain
Raynaud's phenomenon interfered	VAS: 0–100	Raynaud's phenomenon domain
Finger ulcers interfered	VAS: 0–100	Digital ulcers domain
All the ways (pain, discomfort, limitations of daily life, body changes)	VAS: 0–100	General, global health

* C-HAQ = Childhood Health Assessment Questionnaire; C-SHAQ = Childhood Scleroderma Health Assessment Questionnaire; DLco = diffusion capacity of the lungs for carbon monoxide; DUCAS = Digital Ulcer Clinical Assessment Score; EKG = electrocardiogram; HAQ = Health Assessment Questionnaire; LOM = Limitation of Motion; LV = Left Ventricular; SHAQ = Scleroderma Health Assessment Questionnaire; SpO2 = oxygen saturation; VAS = visual analog scale.

general agreement, voters recommended having instructions in the protocol or in the "Manual of Procedures" as to how the PhGA of disease activity was to be done and specified that it was to be performed by the same investigator at each visit. Voters also unanimously agreed that either child or parent may answer the PtGA of disease activity (age and child dependent), so long as it is consistent throughout the protocol. Patients ages ≥8 years are encouraged to complete patient-reported outcomes as is routine for several pediatric rheumatology registry studies (9).

Domain 2: patient-reported global health and function. Patient-reported outcomes are essential in clinical drug trials. For juvenile SSc, several patient-reported outcomes, including quality of life (QoL) measures, were voted on in 2017 and resupported in 2018. The patient-reported outcome measures to include were unanimously agreed (22 of 22 agreed) to be the Childhood Health Assessment Questionnaire (C-HAQ) (10,11) and the scleroderma-specific VAS, derived from the Scleroderma-HAQ disability index (12,13) (Table 1 and Supplementary Appendices A and B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25171, for detail of questions).

The C-HAQ is a standard patient-reported outcome, a child-directed assessment of function, modified from the multiply validated adult HAQ DI (10–15). The C-HAQ ascertains results over 8 functional domains and has been used in 2 large juvenile SSc cohorts (the International Inception and the CARRA Juvenile Systemic Sclerosis cohorts), with mean scores of 0.45 and 0.40 (range 0–3), respectively, (1,9). The C-HAQ, although it has floor effects, reflects the domains that are important to the function of the patient. Thus, it correlates with global well-being, health-related QoL, and organ systems of importance to patients with juvenile SSc (9). The group voted unanimously (22 of 22) in favor of including the C-HAQ for SSc patients age <16 years and the HAQ DI for patients age ≥16 years.

No formal MCID has been defined for the C-HAQ in juvenile SSc, but the Juvenile Systemic Sclerosis Inceptions and the CARRA network cohorts demonstrated that juvenile SSc patients improved by 24% and 44% over 1 and 2 years, respectively, and corresponded with improvements in PtGA and PhGA of disease activity (P = 0.02) (1,9,16). Voting at the consensus meeting then took place in regard to the MCID of C-HAQ in juvenile SSc and voters agreed to apply the MCID and cut points from juvenile idiopathic arthritis (JIA) and adult rheumatoid arthritis to the juvenile SSc cohorts (22 total: 20 agreed, 0 disagreed, 2 did not know). For reference, among 67 JIA patients followed longitudinally, those rated without change had a median C-HAQ change of 0, and for those rated as having change, the MCID was –0.188 for improvement and +0.125 for worsening (17).

The other main group of patient-reported outcome measures discussed were VAS scales from the Scleroderma-HAQ DI. In juvenile SSc, these VAS scales have been piloted in the National Registry of Childhood Onset Scleroderma (NRCOS) cohort (principal investigator Torok; n = 20), one of the few to have direct patient input (unpublished). All of the VAS measurements captured in the Scleroderma-HAQ DI were voted upon in their respective organ or general categories, with a unanimous vote (22 of 22) that they are important patient-reported outcome measures in a juvenile SSc trial (Table 1). The Scleroderma-HAQ DI includes scales for the following components: pain overall, gastrointestinal problems, breathing problems, Raynaud's severity, finger ulcer severity, and PtGA, which capture the SSc patient's perspective on the level of interference with normal activity in these domains over the past week (12). Modifications to the questionnaire for patients ages <8 years may say "your child" instead of "you."

Further discussion regarded the numerous other pediatric QoL instruments available and validated in other connective tissue diseases, particularly JIA. Available QoL instruments include: PedsQL (18), Peds Rheum QL (19), Family QL (20), Child Health Questionnaire (21), and Child Health Questionnaire-9D (22). These measures were not included in the current published juvenile SSc cohort, and so their performance characteristics are unknown in juvenile SSc. Although which QoL instrument is to be used is unknown, voters decided unanimously (22 of 22) that QoL (in addition to the C-HAQ and Scleroderma-HAQ DI VAS) is important to capture in juvenile SSc patients.

Domain 3: skin. The mRSS is a pivotal outcome measure for any therapeutic trial in diffuse adult and juvenile-onset SSc, assessing the degree of skin thickness over 17 body sites (0–3 per skin area, range 0–51). The mRSS is fully validated via Outcome Measures in Rheumatology filters in adult SSc (23). The mRSS has not been formally validated in juvenile SSc, but it has been widely adopted in clinical practice and larger observational cohort studies (1,24). The mRSS was unanimously agreed on at the 2017 and 2018 meetings (22 of 22 agreed) as the only item in the skin domain.

Since the mRSS was developed, studied, and validated in adult SSc patients, with typical average age of onset between 40 to 50 years, a few cutaneous variables to consider in the scoring approach in children were suggested (based on expert opinion). These pediatric rheumatology experts considered additional qualitative features, such as the texture of the skin (i.e., waxy, smooth, hard) compared to other areas in that region of the body, the appearance of the skin (i.e., shininess, yellow/waxy appearance), lack of hair, thin skin with visible veins, dyspigmentation, and atrophy (dermal or subcutis) (2). While the mRSS is to be used in juvenile SSc clinical trials, it needs further thorough examination in juvenile SSc in the future (16). As no MCID has been developed, voters decided to use the absolute mRSS and a statistical change as a measure of skin response in a juvenile SSc trial (22 of 22 agreed) (Table 1).

Domain 4: Raynaud's phenomenon. SSc-associated Raynaud's phenomenon (SSc-RP) is the most common disease-specific manifestation of SSc (25). SSc-RP was ranked by adult patients as having the highest impact on QoL and perception of illness severity (26). RP was recorded in 75% of the patients in the juvenile SSc inception cohort (27). In a clinical trial, RP should be measured in a standardized manner to assess whether a proposed new treatment is effective. Raynaud's outcomes are primarily patient reported, including frequency, severity, and duration, but may be confounded by pain and coping strategies (25,28).

In the 2017 juvenile SSc meeting, 24 of the 25 participants voted that RP should be assessed. After some discussion regarding the Raynaud's condition score, Raynaud's VAS from the Scleroderma-HAQ DI, and the physician's assessment of Raynaud's phenomenon, the Raynaud's VAS from the Scleroderma-HAQ DI was agreed on for a juvenile SSc trial (22 of 22) (Table 1). As no MCID was available, voters agreed (22 of 22) to use a statistically significant difference in the VAS across timepoints as a useful measure in juvenile SSc trials.

Domain 5: digital ulceration. SSc-related digital ulcers (DUs) are a frequent and disabling clinical complication of juvenile SSc, affecting approximately 50% of patients in the cohort of 150 patients (1). DUs occur most frequently on the fingers or toes and can be the consequence of endothelial damage, trauma, or calcinosis. DUs impair hand function and compromise patients' QoL (29). To measure the burden of finger/digital/skin ulcers, the DU clinical assessment score (DUCAS) was developed and validated in adult SSc patients (30). The DUCAS captures the number of DUs, new DUs, gangrene, surgery needed, infection, unscheduled hospitalization for DUs, and analgesics for DU pain (most in a yes/no fashion). The DUCAS plus the digital ulcer Scleroderma-HAQ DI VAS encompass the items suggested in a survey of the European Scleroderma Trials and Research group regarding the DU impact in SSc (31). Voters unanimously decided (22 of 22) to include the DUCAS score and the digital ulcer Scleroderma-HAQ DI VAS as an outcomes measure for digital ulcers in a juvenile SSc trial (Table 1).

Domain 6: musculoskeletal system. Musculoskeletal manifestations, including joint, muscle, and/or tendon involvement, occur in 75–82% of juvenile SSc patients, with 19% having documented inflammatory arthritis in prospective cohort studies (1,9,24). In 2017, several variables constituting musculoskeletal involvement were considered, including swollen joint count, limited joint range of motion, change in muscle strength assessed by the childhood myositis assessment scale or manual muscle testing, new occurrence of tendon friction rubs, and change in muscle enzyme levels (creatine kinase, aldolase). The group reached consensus on including the swollen joint count and not the other discussed variables.

The swollen joint count variable was voted to be included at the 2017 meeting and was discussed again at the 2018 meeting, though with an emphasis on the fact that measuring swollen joints alone captured only a portion of the musculoskeletal involvement in juvenile SSc (e.g., not capturing tenosynovitis, contracture), while also missing inflammation, because joint swelling is difficult to measure in SSc (32). Ultimately, voters decided by a unanimous vote (22 of 22) to collect musculoskeletal involvement in a juvenile SSc trial, as the number of joints that have either joint swelling or limitation in range of motion associated with joint pain or tenderness that is considered secondary to juvenile SSc, thus including tenosynovitis (Table 1). The joint count will be called the "active joint count," will be a total score, and will be very similar to the joint counts in JIA (thus not requiring special training).

Domain 7: cardiac involvement. Although in juvenile SSc cardiac involvement is relatively infrequently detected clinically (5–15%), it is one of the major causes of noninfectious mortality in juvenile SSc (24,33). A consensus meeting among European cardiologists and rheumatologists (34) indicated the need to examine for arrhythmias (electrocardiogram, Holter monitor), and to include an imaging measure to examine fibrosis (i.e., magnetic resonance imaging of the heart), plus patient response outcomes, and echocardiogram to define cardiac involvement in SSc.

Several cardiac variables were discussed in the context of the 2017 and 2018 juvenile SSc meeting, and there was 100% agreement (22 of 22) on the following parameters (Table 1): 1) a

measure of ejection fraction was appropriate as an inclusion measure; 2) new onset of left ventricular failure and/or new "clinically important arrhythmia (malignant/non-benign)" were appropriate measures defining lack of response in a juvenile SSc trial; 3) the development of pulmonary hypertension "by accepted criteria" is a sign of nonresponse; 4) the development of new carditis should be removed from consideration as not well defined; and 5) the N-terminal pro b-type natriuretic peptide (NT-proBNP), not validated in juvenile SSc, was to go into the research agenda. Participants noted that 2 of these consensus items are included as the step 1 CRISS criteria for adult SSc: new-onset left ventricular failure and new-onset pulmonary hypertension, though both are specified further in adult SSc with "≤45% ejection fraction requiring treatment" and "measured via right cardiac catheterization requiring treatment," respectively (5).

Domain 8: pulmonary involvement. Interstitial lung disease (ILD) occurs in approximately 50% of patients in juvenile SSc (1,35). It is a major reason for mortality in adult patients with SSc (5,36). Screening for ILD in adult and pediatric SSc patients traditionally includes a pulmonary function test (PFT) with FVC and single-breath diffusion capacity for carbon monoxide (DL_{CO}) (37,38). In children, assessment of FVC is fairly standardized from age 3 years, while DL_{CO} is more reliable starting at age 8 years (39,40). The combination of high-resolution computed tomography (HRCT; low radiation protocols) and PFTs are now used to both detect and follow ILD progression and regression in adults (37,41). In children, HRCT has been eschewed because there is concern regarding radiation, though this is now being reconsidered (35). The 6-minute walk test (6MWT) is a sensitive measure, with an MCID of 10 meters (42,43), and normal values for healthy children exist for comparison (44), although in adults with SSc the 6MWT is not responsive to treatment, as it is confounded by joint contracture, muscle weakness, and fatigue.

At the 2017 consensus group meeting, voters agreed that the core CRISS variables, including the change in FVC (5), were appropriate for juvenile SSc, and in the 2018 consensus meeting there was 100% consensus to include FVC and age-eligible DL_{CO} in juvenile SSc trials. The group decided to include the 6MWT assessment in the core set (18 of 21 agreed, 3 of 21 disagreed), measured as absolute meters using within-patient changes for statistical comparisons (Table 1). Because there remained concerns of an increased risk of malignancy after repeated HRCT of the lungs (45), the group unanimously rejected it as a required outcome measure in a juvenile SSc trial.

Domain 9: renal involvement. The course of renal involvement in juvenile SSc is usually benign, but a broad spectrum of renal manifestations exist in juvenile SSc, from mild proteinuria to acute renal failure. The most severe type is characterized by new-onset hypertension accompanied by acute kidney injury, proteinuria, hematuria or signs of microangiopathy (thrombocytopenia or hemolysis), or scleroderma renal crisis, which is a rare event in children (1) but it remains a major risk factor for mortality.

The consensus group agreed unanimously (21 of 21) to include the new occurrence of scleroderma renal crisis as an outcome measure criterion for a juvenile SSc trial (Table 1). This occurrence is also an adult CRISS step 1 criterion, which would consider the patient as not improved (5). The criterion was adjusted, accounting for the definition of high blood pressure in children and adolescents (46) and the Kidney Disease: Improving Global Outcomes definition of acute kidney injury (47). Other items related to renal involvement, namely new diagnosis of hypertension, new persistent proteinuria, and decrease of the glomerular filtration rate, were unanimously (21 of 21) rejected by the group as outcome measure criteria for juvenile SSc treatment trials, because those items lacked specificity and/or had a low prevalence in juvenile SSc patients.

Domain 10: gastrointestinal involvement. Gastrointestinal manifestations of SSc have been reported in 25–92% of children and are associated with poor QoL (9,48). Gastrointestinal manifestations in adult SSc patients range from mild oropharyngeal dysphagia to malnutrition (15–56%) and increased mortality (49). Malnutrition is a major concern in the growing child and has been shown to predict mortality in other pediatric chronic illnesses with gastrointestinal absorption issues, such as chronic kidney disease (50). Typical measures in children to assess malnutrition include midarm circumference and triceps skinfold thickness (51), but in juvenile SSc these measures may be confounded by skin manifestations. Another indicator of malnutrition in children, very low body mass index (BMI) (Z score ≤2), indicating moderate to severe malnutrition, can be used in juvenile SSc, with very low BMI documented in 14% of the juvenile SSc CARRA registry patients and correlating with poor QoL measures (9). Multiple other nonspecific laboratory tests (vitamins, pre-albumin level, etc.) may not be reliable in juvenile SSc. Voters unanimously agreed in both the 2017 and 2018 (21 of 21) consensus meetings to include the BMI as a single assessment for response regarding gastrointestinal involvement (Table 1).

Domain 11: biomarkers. No peripheral blood biomarker has been fully validated to the extent that it can be used to measure response in a juvenile SSc trial. Voters unanimously agreed (21 of 21) that it is appropriate to collect biosamples, when possible and available, though a particular serologic biomarkers was not targeted (Table 1).

Domain 12: growth and development. In growing children, normal growth and development is important. In the 2017 and 2018 consensus meetings, both a delay in sexual maturation and a decrease in growth velocity were considered as potential outcome measures for a juvenile SSc study; however, voters

unanimously rejected both as included outcome measures (Table 1). The consensus was that too many factors contribute to growth and development (e.g., sex, age, nutrition) to be reliable as measures of response to treatment in a juvenile SSc trial.

DISCUSSION

In JIA, guidance for measurements and clinical trials is available (52). The present effort is the first such guidance in juvenile SSc (Table 1 and Supplementary Appendices A and B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25171). We specifically aimed this proposal at a 12-month, open label juvenile SSc clinical treatment trial. It was not aimed at clinical practice or other trial designs (e.g., double-blind design), because this design is common in pediatric rheumatology and is a simple design to carry through.

This proposal has some significant strengths. It called together diverse medical specialties concerned with SSc as well as patients, and it built on knowledge of the literature (mostly adult SSc and JIA studies). Also, this proposal was developed over several years and included updated data from 3 juvenile SSc registries (Inceptions, CARRA, and NRCOS) as well as a review of the literature, thus supplying as much factual background as possible and on an ongoing basis. Of note, all the core variables for the composite validated adult SSc outcome measure, CRISS, were captured in our juvenile SSc international consensus (5). Our juvenile SSc consensus measures include 3 of the 4 nonresponse criteria of CRISS step 1 (only a decrease of FVC ≥15% was not included in juvenile SSc) (Table 1), and all components of CRISS step 2 (mRSS, FVC%, PhGA, and PtGA). One of the next steps of this group is considering the validation of CRISS in juvenile SSc.

There are also limitations. This proposal was oriented toward a 12-month, open label clinical trial, and additional considerations would be needed if one were to consider a single-blind or double-blind study design. Some measures were dependent on expert opinion alone (e.g., mRSS) and will need validation. Some novel tools in juvenile SSc, such as capillaroscopy and sonography, have only been used in juvenile SSc in observational setting and are a matter of future research.

The goal in the near future is to pilot these outcomes (see Supplementary Appendices A and B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25171) in the juvenile SSc cohorts, with particular focus on new or established patients starting medications, to evaluate the change of these outcomes in juvenile SSc. Both individual outcomes will be evaluated as well as a composite measure, with options to weigh measures.

ACKNOWLEDGMENTS

The authors would like to thank the following people for their contributions to this study: Nicola Helmus for serving as the main administrator of the meetings and one of the three scribes; Drs. Muthana Al-Obaidi, Eileen Baildem, Antonia Höger, Valentina Leone, and Brandi Stevens, all of whom are pediatric rheumatologists that attended and voted at either the 2017 or 2018 meeting; Dr. Amanda Saracino, an adult dermatologist and Dr. Richard Silver, an adult rheumatologist, who attended and voted at the 2017 meeting; Consultants Dr. Lusine Ambartsumyan, a pediatric gastroenterologist, and Dr. Christian Beyer, a pediatric cardiologist, both who were present and gave guidance during the 2018 meeting.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Foeldvari, Torok, Klotsche, Costa-Reis, Denton, Hinrichs, Li, Pilkington, Kasapcopur, Smith, Furst.

Acquisition of data. Foeldvari, Torok, Anton, Blakley, Constantin, Curran, Cutolo, Denton, Fligelstone, Ingegnoli, Li, Němcová, Orteu, Pilkington, Smith, Stevens, Klotsche, Khanna, Costa-Reis, Del Galdo, Hinrichs, Kasapcopur, Pain, Ruperto, Zheng, Furst.

Analysis and interpretation of data. Foeldvari, Torok, Klotsche, Costa-Reis, Del Galdo, Denton, Cutolo, Khanna, Ruperto, Smith, Furst.

REFERENCES

1. Foeldvari I, Klotsche J, Kasapcopur O, et al. Differences sustained between diffuse and limited forms of juvenile systemic sclerosis in an expanded international cohort. Arthritis Care Res (Hoboken) 2022; 74:1575–84.
2. Foeldvari I, Torok KS. Review for best practice in clinical rheumatology juvenile systemic sclerosis: updates and practice points [review]. Best Pract Res Clin Rheumatol 2021;35:101688.
3. Medsger TA Jr, Silman AJ, Steen VD, et al. A disease severity scale for systemic sclerosis: development and testing. J Rheumatol 1999;26: 2159–67.
4. La Torre F, Martini G, Russo R, et al. A preliminary disease severity score for juvenile systemic sclerosis. Arthritis Rheum 2012;64: 4143–50.
5. Khanna D, Berrocal VJ, Giannini EH, et al. The American College of Rheumatology provisional composite response index for clinical trials in early diffuse cutaneous systemic sclerosis. Arthritis Care Res (Hoboken) 2016;68:167–78.
6. Valentini G, Bencivelli W, Bombardieri S, et al. European Scleroderma Study Group to define disease activity criteria for systemic sclerosis. III. Assessment of the construct validity of the preliminary activity criteria. Ann Rheum Dis 2003;62:901–3.
7. Valentini G, Iudici M, Walker UA, et al. The European Scleroderma Trials and Research group (EUSTAR) task force for the development of revised activity criteria for systemic sclerosis: derivation and validation of a preliminarily revised EUSTAR activity index. Ann Rheum Dis 2017; 76:270–6.
8. Klotsche J, Foeldvari I, Kasapcopur O, et al. SAT0500 How the adult CRISS works in pediatric jSSc patients: results from the Juvenile Scleroderma Inception Cohort. Ann Rheum Dis 2020;79:1206.
9. Stevens BE, Torok KS, Li SC, et al. Clinical characteristics and factors associated with disability and impaired quality of life in children with juvenile systemic sclerosis: results from the Childhood Arthritis and Rheumatology Research Alliance Legacy Registry. Arthritis Care Res (Hoboken) 2018;70:1806–13.

10. Ramey DR, Raynauld JP, Fries JF. The health assessment questionnaire 1992: status and review. Arthritis Care Res (Hoboken) 1992;5:119–29.
11. Singh G, Athreya BH, Fries JF, et al. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 1994;37:1761–9.
12. Steen VD, Medsger TA Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. Arthritis Rheum 1997;40:1984–91.
13. Johnson SR, Hawker GA, Davis AM. The Health Assessment Questionnaire disability index and scleroderma health assessment questionnaire in scleroderma trials: an evaluation of their measurement properties. Arthritis Rheum 2005;53:256–62.
14. Ruperto N, Ravelli A, Pistorio A, et al. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries: review of the general methodology. Clin Exp Rheumatol 2001;19 Suppl 23:S1–9.
15. Poole JL, Williams CA, Bloch DA, et al. Concurrent validity of the Health Assessment Questionnaire Disability Index in scleroderma. Arthritis Care Res (Hoboken) 1995;8:189–93.
16. Foeldvari I, Klotsche J, Kasapcopur O, et al. After 24 months observation period the patients related outcomes improve significantly in the Juvenile Scleroderma Inceptions Cohorte [abstract]. Arthritis Rheumatol 2018;70 Suppl 9. URL: https://acrabstracts.org/abstract/after-24-months-observation-period-the-patients-related-outcomes-improve-significantly-in-the-juvenile-scleroderma-inceptions-cohorte/.
17. Brunner HI, Klein-Gitelman MS, Miller MJ, et al. Minimal clinically important differences of the childhood health assessment questionnaire. J Rheumatol 2005;32:150–61.
18. Varni JW, Burwinkle TM, Seid M. The PedsQL 4.0 as a school population health measure: feasibility, reliability, and validity. Qual Life Res 2006;15:203–15.
19. Varni JW, Seid M, Smith Knight T, et al. The PedsQL in pediatric rheumatology: reliability, validity, and responsiveness of the Pediatric Quality of Life Inventory Generic Core Scales and Rheumatology Module. Arthritis Rheum 2002;46:714–25.
20. Varni JW, Sherman SA, Burwinkle TM, et al. The PedsQL Family Impact Module: preliminary reliability and validity. Health Qual Life Outcomes 2004;2:55.
21. Landgraf JM, Maunsell E, Speechley KN, et al. Canadian-French, German and UK versions of the Child Health Questionnaire: methodology and preliminary item scaling results. Qual Life Res 1998;7:433–45.
22. Stevens KJ. Working with children to develop dimensions for a preference-based, generic, pediatric, health-related quality-of-life measure. Qual Health Res 2010;20:340–51.
23. Clements P, Lachenbruch P, Siebold J, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. J Rheumatol 1995;22:1281–5.
24. Scalapino K, Arkachaisri T, Lucas M, et al. Childhood onset systemic sclerosis: classification, clinical and serologic features, and survival in comparison with adult onset disease. J Rheumatol 2006;33:1004–13.
25. Pauling JD, Saketkoo LA, Matucci-Cerinic M, et al. The patient experience of Raynaud's phenomenon in systemic sclerosis [review]. Rheumatology (Oxford) 2019;58:18–26.
26. Frantz C, Avouac J, Distler O, et al. Impaired quality of life in systemic sclerosis and patient perception of the disease: a large international survey. Semin Arthritis Rheum 2016;46:115–23.
27. Foeldvari I, Klotsche J, Torok KS, et al. Characteristics of the first 80 patients at timepoint of first assessment included in the Juvenile Systemic Sclerosis Inception Cohort. J Scleroderma Relat Disord 2018;4:1–13.
28. Gladue H, Maranian P, Paulus HE, et al. Evaluation of test characteristics for outcome measures used in Raynaud's phenomenon clinical trials. Arthritis Care Res (Hoboken) 2013;65:630–6.
29. Hughes M, Pauling JD. Exploring the patient experience of digital ulcers in systemic sclerosis [review]. Semin Arthritis Rheum 2019;48:888–94.
30. Bruni C, Ngcozana T, Braschi F, et al. Preliminary validation of the Digital Ulcer Clinical Assessment Score in systemic sclerosis. J Rheumatol 2019;46:603–8.
31. Blagojevic J, Bellando-Randone S, Abignano G, et al. Classification, categorization and essential items for digital ulcer evaluation in systemic sclerosis: a DeSScipher/European Scleroderma Trials and Research group (EUSTAR) survey. Arthritis Res Ther 2019;21:35.
32. Low AH, Lax M, Johnson SR, et al. Magnetic resonance imaging of the hand in systemic sclerosis. J Rheumatol 2009;36:961–4.
33. Martini G, Vittadello F, Kasapcopur O, et al. Factors affecting survival in juvenile systemic sclerosis. Rheumatology (Oxford) 2009;48:119–22.
34. Bruni C, Buch MH, Furst DE, et al. Primary systemic sclerosis heart involvement: a systematic literature review and preliminary data-driven, consensus-based WSF/HFA definition. J Scleroderma Relat Disord 2022;7:24–32.
35. Foeldvari I, Klotsche J, Hinrichs B, et al. Underdetection of interstitial lung disease in juvenile systemic sclerosis. Arthritis Care Res (Hoboken) 2022;74:364–70.
36. Cottin V, Brown KK. Interstitial lung disease associated with systemic sclerosis (SSc-ILD) [review]. Respir Res 2019;20:13.
37. Volkmann E. Natural history of systemic sclerosis-related interstitial lung disease: how to identify a progressive fibrosing phenotype. J Scleroderma Relat Disord 2020;5 Suppl:31–40.
38. Showalter K, Hoffmann A, Rouleau G, et al. Performance of forced vital capacity and lung diffusion cutpoints for associated radiographic interstitial lung disease in systemic sclerosis. J Rheumatol 2018;45:1572–6.
39. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. Eur Respir J 2012;40:1324–43.
40. Beydon N, Davis SD, Lombardi E, et al. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children [review]. Am J Respir Crit Care Med 2007;175:1304–45.
41. Silver KC, Silver RM. Management of systemic-sclerosis-associated interstitial lung disease [review]. Rheum Dis Clin North Am 2015;41:439–57.
42. Pugnet G, Marjanovic Z, Deligny C, et al. Reproducibility and utility of the 6-minute walk test in systemic sclerosis. J Rheumatol 2018;45:1273–80.
43. Vandecasteele E, Melsens K, De Keyser F, et al. A prospective, longitudinal study evaluating the baseline six-minute walk test as an individual reference value in systemic sclerosis patients. Clin Exp Rheumatol 2018;36 Suppl 113:95–101.
44. Lammers AE, Hislop AA, Flynn Y, et al. The 6-minute walk test: normal values for children of 4–11 years of age. Arch Dis Child 2008;93:464–8.
45. Miglioretti DL, Johnson E, Williams A, et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. JAMA Pediatr 2013;167:700–7.
46. Flynn JT, Kaelber DC, Baker-Smith CM, et al. Clinical practice guideline for screening and management of high blood pressure in children and adolescents. Pediatrics 2017;140:e20171904.

47. Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. Nephron Clin Pract 2012;120:c179–84.

48. Martini G, Foeldvari I, Russo R, et al. Systemic sclerosis in childhood: clinical and immunologic features of 153 patients in an international database. Arthritis Rheum 2006;54:3971–8.

49. Gyger G, Baron M. Systemic sclerosis: gastrointestinal disease and its management [review]. Rheum Dis Clin North Am 2015;41:459–73.

50. Tu J, Cheung WW, Mak RH. Inflammation and nutrition in children with chronic kidney disease [review]. World J Nephrol 2016;5:274–82.

51. Duggan C, Watkins J, Walker A, et al. Clinical assessment of nutritional status. In: Nutrition in pediatrics: basic science, clinical applications. BC Decker; 2008. p. 5–13.

52. Lovell DJ, Ruperto N, Giannini EH, et al. Advances from clinical trials in juvenile idiopathic arthritis [review]. Nat Rev Rheumatol 2013;9: 557–63.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2463–2471
DOI 10.1002/acr.25167
© 2023 The Authors. *Arthritis Care & Research* published by Wiley Periodicals LLC on behalf of American College of Rheumatology.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

American College of Rheumatology
Empowering Rheumatology Professionals

Serum Urate Monitoring Among Older Adults With Gout: Initiating Urate-Lowering Therapy in Ontario, Canada

Timothy S. H. Kwok,[1] Bindee Kuriya,[1] Gillian Hawker,[2] Ping Li,[3] Gregory Choy,[4] and Jessica Widdifield[5]

Objective. To assess the proportion of, and factors associated with, older adults with gout receiving a serum urate (SUA) test after starting urate-lowering therapy (ULT).

Methods. We performed a population-based retrospective cohort study in Ontario, Canada in patients ages ≥66 years with gout, newly dispensed ULT between 2010 and 2019. We characterized patients with SUA testing within 6 and 12 months after ULT dispensation. Multilevel logistic regression clustered by ULT prescriber evaluated the factors associated with SUA monitoring within 6 months.

Results. We included 44,438 patients with a mean ± SD age of 76.0 ± 7.3 years and 64.4% male. Family physicians prescribed 79.1% of all ULTs. SUA testing was lowest in 2010 (56.4% at 6 months) and rose over time to 71.3% in 2019 ($P < 0.0001$). Compared with rheumatologists, family physicians (odds ratio [OR] 0.26 [95% confidence interval (95% CI) 0.23–0.29]), internists (OR 0.34 [95% CI 0.29–0.39]), nephrologists (OR 0.37 [95% CI 0.30–0.45]), and other specialties (OR 0.25 [95% CI 0.21–0.29]) were less likely to test SUA, as were male physicians (OR 0.87 [95% CI 0.83–0.91]). Patient factors associated with lower odds of SUA monitoring included rural residence (OR 0.81 [95% CI 0.77–0.86]), lower socioeconomic status (OR 0.91 [95% CI 0.85–0.97]), and patient comorbidities. Chronic kidney disease, hypertension, diabetes mellitus, and coprescription of colchicine/oral corticosteroids (OR 1.31 [95% CI 1.23–1.40]) were correlated with increased SUA testing.

Conclusion. SUA testing is suboptimal among older adults with gout initiating ULT but is improving over time. ULT prescriber, patient, and prescription characteristics were correlated with SUA testing.

INTRODUCTION

A treatment principle of gout involves the lowering of serum urate (SUA) below crystallization thresholds to avoid sequelae of hyperuricemia such as tophi, nephrolithiasis, and inflammatory arthritis (1,2). This principle has subsequently been endorsed in current rheumatology clinical practice guidelines, promoting a treat-to-target urate level strategy with urate-lowering therapy (ULT) (3,4). A key step in guiding the titration of ULT to target levels is the monitoring of SUA. This monitoring is critical in gout management, as multiple dose escalations may be needed to reach target levels (5). Consequently, SUA monitoring has been endorsed in the American College of Rheumatology (ACR) Electronic Clinical Quality Measures for Gout, with the recommendation that SUA should be measured, at minimum, within 6 months after ULT initiation (6).

To date, there have been few studies evaluating adherence to SUA monitoring benchmarks, especially in the older adult population. This lack of evidence presents an important knowledge gap, as the burden of gout increases with age (7). Studies thus far have uniformly shown SUA monitoring to occur infrequently after starting ULT, ranging from 17% to 45% within 6 months (8,9). However, past analyses have not studied temporal trends and did not analyze associated patient- and physician-level factors of SUA monitoring, nor did they focus on the older adult population. Our aim was to fill this knowledge gap and describe the proportion of older adults with gout undergoing SUA testing

The opinions, results, and conclusions reported in this article are those of the authors and are independent of the data sources; no endorsement is intended or should be inferred.

Supported by the Sunnybrook Foundation and by ICES (formerly known as the Institute for Clinical Evaluative Sciences), which is funded by the Ontario Ministry of Health and the Ministry of Long-Term Care. Dr. Kwok's work was supported by a Canada Graduate Scholarship from the Canadian Institutes of Health Research. Dr. Widdifield's work was supported by the Arthritis Society Stars Career Development Award (STAR-19-0610).

[1]Timothy S. H. Kwok, MD, MSc, FRCPC, Bindee Kuriya, MD, SM, FRCPC: University of Toronto, Toronto, Ontario, Canada; [2]Gillian Hawker, MD, MSc, FRCPC: University of Toronto and Women's College Hospital, Toronto, Ontario, Canada; [3]Ping Li, PhD: ICES, Toronto, Ontario, Canada; [4]Gregory Choy, MD, FRCPC: University of Toronto and Sunnybrook Research Institute, Toronto, Ontario, Canada; [5]Jessica Widdifield, PhD: University of Toronto, ICES, and Sunnybrook Research Institute, Toronto, Ontario, Canada.

Author disclosures are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25167.

Address correspondence via email to Timothy S. H. Kwok, MD, MSc, FRCPC, at timothysh.kwok@mail.utoronto.ca.

Submitted for publication February 1, 2023; accepted in revised form May 25, 2023.

SIGNIFICANCE & INNOVATIONS

- Serum urate (SUA) monitoring is variable and suboptimal (according to most international recommendations) but is improving over time in older adult patients with gout.
- Rheumatologists and family physicians have the highest and lowest percentages, respectively, of their patients having SUA monitoring after starting urate-lowering therapy (ULT).
- Chronic kidney disease, hypertension, diabetes mellitus, ULT prescribed by rheumatologists at lower starting doses, in more recent years, and colchicine/oral corticosteroid prophylaxis were associated with higher odds of having SUA monitoring.
- Increasing patient age, rural residence, low socioeconomic status, male ULT prescriber, chronic obstructive pulmonary disease, and prior cardiovascular events were associated with lower odds of having SUA monitoring.

within 6 months after ULT initiation. We also analyzed patient- and physician-level factors associated with SUA testing.

PATIENTS AND METHODS

Setting and study design. This study was conducted in Ontario, Canada's most populous province. Ontario residents of all ages are insured under the Ontario Health Insurance Plan (OHIP), a single-payer health care system that covers medically necessary services and procedures (such as laboratory tests). These contacts for health services are recorded in administrative databases that enable comprehensive evaluations of health services. Within Ontario, adults ages 65 years and over also automatically qualify for the Ontario Drug Benefit program for prescription medications.

We analyzed a retrospective population-based cohort of patients with gout age ≥66 years who were dispensed ULT between January 1, 2010, and March 31, 2019, sampled from health administrative data. The primary outcome was the presence of SUA testing within 6 and 12 months of ULT initiation.

This study was approved by a privacy impact assessment at ICES (formerly called the Institute for Clinical Evaluative Sciences). The use of the data was authorized under section 45 of Ontario's Personal Health Information Protection Act, which does not require review by a research ethics board.

Data sources and definitions. Data sources for this study included the province-wide Registered Persons Database (to ascertain patient-level demographic information), the OHIP claims database (to ascertain diagnosis codes during physician encounters and SUA tests performed), the Ontario Drug Benefit pharmacy claims database (to ascertain ULT dispensations), the Canadian Institute for Health Information Discharge Abstract Database (to ascertain inpatient diagnosis codes), the National Ambulatory Care Reporting System (to ascertain diagnosis codes during emergency department [ED] encounters), the Ontario Laboratories Information System (to ascertain SUA values), and the ICES Physician Database to identify specialty type and relevant physician characteristics of ULT prescribers.

Patient sampling commenced with the requirement of individuals age ≥66 years having a ULT dispensed (a prescription that was filled at a pharmacy) between January 1, 2010, and March 31, 2019, and who had a diagnosis code for gout (International Classification of Diseases [ICD], Ninth Revision 274 or 712, and ICD-10 M10). The date of the first ULT prescription served as the index date.

Patients were excluded if they were not permanent residents of Ontario, if they had missing data on age, sex, or location of residence, or had prior ULT claims. Patients with ULT dispensations in the 1-year period prior to age 66 years or January 1, 2010 were thus excluded, to identify new ULT starts. To ensure that the cohort was taking ULT for a diagnosis of gout, individuals without at least 1 gout diagnosis code were excluded. To further exclude patients taking ULT for conditions other than gout, we excluded patients with certain types of acute hematologic malignancy (i.e., all hematologic malignancies except for multiple myeloma and indolent lymphomas), tumor lysis syndrome (TLS) or end-stage renal disease (ESRD) with a look-back window of 2 years before the index date. Multiple myeloma and indolent lymphomas did not form a part of our exclusion criteria, as they are not strongly associated with TLS (10). Patients with an acute hematologic malignancy associated with TLS or TLS itself were excluded, as these patients may have an alternate indication for ULT other than gout, in the prevention or treatment of TLS (10). Patients with ESRD were excluded as allopurinol may have been used in preventing renal disease progression, as guided by historical evidence (11). Patients with ULT dispensed and a gout diagnosis code and who also had chronic kidney disease (CKD) were not excluded, given CKD's strong associations with hyperuricemia (12). The diagnosis codes used to ascertain gout, hematologic malignancies, TLS, and ESRD are detailed in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25167.

Covariates were measured on the index date of ULT dispensation. Patient-level variables included age in years, sex, income quintile, rural versus urban residence, calendar year of ULT prescription, coprescription of colchicine/oral corticosteroids at the time of or within 2 weeks of ULT prescription, ULT type/dose, and comorbidities, including CKD, hypertension, prior cardiovascular events, diabetes mellitus, chronic obstructive pulmonary disease (COPD), nephrolithiasis, or an incident cancer diagnosis in the past 2 years. Patient-level health care usage factors included ED visits for gout within 2 years prior to the index date, and whether patients had a designated primary care provider

(PCP). Physician-level variables included age, sex, volume of gout visits, and specialty of the ULT prescriber. Gout practice volume was defined by the total patient visits per year for gout during a 12-month observation window and was described using quintiles. All covariates including comorbid conditions were identified using hospital inpatient and outpatient diagnosis codes in the 2 years prior to cohort entry, and where possible were derived from previously validated case definitions, with details in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25167.

Statistical analysis. Descriptive statistics determined the proportion of patients with at least 1 SUA test performed within 6 months of starting ULT. The characteristics of the patient populations with and without SUA tests were compared using standardized differences (computed as the difference in means or proportions divided by the SE) with >0.1 (10%) considered a significant effect size (13,14). We also examined the distribution of the first SUA test (after index ULT dispensation) by month. Annual trends in the proportion of patients with SUA monitoring by 6 and 12 months after starting ULT were also determined, overall and stratified by prescriber specialty. The Cochrane-Armitage trend test was used to evaluate trend significance. Multilevel logistic regression clustered by ULT prescriber evaluated factors associated with optimal SUA monitoring by 6 months. This logistic regression model was done using generalized estimating equations with an independent correlation structure, as we were interested in analyzing the first SUA performed. Patient-level variables included demographic information, including age in years, sex, income quintile (discerned based on the patients' postal code and census neighborhood income, with the fifth quintile representing the highest income), rural versus urban residence (defined using the Rurality Index for Ontario), calendar year of ULT prescription, coprescription of colchicine/oral corticosteroids, ULT type/dose, and comorbidities. Patient-level health care usage factors included ED visits for gout within 2 years prior to the index date, and whether patients had a designated PCP (defined as not rostered, rostered, or virtually rostered). Physician-level variables included age, sex, and practice characteristics, including the volume of gout visits and specialty of the ULT prescriber. Individuals with missing ULT prescriber physician data were excluded from the multivariable analyses.

Recognizing the potential low sensitivity of gout diagnosis codes accompanying individuals with a ULT claim, we performed a sensitivity analysis to see whether a secondary cohort who met entry criteria but did not have gout diagnosis codes was similar to our study population. To do so, we created a secondary cohort in which a gout diagnosis code within 2 years prior to ULT initiation was not required for cohort entry. We compared our primary cohort with gout diagnosis codes to the secondary cohort without gout diagnosis codes, in terms of differences in patient and prescriber characteristics as well as the proportions of patients undergoing SUA testing. The data sets were linked using unique encoded identifiers and analyzed at ICES using SAS, version 9.4.

RESULTS

After pertinent exclusion criteria were applied, the primary cohort comprised 44,438 patients (Figure 1). Overall, 28,473 patients (64.1%) with gout had at least 1 SUA test within 6 months of starting ULT. Patients who had SUA testing within 6 months after ULT initiation (versus those who did not) were more likely to have been prescribed ULT in recent years within the study time period, to reside in urban areas, to have been prescribed colchicine or oral corticosteroids at the time of or within 2 weeks of ULT prescription, to be prescribed lower doses of allopurinol, to have CKD or diabetes mellitus, and to have had their ULT prescribed by rheumatologists as opposed to family physicians (Table 1).

The overall percentage of patients with SUA testing within 6 and 12 months of receiving ULT was 64.1% and 75.2%,

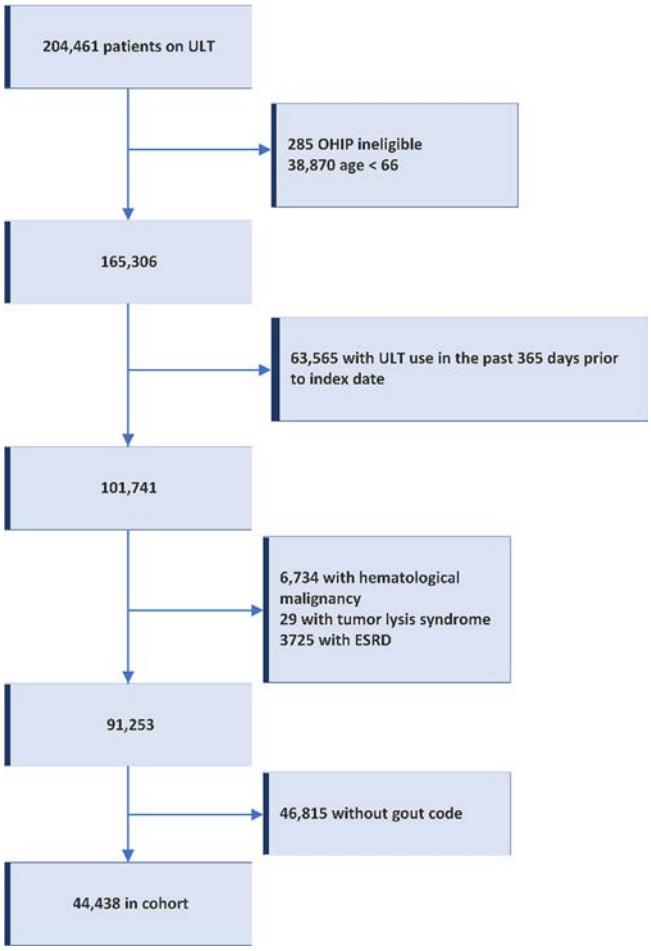


Figure 1. Flow diagram for cohort creation. ESRD = end-stage renal disease; OHIP = Ontario Health Insurance Plan; ULT = urate-lowering therapy. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25167/abstract.

Table 1. Comparison of baseline patient characteristics for those with and without SUA testing by 6 months after ULT initiation*

Characteristics	With SUA testing (n = 28,473)	Without SUA testing (n = 15,965)	Standardized difference
Patient characteristics			
Age, mean ± SD years	76.0 ± 7.2	76.1 ± 7.5	0
Male	18,198 (63.9)	10,435 (65.4)	0.03
Female	10,275 (36.1)	5,530 (34.6)	–
Income quintile			
1 (lowest)	5,512 (19.4)	3,202 (20.1)	0.02
2	6,103 (21.4)	3,444 (21.6)	0
3	5,667 (19.9)	3,396 (21.3)	0.03
4	5,701 (20.0)	3,066 (19.2)	0.02
5 (highest)	5,421 (19.0)	2,806 (17.6)	0.04
Rural residence	3,703 (13.0)	2,610 (16.3)	0.10
Chronic kidney disease	6,104 (21.4)	2,304 (14.4)	0.18
Hypertension	25,469 (89.4)	13,995 (87.7)	0.06
Diabetes mellitus	11,082 (38.9)	5,441 (34.1)	0.10
Chronic obstructive pulmonary disease	7,636 (26.8)	4,431 (27.8)	0.02
Prior cardiovascular event	2,917 (10.2)	1,701 (10.7)	0.01
Nephrolithiasis	615 (2.2)	311 (1.9)	0.01
Cancer	905 (3.2)	511 (3.2)	0
Prescription characteristics			
Calendar year of ULT prescription			
2010	3,020 (10.6)	2,332 (14.6)	0.12
2011	2,975 (10.4)	2,037 (12.8)	0.07
2012	2,812 (9.9)	1,887 (11.8)	0.06
2013	3,084 (10.8)	1,785 (11.2)	0.01
2014	2,874 (10.1)	1,514 (9.5)	0.02
2015	3,112 (10.9)	1,627 (10.2)	0.02
2016	3,385 (11.9)	1,639 (10.3)	0.05
2017	3,246 (11.4)	1,486 (9.3)	0.07
2018	3,309 (11.6)	1,394 (8.7)	0.10
2019†	656 (2.3)	264 (1.7)	0.05
Coprescription of colchicine/oral corticosteroids within 2 weeks of ULT dispensation	3,792 (13.3)	1,514 (9.5)	0.12
ULT type			
Allopurinol	28,364 (99.6)	15,902 (99.6)	0
Febuxostat	65 (0.2)	23 (0.1)	0.02
Probenecid	44 (0.2)	40 (0.3)	0.02
ULT dose			
Allopurinol: ≤50 mg/day	879 (3.1)	220 (1.4)	0.12
Allopurinol: 51–100 mg/day	12,956 (45.7)	6,128 (38.5)	0.14
Allopurinol: >100 mg/day	14,529 (51.2)	9,554 (60.1)	0.18
Febuxostat: ≤40 mg/day	7 (10.8)	‡	–
Febuxostat: >40 mg/day	58 (89.2)	‡	–
Health services characteristics			
Emergency department visits for gout within 2 years prior to index date			
0	23,948 (84.1)	13,414 (84.0)	0
1	3,763 (13.2)	2,082 (13.0)	0.01
≥2	762 (2.7)	469 (2.9)	0.02
Presence of primary care provider§			
Not rostered	208 (0.7)	124 (0.8)	0.01
Rostered	24,883 (87.4)	13,453 (84.3)	0.09
Virtually rostered	3,382 (11.9)	2,388 (15.0)	0.09
ULT prescriber characteristics			
Specialty, no. (% of total)			
Internal medicine	1,635 (5.7)	707 (4.4)	0.06
Nephrology	789 (2.8)	244 (1.5)	0.09
Family medicine	21,640 (76.0)	13,506 (84.6)	0.22
Rheumatology	2,441 (8.6)	346 (2.2)	0.29
Other specialty	1,012 (3.6)	627 (3.9)	0.02
Unknown	956 (3.4)	535 (3.4)	0

* Values are the number (%) unless indicated otherwise. Percentages may not add up to 100% due to missing data related to postal codes (income quintile). SUA = serum urate; ULT = urate-lowering therapy.
† Not a full calendar year due to the observation period ending on March 31, 2019, as per study design.
‡ Exact value cannot be provided due to adjacent small cell.
§ Having a primary care provider defined as being rostered (officially enrolled) or virtual rostered (repeatedly seeing same physician) at any time within 2 years prior to the index date.

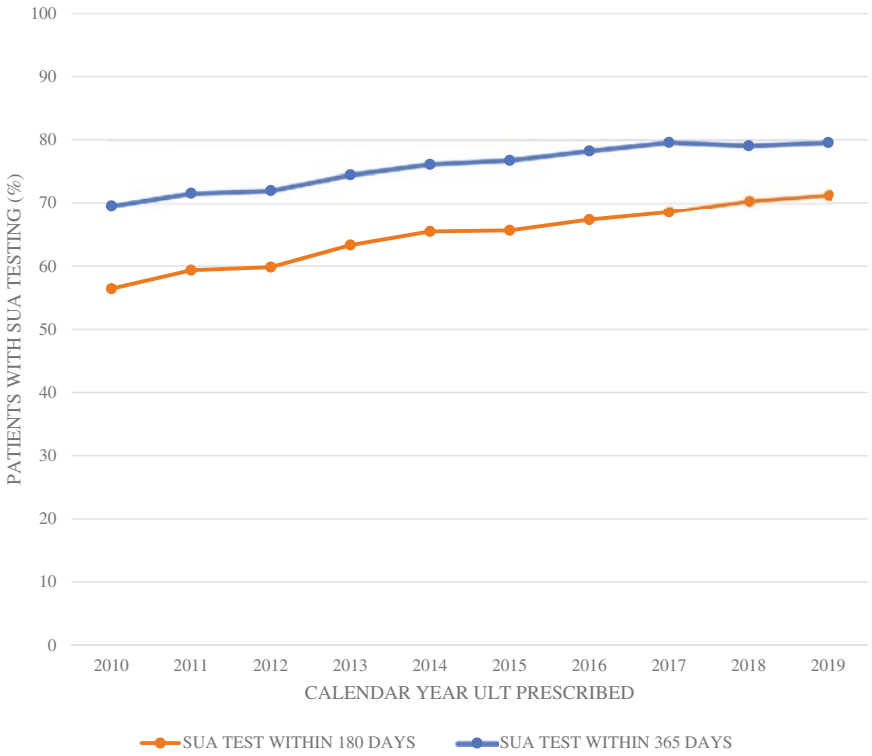


Figure 2. Trends in the annual percentage of patients with serum urate (SUA) tests performed within 6 months (180 days) and 12 months (365 days) of urate-lowering therapy (ULT) initiation between 2010 and 2019.

respectively, and increased significantly over the study period ($P < 0.0001$), from 56.4% and 69.5%, respectively, in 2010, to 71.3% and 79.6% in 2019 (Figure 2). SUA testing after ULT prescription increased over time for all physician specialties (Figure 3). SUA testing was consistently highest among rheumatologists, followed by nephrologists and internists. Family medicine and other specialty prescribers (i.e., not family medicine, internal medicine, nephrology, or rheumatology) were less likely to perform SUA

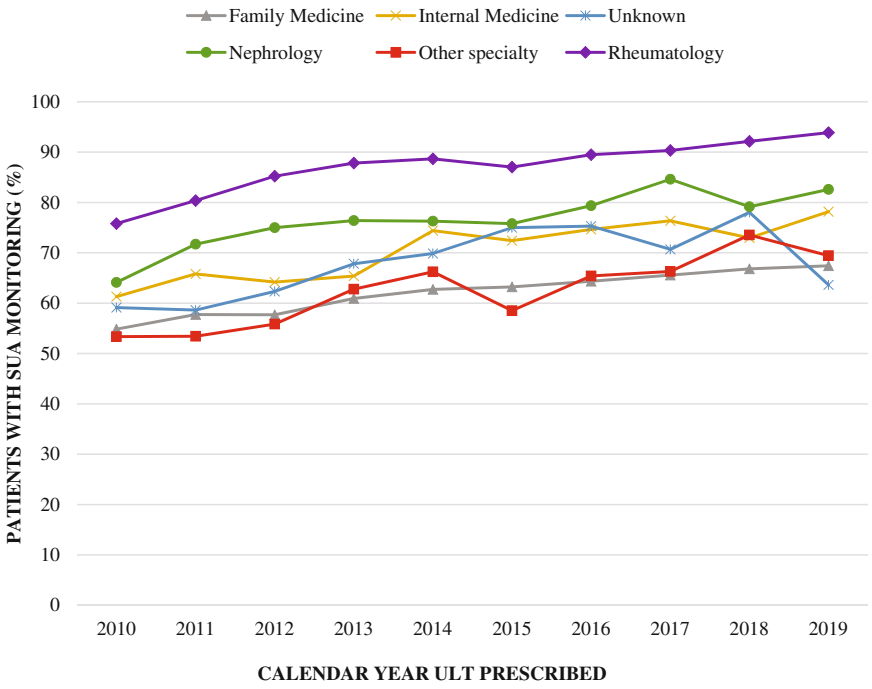


Figure 3. Percentage of gout patients with serum urate (SUA) testing within 6 months after index urate-lowering therapy (ULT) prescription, stratified by prescriber primary specialty.

testing within 6 months (only 54.9%). The percentage of patients with SUA tests within 6 months modestly improved to 67.4% in 2019 for family medicine. This finding drastically contrasts to rheumatologists prescribing ULT, as 75.8% of patients had SUA testing within 6 months in 2010, increasing to 93.9% in 2019. Within the 6- and 12-month observation window, 28,437 and 33,439 patients, respectively, had SUA tests performed. Of these, 12,314 patients (36.8%) had tests performed within the first month and 28,473 (85.1%) within the first 6 months.

In hierarchical logistic regression analyses (Table 2), increasing patient age (OR 0.99 [95% CI 0.99–1.00]), residing in rural areas (OR 0.81 [95% CI 0.77–0.86]), and being in the lowest income quintile (OR 0.91 [95% CI 0.85–0.97]) were associated with lower odds of having SUA testing. Moreover, patients with concomitant CKD (OR 1.40 [95% CI 1.32–1.49]), hypertension (OR 1.11 [95% CI 1.04–1.18]), and diabetes mellitus (OR 1.17 [95% CI 1.12–1.22]) had higher odds, while patients with COPD (OR 0.94 [95% CI 0.90–0.99]) and prior cardiovascular events (OR 0.85 [95% CI 0.79–0.91]) had lower odds of SUA testing. From a ULT prescriber perspective, patients of male physicians were less likely to have SUA testing (OR 0.87 [95% CI 0.83–0.91]). Patients with ULT prescriptions from family physicians (OR 0.26 [95% CI 0.23–0.29]), internists (OR 0.34 [95% CI 0.29–0.39]), nephrologists (OR 0.37 [95% CI 0.30–0.45]), and all other specialties (OR 0.25 [95% CI 0.21–0.29]) were less likely to have SUA testing compared with rheumatologists. ULT prescription characteristics also played a role in SUA testing, with allopurinol starting doses of 51–100 mg (OR 0.72 [95% CI 0.61–0.84]), >100 mg (OR 0.55 [95% CI 0.47–0.65]), febuxostat (OR 0.57 [95% CI 0.34–0.97]) and probenecid (OR 0.43 [95% CI 0.27–0.70]) all being associated with a lower OR compared to a starting dose of allopurinol of 50 mg. Patients with ULT prescribed in 2015–2019 (OR 1.26 [95% CI 1.21–1.31]) compared to 2010–2014 and the coprescription of colchicine/oral corticosteroids (OR 1.31 [95% CI 1.23–1.40]) were both associated with a higher likelihood of SUA testing. Sensitivity analyses demonstrated similar results (see Supplementary Table 2 and Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25167).

DISCUSSION

We evaluated SUA testing after starting ULT in older adults with gout in Ontario, Canada. Among 44,438 gout patients who initiated ULT between 2010 and 2019, 64% of patients had SUA testing within 6 months of ULT initiation. SUA testing improved over time for all prescribers at all time points, while rheumatologists and family physicians had the highest and lowest rates of testing, respectively.

Our rates of SUA testing are higher than those of others in the literature. In particular, a prior retrospective claims analysis found

Table 2. Logistic regression analysis for factors and their association with SUA testing within 6 months after starting ULT in older adult patients with gout*

Characteristic	Adjusted OR (95% CI)	*P*
ULT prescribing physician-level factors		
Age <45 years (ref. ≥65 years)	1.02 (0.96–1.09)	0.504
Age 45–64 years (ref. ≥65 years)	1.02 (0.97–1.08)	0.411
Male (ref. female)	0.87 (0.83–0.91)	<0.001
Gout practice volume (ref. 4, top quartile)		
1	1.05 (0.98–1.12)	0.163
2	1.06 (0.99–1.13)	0.055
3	1.06 (0.99–1.12)	0.072
Specialty (ref. rheumatology)		
Family medicine	0.26 (0.23–0.29)	<0.001
Internal medicine	0.34 (0.29–0.39)	<0.001
Nephrology	0.37 (0.30–0.45)	<0.001
Other specialty	0.25 (0.21–0.29)	<0.001
Patient-level factors		
Age (by 1 year of age)	0.99 (0.99–1.00)	<0.001
Male (ref. female)	0.99 (0.95–1.04)	0.822
Income quintile (ref. 5, highest)		
1 (lowest)	0.91 (0.85–0.97)	0.006
2	0.93 (0.88–1.00)	0.037
3	0.88 (0.82–0.94)	<0.001
4	0.99 (0.93–1.06)	0.790
Rural residence (ref. urban residence)	0.81 (0.77–0.86)	<0.001
Comorbidities (ref. absence of specific comorbidity)		
Chronic kidney disease	1.40 (1.32–1.49)	<0.001
Hypertension	1.11 (1.04–1.18)	0.002
Diabetes mellitus	1.17 (1.12–1.22)	<0.001
Chronic obstructive pulmonary disease	0.94 (0.90–0.99)	0.009
Prior cardiovascular event	0.85 (0.79–0.91)	<0.001
Nephrolithiasis	1.06 (0.92–1.22)	0.434
Cancer	0.94 (0.83–1.05)	0.250
Prescription-level factors		
Calendar year of ULT prescription 2015–2019 (ref. 2010–2014)	1.26 (1.21–1.31)	<0.001
Coprescription of colchicine/oral corticosteroids within 2 weeks of ULT prescription (ref. no coprescription)	1.31 (1.23–1.40)	<0.001
ULT dose (ref. allopurinol ≤50 mg daily)		
Allopurinol 51–100 mg	0.72 (0.61–0.84)	<0.001
Allopurinol >100 mg	0.55 (0.47–0.65)	<0.001
Febuxostat	0.57 (0.34–0.97)	0.039
Probenecid	0.43 (0.27–0.70)	0.001
Health care usage factors		
ED visits for gout within 2 years prior to index date (ref. none)		
1	1.00 (0.94–1.07)	0.937
≥2	0.91 (0.81–1.04)	0.161
Presence of primary care provider (ref. absence)		
Rostered	1.18 (0.93–1.49)	0.175
Virtually rostered	0.93 (0.73–1.18)	0.555

* 95% CI = 95% confidence interval; ED = emergency department; OR = odds ratio; ref. = reference; SUA = serum urate; ULT = urate-lowering therapy.

that 17% of patients with gout on allopurinol had SUA measured within 6 months of ULT initiation (8). This finding was contrasted to a study of American veterans, where 62% of patients had SUA tested within 6 months (9). In addition, an electronic medical record–based study found that only 45% patients had SUA testing within 6 months of starting allopurinol (15). These findings are not just isolated to the US, as prior studies in New Zealand have also demonstrated similar trends in gout care (16). Hence, our data suggest a practice variation in SUA testing after starting ULT as measured against current quality standards (6).

We identified a significant improvement over time in SUA testing within 6 months of ULT dispensation, at 56% in 2010 and increasing over time to 71% by 2019. Additionally, we found that a large proportion of SUA testing performed after index ULT prescription was within the first month, at 36.8% of patients within a 12 month period. These data suggest that for the subset of prescribers who are indeed monitoring SUA after ULT prescription for their patients with gout, monitoring begins very early. More importantly, whether this close monitoring translates to dose adjustments with regard to ULT dose titration is unknown. The frequency of ongoing monitoring, factors related to ongoing monitoring, and whether waning in monitoring occurs over time warrant further research. To our knowledge, this is the first study reporting temporal trends in SUA monitoring in North America. Whether or not improvements over time are reflective of sequentially published clinical practice guidelines is unknown. A recent study from the UK assessed the influence of gout clinical practice guidelines on treat-to-target benchmarks and found no effect, showing that secular trends in improvements began prior to guideline dissemination (17).

The first references to a treat-to-target paradigm recommending the monitoring of SUA levels were the 2006 EULAR guidelines and subsequently the 2007 British Society for Rheumatology (BSR) guidelines (18,19). The subsequent chronology of guidelines that occurred during our study observation period included the 2012 ACR guidelines, which was the first set of guidelines for gout developed by a North American rheumatology society, and subsequently the 2017 American College of Physicians (ACP) guidelines (20,21). Subsequent guidelines were published during the study period and were developed in Europe, thus leading to a questionable effect on our study population in Canada. Since multiple international guidelines were published during our study period, assessing the temporal influence of any specific one is difficult.

We found that the rate of postprescription testing was significantly higher for specialists than for PCPs. There has been a paucity of prior studies assessing the effect of physician specialty on SUA monitoring. A prior study found that patients of specialist ULT prescribers were more likely to undergo SUA monitoring compared with patients receiving care from generalist practitioners (68% versus 39%) (9). Gout management within primary care has been previously studied in a US survey of physicians, which demonstrated that SUA monitoring is suboptimal according to international gout guidelines (22). There are many potential explanations. Rheumatologists have the most advanced inflammatory musculoskeletal disorder medical training in comparison to the other specialties and may also be more likely to remain involved in the ongoing care of gout patients as the principal care physician overseeing the patient's gout management, whereas other specialties (such as nephrologists) may be prioritizing the management of the comorbid illness (e.g., kidney disease). While these other specialties may initiate ULT if indicated, there may be an expectation for the PCP to provide ongoing gout care management. Moreover, training in musculoskeletal medicine is limited in medical school curricula, with knowledge gaps in musculoskeletal medicine demonstrated among early career physicians and PCPs (23). The other possibility is that the various specialties are adhering to different sets of guidelines, with family medicine and internal medicine potentially subscribing to the ACP guidelines that do not recommend for or against SUA monitoring, while rheumatologists are subscribing to the EULAR, BSR, and ACR guidelines, which recommend a treat-to-target SUA-level strategy (21).

For patient-level factors, increasing patient age, residence in a rural area, and being in the lowest income quintile were associated with lower odds of having SUA testing. Considering the older age of patients in this study with a high prevalence of comorbidities, the increasing complexity of multimorbidity management may have impacted the quality of gout care due to competing demands from other health conditions. Rural residence is often associated with lower quality of care due to reduced health services available in rural areas compared to urban centers, and other sociodeterminants of health differences across regions (24). The relationship between comorbidities and SUA testing varied by condition. Those with concomitant CKD, hypertension, and diabetes mellitus had higher odds of SUA testing, while patients with COPD and prior cardiovascular events had lower odds. These findings can possibly be explained by the nature of each comorbidity. Namely, conditions such as CKD, hypertension, and diabetes mellitus all require routine blood work for monitoring as per their respective clinical practice guidelines, whether stemming from the underlying disease itself or from associated pharmacotherapy (25–27). Thus, the health care provider may have added SUA to the monitoring blood work perhaps not solely for the patient's underlying gout, and the patient may have already been present for blood work for other reasons besides gout. However, why patients with prior cardiovascular events (who should be undergoing serial laboratory monitoring for cholesterol levels, for example) were less likely to have SUA testing is unclear. Yet within Canada, significant deficiencies in cardiovascular care have been reported, highlighting a need for quality improvement initiatives to strengthen health service provisions to patients experiencing major cardiovascular events (28).

Moreover, ULT prescription characteristics played a role in SUA testing, with lower allopurinol starting doses at 50 mg daily

being associated with a higher uptake of SUA monitoring. These findings may have been explained by the fact that prescribers who initiated lower doses of ULT may have been more likely to dose escalate as per guideline recommendations by monitoring SUA levels (3,4). Patients who were coprescribed colchicine/oral corticosteroids were also associated with a higher likelihood of SUA testing. Hence, patients who received SUA monitoring may be the same patients whose ULT prescribers adhered to guideline concordant care in terms of initiating gout flare prophylaxis therapy when starting ULT.

A major study strength is the comprehensiveness of our data within a universal health care program, where patients access the majority of medical services and obtain laboratory investigations. We also combined diagnosis codes alongside ULT data to strengthen the ascertainment of our study population and performed sensitivity analyses on our case ascertainment approach. However, the use of health administrative data is marred by certain limitations, as the data were not collected specifically for health research purposes. First, the accuracy of identifying gout using Ontario administrative data has not been quantified. While all individuals in this study were required to have both gout diagnosis codes and, importantly, ULT dispensing claims, which greatly reduced the potential for misclassification bias, we were not able to use gout classification criteria for case definition (29). Conversely, our approach to ascertain gout patients (which prioritizes specificity and positive predictive value) possibly accompanied lower sensitivity at identifying all gout patients. Due to the limitations of administrative data, we lacked clinical information on gout severity and gout flares, which may have influenced both physician and patient behavior in SUA testing. Moreover, we were unable to identify tests ordered by physicians in which patients failed to have their SUA performed. Additionally, we did not assess adherence to ULT, which could have been an important factor in influencing SUA monitoring, as we may have overestimated the degree of SUA monitoring in our cohort in assuming that all patients took ULT as prescribed and should have undergone SUA testing.

We further assumed that physicians who prescribed ULT were those responsible for ordering SUA, in analyzing the role of the prescriber in gout care. In terms of generalizability of our data, since our study was population-based within a universal health care system where medically necessary services, including SUA monitoring, are provided free of cost at the point of access to all citizens, we expected that our rates of monitoring may be greater than other jurisdictions with health care privatization. However, an intrinsic limitation of administrative data is that we only included patients with established gout diagnoses and initiated treatment, and thus we were unable to identify individuals with undiagnosed gout who are undertreated.

In conclusion, we demonstrated that SUA testing after index ULT dispensation among older gout patients was lower than currently recommended by guidelines for gout, though SUA testing appears to be improving over time. Large variation exists in practice patterns for monitoring across different physician specialties, with rheumatologists and family physicians having the highest and lowest percentages of their patients having SUA testing, respectively. Last, our study suggests that patient, prescription, health services, and ULT prescriber factors are correlated with SUA testing. A need exists to further study the causal mechanisms behind which factors ultimately influence SUA testing and to pinpoint tangible areas for improvement.

ACKNOWLEDGMENTS

The authors thank IQVIA Solutions Canada Inc. for use of their Drug Information File. Parts of this material are based on data and information compiled and provided by the Ministry of Health, Ontario Health, and the Canadian Institute for Health Information. This document used data adapted from the Statistics Canada Postal Code Conversion File, which is based on data licensed from Canada Post Corporation, and/or data adapted from the Ontario Ministry of Health Postal Code Conversion File, which contains data copied under license from Canada Post Corporation and Statistics Canada.

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. Kwok 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. Kwok, Kuriya, Hawker, Widdifield.

Acquisition of data. Kwok, Li, Widdifield.

Analysis and interpretation of data. Kwok, Kuriya, Hawker, Li, Choy, Widdifield.

REFERENCES

1. Stamp L, Morillon MB, Taylor WJ, et al. Serum urate as surrogate endpoint for flares in people with gout: a systematic review and meta-regression analysis. Semin Arthritis Rheum 2018;48:293–301.
2. Kwok TS, Xu VY, Lake SL. Gout. CMAJ 2021;193:E171.
3. Richette P, Doherty M, Pascual E, et al. 2016 updated EULAR evidence-based recommendations for the management of gout [review]. Ann Rheum Dis 2017;76:29–42.
4. FitzGerald JD, Dalbeth N, Mikuls T, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Care Res (Hoboken) 2020;72:744–60.
5. Jennings CG, Mackenzie IS, Flynn R, et al. Up-titration of allopurinol in patients with gout. Semin Arthritis Rheum 2014;44:25–30.
6. FitzGerald JD, Mikuls TR, Neogi T, et al. Development of the American College of Rheumatology electronic clinical quality measures for gout. Arthritis Care Res (Hoboken) 2018;70:659–71.
7. Hanly JG, Skedgel C, Sketris I, et al. Gout in the elderly: a population health study. J Rheumatol 2009;36:822–30.
8. Sarawate CA, Brewer KK, Yang W, et al. Gout medication treatment patterns and adherence to standards of care from a managed care perspective. Mayo Clin Proc 2006;81:925–34.
9. Hughes JC, Wallace JL, Bryant CL, et al. Monitoring of urate-lowering therapy among US veterans following the 2012 American College of Rheumatology guidelines for management of gout. Ann Pharmacother 2017;51:301–6.
10. Belay Y, Yirdaw K, Enawgaw B. Tumor lysis syndrome in patients with hematological malignancies [review]. J Oncol 2017;2017:9684909.

11. Goicoechea M, de Vinuesa SG, Verdalles U, et al. Effect of allopurinol in chronic kidney disease progression and cardiovascular risk. Clin J Am Soc Nephrol 2010;5:1388–93.
12. Weiner DE, Tighiouart H, Elsayed EF, et al. Uric acid and incident kidney disease in the community. J Am Soc Nephrol 2008;19:1204–11.
13. Austin PC. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. Commun Stat Simul Comput 2009;38:1228–34.
14. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med 2009;28:3083–107.
15. Kerr GS, Richards JS, Nunziato CA, et al. Measuring physician adherence with gout quality indicators: a role for natural language processing. Arthritis Care Res (Hoboken) 2015;67:273–9.
16. Jackson G, Dalbeth N, Te Karu L, et al. Variation in gout care in Aotearoa New Zealand: a national analysis of quality markers. N Z Med J 2014;127:37–47.
17. Russell MD, Rutherford AI, Ellis B, et al. Management of gout following 2016/2017 European (EULAR) and British (BSR) guidelines: an interrupted time-series analysis in the United Kingdom. Lancet Reg Health Eur 2022;18:100416.
18. Zhang W, Doherty M, Bardin T, et al. EULAR evidence based recommendations for gout. Part II: management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT) [review]. Ann Rheum Dis 2006;65:1312–24.
19. Jordan KM, Cameron JS, Snaith M, et al. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of gout. Rheumatology (Oxford) 2007;46:1372–4.
20. Khanna D, Fitzgerald JD, Khanna PP, et al. 2012 American College of Rheumatology guidelines for management of gout. Part 1: systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. Arthritis Care Res (Hoboken) 2012;64:1431–46.
21. Qaseem A, Harris RP, Forciea MA, et al. Management of acute and recurrent gout: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017;166:58–68.
22. Harrold LR, Mazor KM, Negron A, et al. Primary care providers' knowledge, beliefs and treatment practices for gout: results of a physician questionnaire. Rheumatology (Oxford) 2013;52:1623–9.
23. Al-Nammari SS, James BK, Ramachandran M. The inadequacy of musculoskeletal knowledge after foundation training in the United Kingdom. J Bone Joint Surg Br 2009;91:1413–8.
24. Kapral MK, Austin PC, Jeyakumar G, et al. Rural-urban differences in stroke risk factors, incidence, and mortality in people with and without prior stroke. Circ Cardiovasc Qual Outcomes 2019;12:e004973.
25. Diabetes Canada Clinical Practice Guidelines Expert Committee, Berard LD, Siemens R, et al. Monitoring glycemic control. Can J Diabetes 2018;42 Suppl 1:S47–53.
26. Rabi DM, McBrien KA, Sapir-Pichhadze R, et al. Hypertension Canada's 2020 comprehensive guidelines for the prevention, diagnosis, risk assessment, and treatment of hypertension in adults and children [review]. Can J Cardiol 2020;36:596–624.
27. Akbari A, Clase CM, Acott P, et al. Canadian Society of Nephrology commentary on the KDIGO clinical practice guideline for CKD evaluation and management [review]. Am J Kidney Dis 2015;65:177–205.
28. Tran DT, Welsh RC, Ohinmaa A, et al. Quality of acute myocardial infarction care in Canada: a 10-year review of 30-day in-hospital mortality and 30-day hospital readmission. Can J Cardiol 2017;33:1319–26.
29. Neogi T, Jansen TL, Dalbeth N, et al. 2015 gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2015;74:1789–98.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2472–2480
DOI 10.1002/acr.25178
© 2023 The Authors. *Arthritis Care & Research* published by Wiley Periodicals LLC on behalf of American College of Rheumatology.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

American College of Rheumatology
Empowering Rheumatology Professionals

Sex Differences in Pain and Quantitative Sensory Testing in Patients With Rheumatoid Arthritis

Kelly Vogel,[1] Lutfiyya N. Muhammad,[1] Jing Song,[1] Tuhina Neogi,[2] Clifton O. Bingham,[3] Marcy B. Bolster,[4] Wendy Marder,[5] Alyssa Wohlfahrt,[6] Daniel J. Clauw,[5] Dorothy Dunlop,[1] and Yvonne C. Lee[1]

Objective. Women with rheumatoid arthritis (RA) have higher pain and worse functional outcomes compared to men, even when treated with similar medications. The objective of this study was to identify sex differences in pain intensity, pain interference, and quantitative sensory tests (QST), which are independent of inflammation, in patients with RA.

Methods. This study is a post hoc analysis of participants in the Central Pain in Rheumatoid Arthritis cohort. Pain intensity was assessed using a 0–10 numeric rating scale. Pain interference was measured using a Patient-Reported Outcomes Measurement Information System computerized adaptive test. QST included pressure pain detection thresholds, temporal summation, and conditioned pain modulation. Women and men were compared using multiple linear regression, adjusted for age, education, race, research site, depression, obesity, RA disease duration, swollen joint count, and C-reactive protein.

Results. Mean ± SD pain intensity was 5.32 ± 2.29 among women with RA, compared to 4.60 ± 2.23 among men with RA (adjusted difference 0.83 [95% confidence interval (95% CI) 0.14, 1.53]). Women with RA had lower pressure pain detection thresholds at the trapezius (adjusted difference –1.22 [95% CI –1.73, –0.72]), wrist (adjusted difference –0.57 [95% CI –1.07, –0.06]), and knee (adjusted difference –1.10 [95% CI –2.00, –0.21]). No statistically significant differences in pain interference, temporal summation, and conditioned pain modulation were observed.

Conclusion. Women reported higher pain intensity and lower pressure pain detection thresholds (higher pain sensitivity) than men. However, pain interference, temporal summation, and conditioned pain modulation did not differ between men and women.

INTRODUCTION

Women are disproportionately affected by autoimmune conditions and chronic pain syndromes (1,2). For example, rheumatoid arthritis (RA) occurs at a female to male ratio of 3:1 (2), and, compared to men with RA, women with RA suffer from more active disease. Composite disease activity scores, such as the Disease Activity Score in 28 joints (DAS28) include objective (e.g., C-reactive protein [CRP] level) and subjective (e.g., patient global assessment of disease activity, tender joint count [TJC]) components (3,4). The subjective components are highly influenced by pain, which is in concurrence with women with RA reporting higher pain ratings (5–7). Research suggests that the disparity in pain is not fully explained by differences in disease severity (i.e., structural damage) or access to treatment (8,9).

Variations in central nervous system (CNS) regulation of pain may be one contributor to sex differences in pain intensity. CNS regulation of pain can be assessed using quantitative sensory tests (QST), including tests of pressure pain detection threshold (PPTs), temporal summation (TS), and conditioned pain modulation (CPM). Studies of healthy individuals have reported that women are more sensitive to a wide range of noxious pain stimuli

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grants R01-AR-064850, K24-AR-080840, and P30-AR-072579 to Dr. Lee, grant P30-AR-072579 to Dr. Muhammad, grant P30-AR-072579 to Ms. Song, grants K24-AR-070892 and P30-AR-072571 to Dr. Neogi, and grant P30-AR-070254 to Dr. Bingham).

[1]Kelly Vogel, MD, Lutfiyya N. Muhammad, PhD, Jing Song, MS, Dorothy Dunlop, PhD, Yvonne C. Lee, MD, MMSc: Northwestern University Feinberg School of Medicine, Chicago, Illinois; [2]Tuhina Neogi, MD, PhD: Boston University School of Medicine, Boston, Massachusetts; [3]Clifton O. Bingham, MD: Johns Hopkins University, Baltimore, Maryland; [4]Marcy B. Bolster, MD: Massachusetts General Hospital, Boston; [5]Wendy Marder, MD, Daniel J. Clauw, MD: University of Michigan Medical School, Ann Arbor; [6]Alyssa Wohlfahrt, MS: Tufts University School of Medicine, Boston, Massachusetts.

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25178.

Address correspondence via email to Yvonne C. Lee, MD, MMSc, at Yvonne.Lee@northwestern.edu.

Submitted for publication September 6, 2022; accepted in revised form June 22, 2023.

SIGNIFICANCE & INNOVATIONS

- Women with rheumatoid arthritis (RA) reported more intense pain than men, but there were no statistically significant sex differences in pain interference.
- This study is the first to report that, compared to men with RA, women with RA are more sensitive to pressure applied at both articular and nonarticular sites.
- Differences in pain intensity and pressure pain detection thresholds remained despite controlling for peripheral inflammation (assessed by swollen joint counts) and systemic inflammation (measured by C-reactive protein levels).

(pressure, heat, cold, ischemic, electrical, etc.) than men (10). In addition, some studies suggest that healthy women have higher TS of pain than healthy men, although these studies reveal more mixed results than the studies on pain sensitivity (11). To our knowledge, only one small study of 18 women and men has examined sex differences in PPTs in patients with RA and did not find significant differences in PPTs between men and women (12).

In this study, we examine patient-reported pain intensity, pain interference, and QST-based measures of CNS pain regulation in women compared to men with RA. We hypothesize that women with RA report higher pain intensity and higher pain interference and exhibit greater abnormalities in QST-derived measures of CNS pain regulation than men with RA. Furthermore, we hypothesize that these differences are independent of potential sex-differences in peripheral and systemic inflammation.

PATIENTS AND METHODS

Study population. This study is a post hoc analysis of baseline data from the Central Pain in Rheumatoid Arthritis (CPIRA) cohort. CPIRA is a multicenter, prospective, observational study of participants designed to examine the association between pain mechanisms, pain intensity, and treatment response (13,14). Participants were recruited from 5 US academic medical centers from January 2014 to July 2017. Inclusion criteria for the parent study (and hence this post hoc analysis) were: 1) a diagnosis of RA based on the 2010 American College of Rheumatology/EULAR criteria, and 2) starting or switching a disease-modifying antirheumatic drug (DMARD) due to active RA. Exclusion criteria were: 1) changing doses of centrally acting pain medications (e.g., amitriptyline, gabapentin, or duloxetine) within 3 months of enrollment; 2) >10 mg of prednisone daily or its equivalent; 3) chronic opioid use or any opioid use within 24 hours of study date; 4) systemic autoimmune disease other than RA; 5) severe Raynaud's disease requiring pharmacologic treatment; 6) severe peripheral vascular disease manifested by claudication or ischemic rest pain; or 7) self-reported peripheral neuropathy diagnosis.

The CPIRA study was approved by the institutional review boards (IRBs) at each of the 5 participating academic medical centers (Brigham and Women's Hospital, Massachusetts General Hospital, Johns Hopkins University, University of Michigan, Boston University). The Northwestern University IRB determined that this study, which is a post hoc analysis of CPIRA data, met the criteria for exemption from further IRB review.

Assessment of clinical variables. Baseline clinical variables were assessed at the initial study visit, prior to DMARD initiation or change. These variables included age, sex, race, education, body mass index (BMI), RA disease duration, and patient global assessment. Race and ethnicity were self-reported from a fixed set of categories. For race, this included American Indian or Alaskan Native, Asian, Black or African American, Native Hawaiian or other Pacific Islanders, White, or other. For ethnicity, participants could select Hispanic or Latino or not Hispanic or Latino. Sex was assessed by the question: "What is your gender?", with answer choices being male and female, consistent with National Institutes of Health reporting guidelines. Serum from the baseline visit was analyzed for CRP at a single laboratory. A trained assessor performed a 28 swollen joint count (SJC) and TJC; DAS28 score was calculated based on this information (15).

Patient-reported pain metrics. Pain intensity was assessed by a 0–10 numerical rating scale with the lead question: "In the past 7 days, how would you rate your pain on average?". Pain interference and depression were assessed by the Patient-Reported Outcomes Measurement Information System (PROMIS) computerized assessment tests. Participants were considered to have symptoms consistent with depression if they had a T score of ≥60 on the PROMIS bank v1.0 depression scale (16).

QST. All participants underwent QST at baseline. We utilized 3 types of QSTs: 1) pressure pain detection thresholds (PPTs); 2) temporal summation (TS); and 3) conditioned pain modulation (CPM), as described in previous publications (17,18). Intraclass correlation coefficients for PPT and TS measurements ranged from 0.71 to 0.90, which is considered good to excellent. The intraclass correlation coefficient for CPM was 0.45, which is considered fair (19).

PPTs. A Force 10 FDX algometer (Wagner) was used to assess PPTs at the bilateral knees, wrists, trapezius muscles, and thumbnails. The algometer probe was placed on the center of the target, and force was applied at 0.5 kgf/second until pain was reported. Three trials were performed per side. The PPT was defined as the mean pressure at which pain was reported. Low PPTs at extraarticular sites were considered to represent

increased pain sensitivity due to dysregulation of CNS pain mechanisms or other causes of widespread pain sensitivity, such as circulating factors (cytokines), genetic factors, etc. Low PPTs at joint sites were considered to reflect abnormalities in both central and peripheral processes, including sensitization of the peripheral nociceptor by joint inflammation (20). Differences in PPTs do not represent a single mechanism of CNS dysregulation, and we are unable to entirely disentangle differences in this metric from other confounding psychological and cultural influences.

Temporal summation. Participants were tested with 6 blunt-tipped, punctate probes with forces ranging from 8 mN to 256 mN. Probes of increasing weight were tested on the participant's dorsal forearm until a pain score of 30–40 of 100 was produced. The probe generating a pain score between 30 and 40 was used for further testing. If no such pain rating was achieved, the highest weighted probe was used. The selected probe was tapped 10 times on the dorsal forearm. The participant was asked to rate their pain on a 0–100 scale at taps 1, 5, and 10. TS was calculated by subtracting the pain score at the first tap from the pain score at the tenth tap. The mean TS was calculated by taking the average of 3 trials. Higher TS values were indicative of higher levels of central pain sensitization.

CPM. CPM was assessed using a painful conditioning stimulus to activate the descending inhibitory pain pathways and a test stimulus to assess pain sensitivity. The conditioning stimulus was produced by inserting the participant's right hand into a 5–7°C water bath. The test stimulus was pressure produced by an algometer placed at the center of the contralateral trapezius. PPTs were measured immediately prior to hand submersion in the cold-water bath and after 20 seconds of cold-water submersion. The ratio of the second PPT to the first PPT was calculated. Inefficient (lower) CPM was considered indicative of abnormalities in descending pain inhibition.

Statistical analysis. The primary outcomes were pain intensity and pain interference. Secondary outcomes were PPT at the knees, PPT at the wrists, PPT at the thumbnails, PPT at the trapezius muscles, TS, and CPM. Means and SDs were calculated and stratified by sex. The unadjusted and adjusted differences between women versus men as well as the 95% confidence intervals (95% CIs) were calculated. Difference in pain measures between sexes were examined using multiple linear regression. The diagnostics of the multiple regression models were checked during the construction of the models, and standard assumptions of linear regression were met. All adjusted models included age, education, race, research site, depression, and obesity as covariates because these variables may be related to the pain outcomes (21–25). RA disease duration, SJC, and CRP level were also included in the models to account for the contribution of RA-related characteristics, including peripheral and systemic inflammation, to pain (3,26). Sensitivity analyses were performed, excluding depression and SJC as covariates. Statistical testing used a nominal $\alpha = 0.05$. All analyses were performed using SAS software, version 9.4, and R.

RESULTS

Clinical characteristics. Of the 295 participants included in the parent CPIRA study (242 female and 53 male), 280 individuals (230 female and 50 male) had nonmissing data in all 7 outcomes of interest. Following exclusion of participants with missing predictor data (race, RA duration, or CRP level), 268 individuals remained (220 female and 48 male). Participants excluded for missing data had similar baseline characteristics compared to those included in the cohort, with the exception that excluded participants were less likely to be White and less likely to have some college education or higher (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25178).

Most of the sample were women (82.1%). Most demographic characteristics were balanced between men and women, although mean ± SD age was slightly lower for women (53.82 ± 14.29 years) than men (and 58.06 ± 10.94 years). Compared to women, a slightly higher proportion of men self-reported as White (83.3% versus 73.2%), and a higher proportion of men were obese (47.9% versus 33.2%) (Table 1). The metrics of

Table 1. Baseline characteristics of study participants*

Characteristic	Women (n = 220)	Men (n = 48)
Age, years	53.82 ± 14.29	58.06 ± 10.94
Race, no. (%)		
White	161 (73.2)	40 (83.3)
African/African American/Black	38 (17.3)	4 (8.3)
Asian	11 (5.0)	0 (0)
American Indian/Alaskan Native	0 (0)	1 (2.8)
Unknown	10 (4.6)	3 (6.3)
Ethnicity, no. (%)		
Spanish/Hispanic/Latino	15 (6.8)	3 (6.3)
Not Spanish/Hispanic/Latino	205 (93.2)	45 (93.8)
Some college or higher, no. (%)	169 (76.8)	34 (70.8)
BMI, kg/m^2	28.02 ± 6.65	30.66 ± 6.07
Obesity, no. (%)	73 (33.2)	23 (47.9)
Depression, no. (%)	33 (15.0)	6 (12.5)
Disease duration, years	9.84 ± 11.58	9.81 ± 13.73
CRP, mg/liter	8.28 ± 12.83	7.11 ± 10.01
Swollen joint count	4.92 ± 4.57	5.58 ± 6.00
Tender joint count	10.61 ± 8.53	10.17 ± 7.89
Patient global assessment score†	4.28 ± 2.47	3.81 ± 2.19
DAS28 score†	4.38 ± 1.22	4.24 ± 1.27

* Values are the mean ± SD unless indicated otherwise. BMI = body mass index; CRP = C-reactive protein; DAS28 = Disease Activity Score in 28 joints.
† Women: n = 186; men: n = 41.

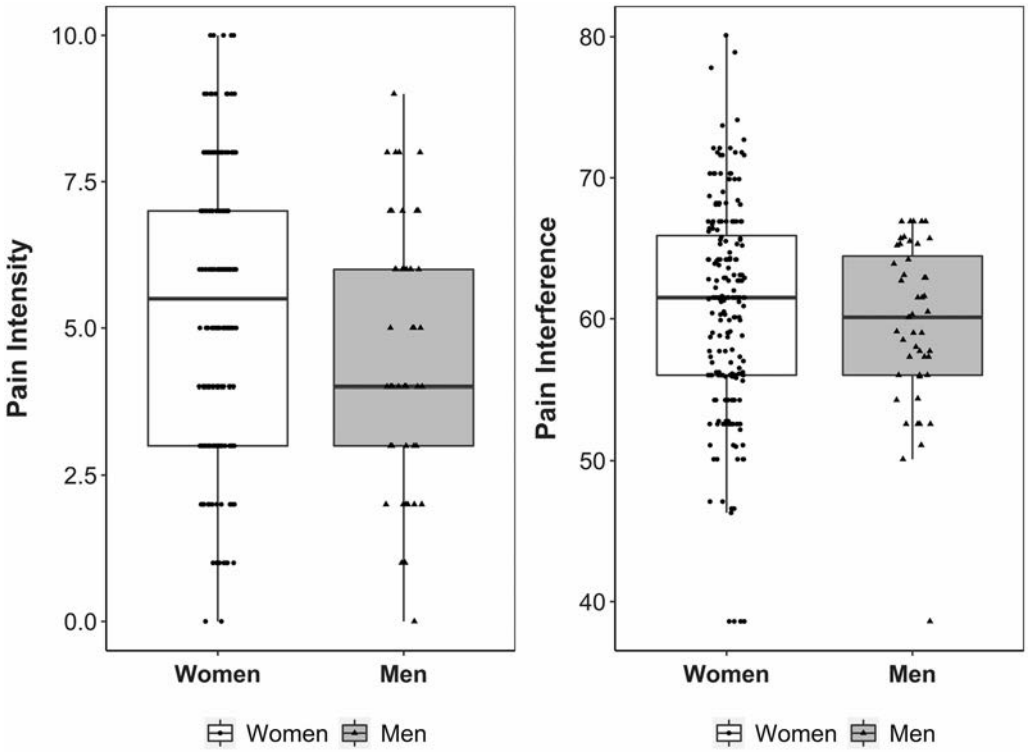


Figure 1. Sex differences in patient-reported pain intensity and interference. Raw data points for women represented by circles and for men by triangles. Within each box, horizontal lines denote the median; boxes extend from the 25th to 75th percentile of each group's distribution of values; and vertical lines extending from the boxes denote the most extreme values within 1.5 interquartile range of the 25th and 75th percentile of each group.

disease severity were balanced between men and women (Table 1). Pain medication use was also similar between men and women (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25178).

Patient-reported pain metrics. Compared to men, women reported greater pain intensity on average (5.32 versus 4.60). This difference was statistically significant (unadjusted difference 0.72 [95% CI 0.004, 1.43]; adjusted difference 0.83 [95% CI 0.14, 1.53]; standardized mean difference 0.31 [95% CI 0.002, 0.63]). PROMIS pain interference scores were not significantly different between men and women (unadjusted difference 0.99 [95% CI −1.25, 3.23]; adjusted difference 1.12 [95% CI −1.04, 3.28]) (Figure 1 and Table 2).

Results of QSTs. Compared to men, women had significantly lower PPTs at the trapezius, wrist, and knee in unadjusted and adjusted analyses. On average, women also had lower PPT at the thumbnail compared to men, although this difference was not statistically significant (Figure 2 and Table 3). There were no statistically significant differences in temporal summation (unadjusted difference 2.06 [95% CI −2.55, 6.67]; adjusted difference: 1.54 [95% CI −3.19, 6.27]) or CPM between women and men (unadjusted difference: 0.08 [95% CI −0.03, 0.19]; adjusted

Table 2. Sex differences in patient-reported pain intensity and pain interference*

Survey	Women (n = 220)	Men (n = 48)	Unadjusted difference women vs. men (95% CI)	Adjusted difference women vs. men (95% CI)†
Pain intensity numeric rating scale	5.32 ± 2.29	4.60 ± 2.23	0.72 (0.004, 1.43)‡	0.83 (0.14, 1.53)‡
PROMIS pain interference scale	60.61 ± 7.40	59.63 ± 5.73	0.99 (−1.25, 3.23)	1.12 (−1.04, 3.28)

* Values are the mean ± SD unless indicated otherwise. 95% CI = 95% confidence interval; PROMIS = Patient-Reported Outcomes Measurement Information System.
† Adjusted for research site, age, education, race, obesity, depression, disease duration, C-reactive protein level, and swollen joint count.
‡ Significant.

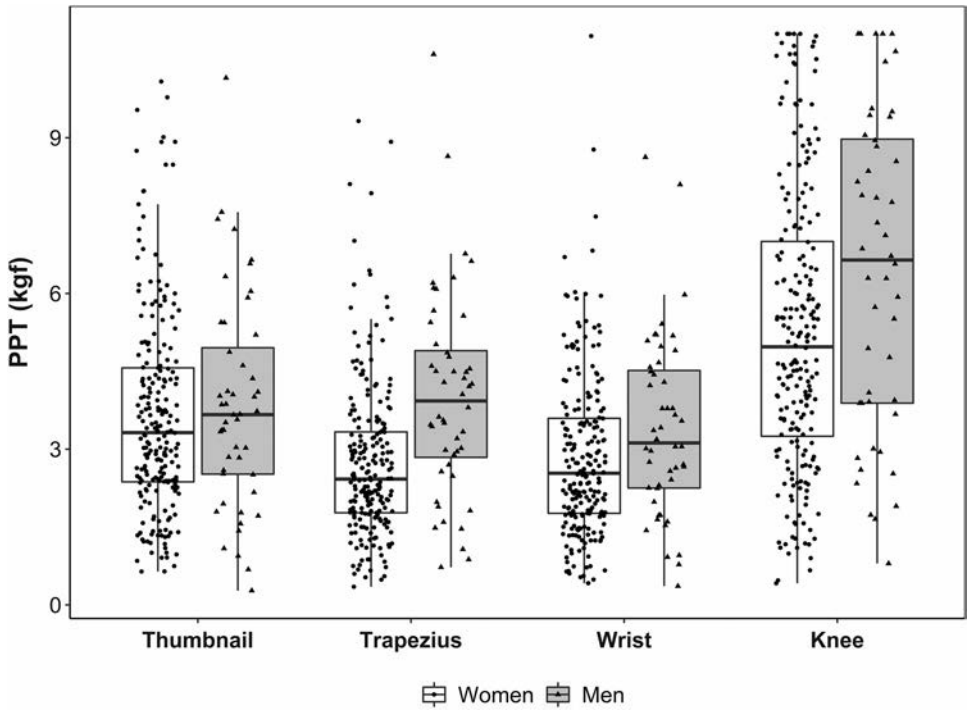


Figure 2. Sex differences in pressure pain detection thresholds. Raw data points for women represented by circles and for men by triangles. Within each box, horizontal lines denote the median; boxes extend from the 25th to 75th percentile of each group's distribution of values; and vertical lines extending from the boxes denote the most extreme values within 1.5 interquartile range of the 25th and 75th percentile of each group.

difference 0.06 [95% CI –0.05, 0.18]). Sensitivity analyses excluding depression and SJC as covariates yielded similar results.

DISCUSSION

The objective of this study was to identify sex differences in pain, which are independent of peripheral and systemic inflammation, in patients with RA. In analyses adjusted for SJC and CRP, women with RA reported higher pain intensity but no statistically significant difference in pain interference. Women also had lower PPTs at the trapezius, wrist, and knee. No statistically significant differences in TS and CPM were observed between men and women.

The finding that women reported higher pain intensity than men is in concordance with a meta-analysis of individuals with inflammatory arthritis. Barnabe and colleagues reported that all but 1 of 24 studies found that women with RA reported higher pain intensity than men with RA (5). The standardized mean difference between women and men was 0.21 (95% CI 0.16, 0.26; $P < 0.001$) (5), which was similar to the standardized mean difference observed in our study (0.31 [95% CI 0.002, 0.63]).

While the difference in pain intensity between men and women was statistically significant, the clinical significance of these findings is less clear. The 95% CI of 0.14 to 1.53 represents a range of potential true differences, which includes values above and below the minimum clinically important difference (MCID) in pain intensity of 1 unit on a scale of 0–10 (27). The relatively wide confidence interval likely reflects variability in the data. This variability is expected because pain is a subjective experience that is often multifactorial in etiology. Given the range in values, however, we cannot conclude that the difference observed in this study is clinically meaningful. Nevertheless, the consistency in our results

Table 3. Sex differences in pressure pain detection thresholds*

Location	Women (n = 220)	Men (n = 48)	Unadjusted difference women vs. men (95% CI)	Adjusted difference women vs. men (95% CI)†
Thumbnail	3.65 ± 1.94	3.85 ± 2.04	–0.19 (–0.81, 0.42)	–0.31 (–0.95, 0.32)
Trapezius	2.74 ± 1.49	4.02 ± 1.98	–1.28 (–1.77, –0.78)‡	–1.22 (–1.73, –0.72)‡
Wrist	2.85 ± 1.56	3.40 ± 1.74	–0.54 (–1.05, –0.04)‡	–0.57 (–1.07, –0.06)‡
Knee	5.28 ± 2.74	6.44 ± 3.06	–1.16 (–2.04, –0.28)‡	–1.10 (–2.00, –0.21)‡

* Values are the mean ± SD unless indicated otherwise. 95% CI = 95% confidence interval.
† Adjusted for research site, age, education, race, obesity, depression, disease duration, C-reactive protein level, and swollen joint count.
‡ Significant.

with the results of aforementioned meta-analysis (5) supports the conclusion that these differences are real. Understanding the mechanisms underlying these differences is necessary to develop targeted therapies to improve pain management in this era of precision medicine.

To our knowledge, no studies have examined sex-differences in pain interference in patients with RA. In this study, the point estimate of the adjusted difference in PROMIS pain interference was 1.12, which is small compared to the MCID of 2–3 (28,29). Furthermore, the 95% CI overlapped 0. Taken in conjunction with the small sex differences in pain intensity, this finding suggests that pain-associated function does not differ in a clinically meaningful manner between sexes. The range of possible values in PROMIS pain interference, however, was relatively wide, with a lower bound of –1.04 and an upper bound of 3.28. As such, the clinical interpretation of this finding remains unclear.

This study is unique because, in addition to examining sex differences in patient-reported pain intensity and pain interference, we also assessed sex differences in QST measures. Consistent with literature in the general population and other chronic pain conditions, women reported lower PPTs compared to men (30,31). The value of QST lies in the information that they, in aggregate, provide on pain phenotype, which, in turn, provides information regarding the neurobiological mechanisms underlying that phenotype. Differences in PPTs at joint sites (wrists and knees) could occur due to sex-based differences in peripheral sensitization, central sensitization, or both (20). The observation that women also reported lower PPTs at non-joint sites, though only statistically significant at the trapezius, suggests that these differences may be partially mediated by sex-based differences in CNS regulation of pain. Alternatively, the underlying neurobiological pathways may not differ between men and women, but rather, the differences in PPTs may be related to differences in psychological and/or cultural mechanisms that influence pain reporting. While PPTs are thought to be objective measures of pain sensitivity, they ultimately still rely on patient reporting of the first sensation of pain.

To probe potential pathways leading to the observed differences in PPTs, we examined sex differences in TS and CPM. To our knowledge, this was the first study to compare TS between men and women with RA. We did not observe significant differences in TS between men and women with RA. Most studies in healthy populations have reported higher TS in women than men, although there have also been several studies showing no differences between groups (11,32). One study of patients with osteoarthritis (OA) reported increased TS in women compared to men (33). Taken in totality, there is insufficient evidence to support sex-based differences in TS among patients with RA, OA, and other arthritic disease populations.

This was the first study to examine sex differences in CPM among patients with RA. We did not observe a difference in CPM between men and women with RA. Similarly, a systematic review by Racine et al reported that most studies involving healthy subjects found no difference in CPM between women and men (32). Interestingly, a recent study reported that inter-individual differences accounted for 24–34% of the variance in CPM while age, sex, and intensity of the conditioning stimulus together only explained <3–12% of the variance (34). Possible inter-individual differences contributing to CPM include genetic variations (e.g., polymorphisms in a serotonin transporter gene), behavioral differences (e.g., cardiovascular reactivity to pain), and psychological traits (e.g., anxiety, depression, catastrophizing) (35–37).

These findings may have important research implications. As previously discussed, our results are consistent with multiple prior studies showing that women with inflammatory arthritis report higher pain intensity than men (5). In addition, our results are internally consistent. Compared to men, women had higher pain intensity and lower PPTs across multiple body sites. This consistency provides confidence that these differences are real. We want to emphasize, however, that the clinical relevance of each specific QST (e.g., PPTs, TS, CPM), taken in isolation, is unclear. No data exist regarding the MCID in PPTs, TS, and CPM in patients with RA. While some data exist for other populations (e.g., patients with mandibular pain, neck pain, back pain), we are not confident in the applicability of these results given differences in study population and testing procedures (38,39). Nevertheless, we believe the data in this manuscript reveal interesting avenues for future research which should be pursued in larger studies, which are specifically designed to examine sex differences in pain and pain mechanisms. Understanding these differences will be important for improving equity of care and developing precision-based medicine approaches to pain management.

Existing literature has identified sex-based neurobiological differences in pain perception, which range from the level of transcription in the dorsal root ganglia to differences in cell types important for pain processing in the brain (40). For example, expression of colony stimulating factor 1, a factor important for inducing CNS changes to promote mechanical hypersensitivity, was higher following chronic constriction injury in the dorsal ganglion of female rodents compared to their male counterparts (41). In addition, animal studies have shown that microglia, a type of central immune cell, are important mediators of pain hypersensitivity in male but not female mice (42). Studies in humans, using PET imaging to tag microglia, may be helpful in further understanding differences in CNS regulation between men and women.

In addition to differences in the neuroimmune regulation of pain, hormonal differences may alter the pain experience through multiple pathways. Testosterone seems to be antinociceptive given the correlation between decreased androgen concentrations and chronic pain, whereas the effect of estradiol and progesterone on pain appears to be more complex (43). Sex hormones may also influence activation of brain opioid receptors. Specifically, women with high estradiol have been found to have

decreased pain sensitivity and increased brain mu-opioid receptor binding compared to women with low estradiol (43).

Furthermore, sex/gender differences in psychological factors (e.g., coping, catastrophizing, affect), as well as societal influences on gender role expectations may have a large impact on pain (44). Depression is more prevalent in women than men and is highly comorbid with pain (24). Symptoms of depression are associated with processes that may augment pain, including cognitive distortions and lower levels of positive reinforcement (45). There may also be a neurobiological underpinning to this association, given that both pain perception and depression involve serotonin and norepinephrine signaling (46). Furthermore, different coping strategies for pain are established from a young age. For example, women score higher on pain catastrophizing, a negative form of coping characterized by amplified negative reactions to pain (24,47,48). Finally, social pressures and stereotypes are also likely to influence differences in response to pain by gender. The gender expectation is that men will exhibit stoicism and thus express less pain (49). Interestingly, Robinson et al showed that gender differences in QSTs are reduced when adjusting for gender role expectations (50).

There are limitations to our study. First, this work was a post hoc analysis. The parent CPIRA study was not designed to look for sex differences in pain, and the question used to assess sex was not explicit about sex assignment at birth versus gender identity. While it asked about "gender," it did not include nonbinary options. Given the timeframe during which data were collected (2014–2017) and the lack of nonbinary answer options, we infer that the participants understood this question to be regarding their sex assigned at birth rather than a reflection of their gender identity. However, we acknowledge that we cannot disentangle the effects of the neurobiological effects of sex from the influence of systemic and structural factors of gender identity that may influence the pain experience. Second, the study population may not be representative of all RA patients. We only enrolled participants switching or initiating DMARD therapy due to active disease. Thus, they likely had more pain and inflammation than the general RA population. Furthermore, the sample size of men in this study was smaller than the sample size for women and may have limited our ability to detect significant sex differences in survey and QST measurements. The difference in sample sizes likely reflects the lower prevalence of men than women with RA. However, we cannot exclude the possibility of a contribution from selection bias. Third, a higher proportion of men than women were taking as needed opioid medications, which could, in theory, influence outcomes. We think this is unlikely given that we excluded all patients who were taking opioids regularly (more often than not over a 3-month period) from the study, thereby excluding individuals who may have long-term changes in pain processing, such as opioid-induced hyperalgesia. Short-term effects of opioids were mitigated by requiring participants to hold their opioids for at least 24 hours prior to study procedures. Finally, while we controlled for SJC and CRP as markers of peripheral and systemic inflammation, more sensitive methods of detecting synovitis (e.g., ultrasound, magnetic resonance imaging) were not utilized. Thus, we are not able to exclude the possibility that residual differences in peripheral and/or systemic inflammation may be confounding our analysis and contributing to observed differences in pain.

Our study has multiple strengths. First, it controlled for differences in disease activity by including SJC and CRP as covariates in the models, allowing us to assess sex differences in pain and pain mechanisms, independent of peripheral and systemic inflammation. Second, this is one of few studies with data on QST in patients with RA, and the first to show a difference between women and men in QSTs in this population. To our knowledge, only one other study examined differences in PPTs between men and women with RA, and that study included only 18 men and 18 women (12).

In conclusion, our work demonstrates sex differences in pain intensity and QST assessments of pain mechanisms in patients with RA independent of inflammation. Women reported higher pain intensity than men. Compared to men, women were more sensitive to pain at articular and nonarticular sites, suggesting differences in peripheral and CNS regulation of pain. These results indicate that sex-based differences impact neurobiological functioning and, ultimately, the pain experience, in patients with RA. These observations are impactful because patients and health care providers frequently consider pain as an indicator of inflammation, and composite disease activity measures include assessments that are influenced by the pain experience (i.e., patient global assessment) and pain sensitivity (i.e., TJC). However, there are no modifications to disease activity thresholds based on sex or gender. Because these assessments directly affect treatment decisions and outcomes, it will be increasingly important for researchers and health care providers to consider the impact of sex and gender on pain, particularly as the field of rheumatology moves toward individualized treatment plans a part of a precision medicine treatment approach.

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. Lee 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. Vogel, Muhammad, Song, Neogi, Bingham, Bolster, Marder, Wohlfahrt, Clauw, Dunlop, Lee.

Acquisition of data. Neogi, Bingham, Bolster, Marder, Wohlfahrt, Lee.

Analysis and interpretation of data. Vogel, Muhammad, Song, Dunlop, Lee.

REFERENCES

1. Mogil JS. Sex differences in pain and pain inhibition: multiple explanations of a controversial phenomenon [review]. Nat Rev Neurosci 2012; 13:859–66.

2. Quintero OL, Amador-Patarroyo MJ, Montoya-Ortiz G, et al. Autoimmune disease and gender: plausible mechanisms for the female predominance of autoimmunity [review]. J Autoimmun 2012;38: J109–19.
3. Intriago M, Maldonado G, Cardenas J, et al. Clinical characteristics in patients with rheumatoid arthritis: differences between genders. ScientificWorldJournal 2019;2019:8103812.
4. Sokka T, Toloza S, Cutolo M, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. Arthritis Res Ther 2009;11:R7.
5. Barnabe C, Bessette L, Flanagan C, et al. Sex differences in pain scores and localization in inflammatory arthritis: a systematic review and metaanalysis [review]. J Rheumatol 2012;39:1221–30.
6. Shin S, Park EH, Kang EH, et al. Sex differences in clinical characteristics and their influence on clinical outcomes in an observational cohort of patients with rheumatoid arthritis. Joint Bone Spine 2021;88: 105124.
7. Hallert E, Bjork M, Dahlstrom O, et al. Disease activity and disability in women and men with early rheumatoid arthritis (RA): an 8-year followup of a Swedish early RA project. Arthritis Care Res (Hoboken) 2012;64:1101–7.
8. Jawaheer D, Olsen J, Hetland ML. Sex differences in response to antitumor necrosis factor therapy in early and established rheumatoid arthritis: results from the DANBIO registry. J Rheumatol 2012;39: 46–53.
9. Hekmat K, Jacobsson LT, Nilsson JA, et al. Changes and sex differences in patient reported outcomes in rheumatoid factor positive RA-results from a community based study. BMC Musculoskelet Disord 2014;15:44.
10. Mogil JS. Qualitative sex differences in pain processing: emerging evidence of a biased literature [review]. Nat Rev Neurosci 2020;21: 353–65.
11. Fillingim RB, King CD, Ribeiro-Dasilva MC, et al. Sex, gender, and pain: a review of recent clinical and experimental findings [review]. J Pain 2009;10:447–85.
12. Gerecz-Simon EM, Tunks ER, Heale JA, et al. Measurement of pain threshold in patients with rheumatoid arthritis, osteoarthritis, ankylosing spondylitis, and healthy controls. Clin Rheumatol 1989;8:467–74.
13. Lee YC, Bingham CO III, Edwards RR, et al. Association between pain sensitization and disease activity in patients with rheumatoid arthritis: a cross-sectional study. Arthritis Care Res (Hoboken) 2018;70: 197–204.
14. Wohlfahrt A, Bingham CO III, Marder W, et al. Responsiveness of patient-reported outcomes measurement information system measures in rheumatoid arthritis patients starting or switching a disease-modifying antirheumatic drug. Arthritis Care Res (Hoboken) 2019;71:521–9.
15. Van der Heijde DM, van 't Hof MA, van Riel PL, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49:916–20.
16. Pilkonis PA, Yu L, Dodds NE, et al. Validation of the depression item bank from the Patient-Reported Outcomes Measurement Information System (PROMIS) in a three-month observational study. J Psychiatr Res 2014;56:112–9.
17. Heisler AC, Song J, Dunlop DD, et al. Association of pain centralization and patient-reported pain in active rheumatoid arthritis. Arthritis Care Res (Hoboken) 2020;72:1122–9.
18. Heisler AC, Song J, Muhammad LN, et al. Association of dysregulated central pain processing and response to disease-modifying antirheumatic drug therapy in rheumatoid arthritis. Arthritis Rheumatol 2020; 72:2017–24.
19. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;64:284–90.
20. Suokas AK, Walsh DA, McWilliams DF, et al. Quantitative sensory testing in painful osteoarthritis: a systematic review and meta-analysis [review]. Osteoarthritis Cartilage 2012;20:1075–85.
21. Lautenbacher S, Peters JH, Heesen M, et al. Age changes in pain perception: a systematic-review and meta-analysis of age effects on pain and tolerance thresholds [review]. Neurosci Biobehav Rev 2017;75: 104–13.
22. Jiang X, Sandberg ME, Saevarsdottir S, et al. Higher education is associated with a better rheumatoid arthritis outcome concerning for pain and function but not disease activity: results from the EIRA cohort and Swedish rheumatology register. Arthritis Res Ther 2015;17:317.
23. Kim HJ, Yang GS, Greenspan JD, et al. Racial and ethnic differences in experimental pain sensitivity: systematic review and meta-analysis [review]. Pain 2017;158:194–211.
24. Mogil JS, Bailey AL. Sex and gender differences in pain and analgesia [review]. Prog Brain Res 2010;186:141–57.
25. Liu Y, Hazlewood GS, Kaplan GG, et al. Impact of obesity on remission and disease activity in rheumatoid arthritis: a systematic review and meta-analysis [review]. Arthritis Care Res (Hoboken) 2017;69: 157–65.
26. Pollard LC, Ibrahim F, Choy EH, et al. Pain thresholds in rheumatoid arthritis: the effect of tender point counts and disease duration. J Rheumatol 2012;39:28–31.
27. Salaffi F, Stancati A, Silvestri CA, et al. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004;8:283–91.
28. Bartlett SJ, Gutierrez AK, Andersen KM, et al. Identifying Minimal and meaningful change in a patient-reported outcomes measurement information system for rheumatoid arthritis: use of multiple methods and perspectives. Arthritis Care Res (Hoboken) 2022;74:588–97.
29. Chen CX, Kroenke K, Stump TE, et al. Estimating minimally important differences for the PROMIS pain interference scales: results from 3 randomized clinical trials. Pain 2018;159:775–82.
30. Neville SJ, Clauw AD, Moser SE, et al. Association between the 2011 Fibromyalgia Survey Criteria and multisite pain sensitivity in knee osteoarthritis. Clin J Pain 2018;34:909–17.
31. Petrini L, Matthiesen ST, Arendt-Nielsen L. The effect of age and gender on pressure pain thresholds and suprathreshold stimuli. Perception 2015;44:587–96.
32. Racine M, Tousignant-Laflamme Y, Kloda LA, et al. A systematic literature review of 10 years of research on sex/gender and pain perception - part 2: do biopsychosocial factors alter pain sensitivity differently in women and men? [review] Pain 2012;153:619–35.
33. Bartley EJ, King CD, Sibille KT, et al. Enhanced pain sensitivity among individuals with symptomatic knee osteoarthritis: potential sex differences in central sensitization. Arthritis Care Res (Hoboken) 2016;68: 472–80.
34. Graeff P, Itter A, Wach K, et al. Inter-individual differences explain more variance in conditioned pain modulation than age, sex and conditioning stimulus intensity combined. Brain Sci 2021;11:1186.
35. Lindstedt F, Berrebi J, Greayer E, et al. Conditioned pain modulation is associated with common polymorphisms in the serotonin transporter gene. PLoS One 2011;6:e18252.
36. Chalaye P, Devoize L, Lafrenaye S, et al. Cardiovascular influences on conditioned pain modulation. Pain 2013;154:1377–82.
37. Ruscheweyh R, Albers C, Kreusch A, et al. The effect of catastrophizing self-statements on pain perception and the nociceptive flexor reflex (RIII reflex). Clin J Pain 2013;29:725–32.
38. Calixtre LB, Oliveira AB, Alburquerque-Sendin F, et al. What is the minimal important difference of pain intensity, mandibular function,

and headache impact in patients with temporomandibular disorders? Clinical significance analysis of a randomized controlled trial. Musculoskelet Sci Pract 2020;46:102108.

39. Jorgensen R, Ris I, Juhl C, et al. Responsiveness of clinical tests for people with neck pain. BMC Musculoskelet Disord 2017;18:548.
40. Delay L, Goncalves Dos Santos G, Dias EV, et al. Sexual dimorphism in the expression of pain phenotype in preclinical models of rheumatoid arthritis [review]. Rheum Dis Clin North Am 2021;47:245–64.
41. Stephens KE, Zhou W, Ji Z, et al. Sex differences in gene regulation in the dorsal root ganglion after nerve injury. BMC Genomics 2019;20:147.
42. Sorge RE, Mapplebeck JC, Rosen S, et al. Different immune cells mediate mechanical pain hypersensitivity in male and female mice. Nat Neurosci 2015;18:1081–3.
43. Bartley EJ, Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings [review]. Br J Anaesth 2013;111:52–8.
44. Salk RH, Hyde JS, Abramson LY. Gender differences in depression in representative national samples: meta-analyses of diagnoses and symptoms. Psychol Bull 2017;143:783–822.
45. Munce SE, Stewart DE. Gender differences in depression and chronic pain conditions in a national epidemiologic survey. Psychosomatics 2007;48:394–9.
46. Meana M. The meeting of pain and depression: comorbidity in women [review]. Can J Psychiatry 1998;43:893–9.
47. Edwards RR, Haythornthwaite JA, Sullivan MJ, et al. Catastrophizing as a mediator of sex differences in pain: differential effects for daily pain versus laboratory-induced pain. Pain 2004;111:335–41.
48. Keefe FJ, Lefebvre JC, Egert JR, et al. The relationship of gender to pain, pain behavior, and disability in osteoarthritis patients: the role of catastrophizing. Pain 2000;87:325–34.
49. Melchior M, Poisbeau P, Gaumond I, et al. Insights into the mechanisms and the emergence of sex-differences in pain [review]. Neuroscience 2016;338:63–80.
50. Robinson ME, Wise EA, Gagnon C, et al. Influences of gender role and anxiety on sex differences in temporal summation of pain. J Pain 2004;5:77–82.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2481–2488
DOI 10.1002/acr.25166
© 2023 American College of Rheumatology. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

American College of Rheumatology
Empowering Rheumatology Professionals

Trends in Fracture Rates Over Two Decades Among Veterans With Ankylosing Spondylitis

Sali Merjanah,[1] Jean W. Liew,[1] John Bihn,[2] Nathanael R. Fillmore,[3] Mary T. Brophy,[4] Nhan V. Do,[4] and Maureen Dubreuil[4]

Objective. There is an increased risk of fracture in individuals with ankylosing spondylitis (AS) compared to the general population, possibly due to systemic inflammatory effects. The use of tumor necrosis factor inhibitors (TNFi) may reduce fracture risk by inhibiting inflammation. We assessed fracture rates in AS versus non-AS comparators and whether these rates have changed since the introduction of TNFi.

Methods. We used the national Veterans Affairs database to identify adults ≥18 years old with ≥1 International Classification of Diseases, Ninth Revision (ICD-9)/ICD-10 code for AS and at least 1 disease-modifying antirheumatic drug prescription. As comparators, we selected a random sample of adults without AS diagnosis codes. We calculated fracture incidence rates for AS and comparators, with direct standardization to the cohort structure in 2017. To compare fracture rates from 2000 to 2002 (pre-TNFi) versus 2004–2020 (TNFi era), we performed an interrupted time series analysis.

Results. We included 3,794 individuals with AS (mean age 53 years, 92% male) and 1,152,805 comparators (mean age 60 years, 89% male). For AS, the incidence rate of fractures increased from 7.9/1,000 person-years in 2000 to 21.6/1,000 person-years in 2020. The rate also increased among comparators, although the ratio of fracture rates (AS/comparators) remained relatively stable. In the interrupted time series, the fracture rate for AS patients in the TNFi era was nonsignificantly increased compared to the pre-TNFi era.

Conclusion. Fracture rates have increased over time for both AS and non-AS comparators. The fracture rate in individuals with AS did not decrease after TNFi introduction in 2003.

INTRODUCTION

There is mounting evidence of the impact of comorbidities on disease activity, functional disability, and mortality in axial spondyloarthritis, including ankylosing spondylitis (AS) (1–3). Abnormalities in bone metabolism, including osteoporosis and fractures, are a concern due to their association with morbidity and mortality. The prevalence of fracture has been estimated to be 3.9% in AS. The risk of vertebral fractures in AS is increased 2- to 4-fold compared to the general population, and the risk of all nonvertebral fractures is 10% higher in AS than age- and sex-matched comparators, in 1 study (4,5). Cervical spine fractures were the leading cause of in-hospital mortality in a study of hospitalized AS patients (6).

The increased risk of vertebral fracture among this population is thought to be primarily due to local bone remodeling, causing some areas of excess bone formation and some areas of decreased bone density (7). Excess bone in the spine leads to higher spinal rigidity and vulnerability to fractures even with trivial trauma. Increased fracture risk at sites outside the spine is thought to be due to effects of systemic inflammation. The proinflammatory cytokine tumor necrosis factor (TNF) has been linked to increased bone resorption through activation of osteoclasts (8–10).

The views expressed are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the US government.

Supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant *R03*-AR-076495).

Dr. Liew's work was supported by the Spondylarthritis Research and Treatment Network seed grant and the Rheumatology Research Foundation Investigator award. Mr. Bihn and Drs. Fillmore, Brophy, and Do's work as supported by the VA Cooperative Studies Program. Dr. Fillmore's work was supported by the American Heart Association (grant 870726). Dr. Dubreuil's work was supported by the SPARTAN Bruckel Award.

[1]Sali Merjanah, MD, Jean W. Liew, MD, MS: Boston University Chobanian and Avedisian School of Medicine, Boston, Massachusetts; [2]John Bihn, BA: VA Boston Healthcare System, Boston, Massachusetts; [3]Nathanael R. Fillmore, PhD: VA Boston Healthcare System and Harvard Medical School, Boston, Massachusetts; [4]Mary T. Brophy, MD, Nhan V. Do, MD, Maureen Dubreuil, MD, MSc: VA Boston Healthcare System and Boston University Chobanian and Avedisian School of Medicine, Boston, Massachusetts.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25166.

Address correspondence via email to Maureen Dubreuil, MD, MSc, at mdubreui@bu.edu.

Submitted for publication September 26, 2022; accepted in revised form May 23, 2023.

SIGNIFICANCE & INNOVATIONS

- The risk of fracture is increased in individuals with ankylosing spondyloarthritis (AS) compared to the non-AS comparators of the same age and sex, possibly due to systemic inflammation. Despite some studies suggesting improved bone density with tumor necrosis factor inhibitor (TNFi) use among AS patients, no large-scale studies have assessed the effect of TNFi on fracture risk in this population.
- In this study, fracture rates increased over the past 2 decades for both individuals with AS and non-AS comparators. These findings suggest increasing detection and diagnosis of fractures over time overall.
- In the interrupted time series, there was no reduction in fracture rates in AS patients with the introduction of TNFi. Therefore, more work needs to be done to explore other potential interventions beyond screening and treating traditional risk factors for fractures among people with AS.

TNF inhibitors (TNFi) are theorized to lower the risk of osteoporosis and fracture by inhibiting TNF, the driver of inflammation in AS (11,12). The first TNFi, etanercept, gained US Food and Drug Administration approval for treatment of AS in 2003, and the use of TNFi is recommended for those with AS who do not have clinical improvement with use of nonsteroidal antiinflammatory (NSAID) drugs. However, there is a paucity of evidence as to whether TNFi have effects on clinical aspects of AS beyond spinal inflammation and disease activity. No large-scale studies have assessed the effect of TNFi on fracture risk in AS patients, despite some studies suggesting increased bone density with TNFi use among AS patients (13). Therefore, we studied the trends in the incidence rates of fractures among those with AS in time periods before and after the introduction of TNFi for treatment of AS, using an interrupted time series (ITS) analysis.

MATERIALS AND METHODS

Data source. This study was conducted using the national Veterans Affairs (VA) Clinical Data Warehouse (CDW), which collates electronic health records of veterans seen at VA facilities nationwide, including both outpatient visits and hospitalizations (14). These data include diagnoses, detailed medication prescriptions and pharmacy fill records, the results of laboratory and imaging studies, procedure reports, and vital status. The longitudinal nature of this data set makes it ideal for the study of chronic disease outcomes (15–17).

Study population. We included adults ≥18 years old who had ≥1 International Classification of Diseases, Ninth Revision (ICD-9) or ICD-10 code for AS from an outpatient data source separated by at least 7 days, and at least 1 disease-modifying antirheumatic drug (DMARD), immunosuppressive treatment (IST), or biologic medication fill or administration. The validity of AS diagnosis based on ICD-9/10 codes has been established and used for epidemiologic studies of AS with a positive predictive value of 71.8% (18,19). To be more confident that individuals classified as having AS truly had disease, we required at least 1 fill or administration of a relevant medication (DMARD, IST, or biologic) within 90 days before or 90 days after an ICD code for AS. We included subjects from 2000 to 2017; the final year for study entry was 2017, to allow for appropriate follow-up time for the outcome of interest.

As non-AS comparators, we selected a 20% random sample of adults without a prior AS diagnosis in each year and applied the same age and other inclusion criteria as for the AS group. Within each study year, we selected an eligible comparator for each individual who fulfilled the definition of AS as defined above.

Subjects were excluded if they had ≥1 ICD-9/10 code for rheumatoid arthritis (RA) or psoriatic arthritis (PsA), were missing age data, or if their age was <18 years on the date of AS diagnosis. Additionally, if subjects had history of the outcome of interest (radius, femoral, or vertebral fracture) prior to cohort eligibility, they were excluded to allow identification of incident outcomes.

Outcomes. The primary outcome of interest, any fracture in the radius, femur, or vertebra, was defined by ICD-9/10 or Current Procedural Terminology (CPT) codes (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25166).

For each AS subject, follow-up time began at the time of AS diagnosis as defined above. Follow-up for each comparator began at the time they met other inclusion criteria (age and VA health care utilization). Subjects were followed until they had a fracture, reached age 90 years, died, were disenrolled from the VA, or at the study end in December 2020.

Statistical analysis. We calculated fracture incidence rates over 6-month time periods from 2000 to 2020 as the number of incident fractures in each cohort divided by the number of person-years of follow-up for that cohort, separately for the AS and comparison groups. For each time period, only patients who received treatment at the VA within the calendar year, as determined by outpatient service and billing data, were included in the analysis.

For AS and comparators, direct standardization was used to standardize rates to the structure of the cohort in the year 2017 (the final year of diagnosis eligibility for the study). Incidence rates were stratified by sex. For male patients for whom the sample size was larger, rates were further stratified for age categories (<50, 50–59, 60–69, 70–79, 80–89 years) and race (White, Black, other, or unknown). In order to assess whether fractures were increasing more in AS than in comparators over time, we

calculated the ratio of fracture rates in the AS group versus comparators in each 6-month study period.

ITS. To assess for changes in fracture rates following the introduction of TNFi for use in AS, we then carried out an ITS analysis comparing the trends in 6-month incidence rates of fractures in 2 periods: the pre-TNFi era (2000–2002) and the TNFi-era (2004–2020). The 1-year interruption period, representing the introduction of TNFi for use in AS in the US, was set as the year 2003, which was when etanercept was approved for this indication. Fractures during the 1-year interruption period were not included in calculations of event rates in the pre-TNFi era (prior to 2003) or the TNFi era (2004–2020).

Upon visually inspecting the data, the AS fracture rate was substantially higher in the first half of 2001 (14.5 AS fractures/1,000 person-years) than all other points during this study period (range 4.5–8 AS fractures/1,000 person-years). Therefore, we excluded this point prior to running the ITS analysis. Additionally, to test the robustness of our results, we performed sensitivity analyses using different interruption periods (2004, 2005, and 2006) for the ITS, as changes in prescriptions and TNFi use implementation likely took place gradually following the approval of etanercept in 2003.

Lastly, to assess whether a certain type of fracture accounted for the overall fracture trend, we looked at the incidence rates of vertebral, femoral, and radius fracture separately using the

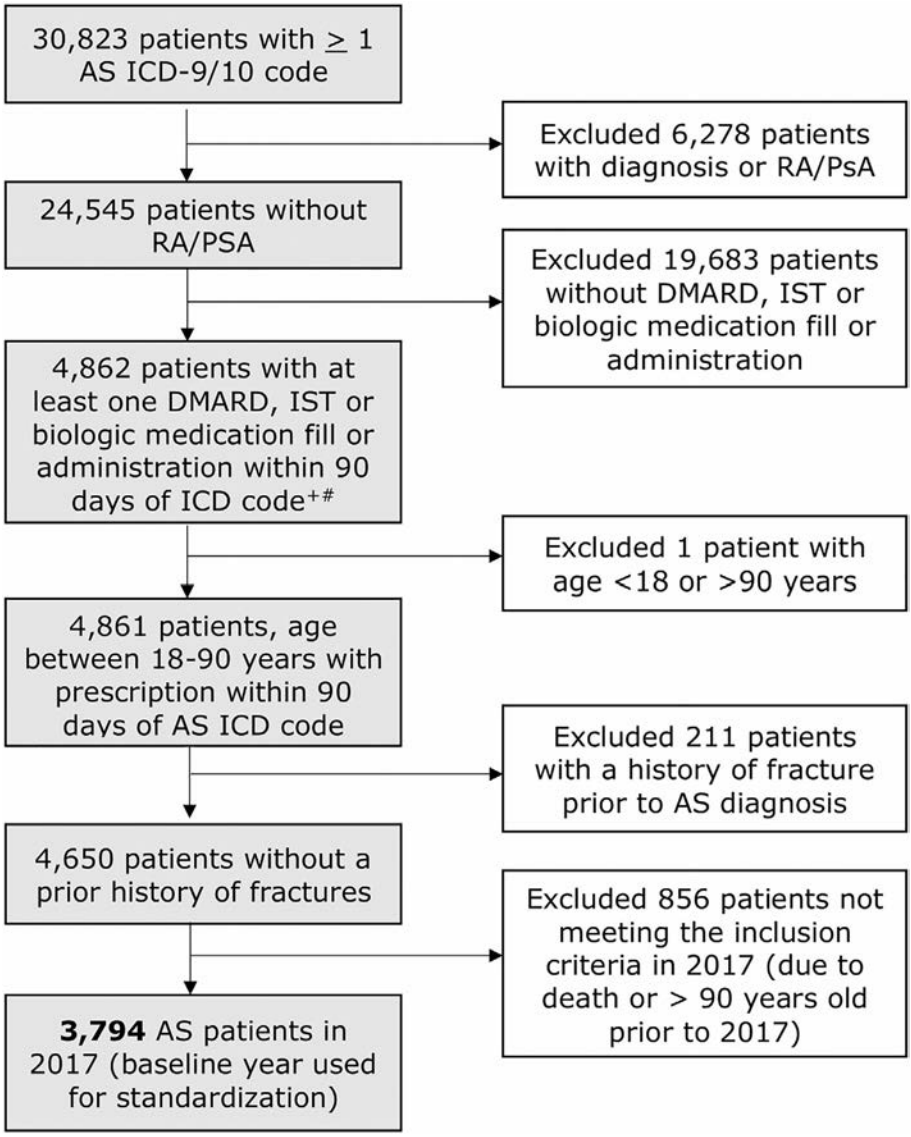


Figure 1. Flow diagram of study inclusion for the ankylosing spondylitis (AS) group. Disease-modifying antirheumatic drugs (DMARDs) and immunosuppressive therapy (IST) included apremilast, azathioprine, chloroquine, hydroxychloroquine, cyclosporine, methotrexate, leflunomide, sulfasalazine, mycophenolate mofetil and cyclophosphamide. Biologics included infliximab, adalimumab, etanercept, golimumab, certolizumab, secukinumab, ixekizumab, ustekinumab, rituximab, anakinra, abatacept, tocilizumab and sarilumab. ICD-9 = International Classification of Diseases (ICD); RA = rheumatoid arthritis. PsA = psoriatic arthritis.

nonstandardized incidence rates for each period, in order to show the change in these fracture rates over time.

RESULTS

Baseline characteristics. We identified 3,794 patients with AS who met the eligibility criteria for the study and 1,152,805 non-AS comparators. The flow diagram of study inclusion is shown in Figure 1.

Demographic and clinical characteristics of AS and comparator cohorts are shown in Table 1. For the AS group, the mean ± SD age was 53 ± 15 years, 92.2% were men, and 77% were White. The comparator group had a mean ± SD age of 60 ± 16 years, 89.4% were men, and 69.8% were White. The mean length of follow-up in patients with AS and in comparators was 5.7 years and 7.9 years, respectively. The use of NSAIDs, glucocorticoids, conventional synthetic DMARDs, TNFi, and non-TNFi biologics was more frequent in the AS group, as was prevalence of extra-musculoskeletal manifestations of AS, including enthesitis, uveitis, inflammatory bowel disease, and psoriasis. The prevalence of comorbidities, including diabetes mellitus, chronic kidney disease, and end stage renal disease was similar between both groups. Obesity (body mass index 30+ kg/m^2) and smoking were more commonly seen in the AS group.

Table 1. Demographic and clinical characteristics of AS and comparator groups*

	AS group (n = 3,794)	Comparison group (n = 1,152,805)
Age, mean ± SD years	53.3 ± 15.2	60.2 ± 16.7
Follow-up time, years	5.7	7.9
Sex†		
Female	7.8	10.6
Male	92.2	89.4
Race†		
Black	12.5	16.8
Other or unknown	10.5	13.4
White	77.0	69.8
Ethnicity†		
Not Hispanic or Latino	90.4	86.2
Hispanic or Latino	6.0	6.1
Unknown	3.6	7.6
NSAIDs	77.8	42.3
Glucocorticoids	54.6	28.5
csDMARDs +IST†	45.3	1.3
TNFi	82.8	0.3
Non-TNFi biologics‡	11.1	0.1
Enthesitis	26.8	14.0
IBD	19.3	5.1
Psoriasis	5.9	2.8
Uveitis	25.5	1.2
Diabetes mellitus	30.8	32.1
Hypertension	63.4	65.5
Chronic kidney disease	10.6	9.9
Peptic ulcer disease	51.4	38.5
Obesity (BMI ≥ 30 kg/m2)	46.0	38.4
Smoking	37.2	29.0
Cancer§	2.4	2.8
Congestive heart failure	7.7	8.0
ESRD	1.9	1.8
HIV	2.7	0.5
Ischemic heart disease	22.9	25.8
Liver disease	14.2	8.4
Pulmonary disease	2.7	3.0

* Values are the %. AS = ankylosing spondyloarthritis; BMI = body mass index; csDMARDs = conventional synthetic disease-modifying antirheumatic drugs; ESRD = end-stage renal disease; IBD = inflammatory bowel disease; IST = immunosuppressive treatment; NSAIDs = nonsteroidal antiinflammatory drugs; TNFi = tumor necrosis factor inhibitor.
† csDMARDS and IST included apremilast, azathioprine, chloroquine, hydroxychloroquine, cyclosporine, methotrexate, leflunomide, sulfasalazine, mycophenolate mofetil, and cyclophosphamide.
‡ Non-TNFi biologics included secukinumab, ixekizumab, ustekinumab, rituximab, anakinra, abatacept, tocilizumab, and sarilumab.
§ Excluding non-melanomatous skin cancer.

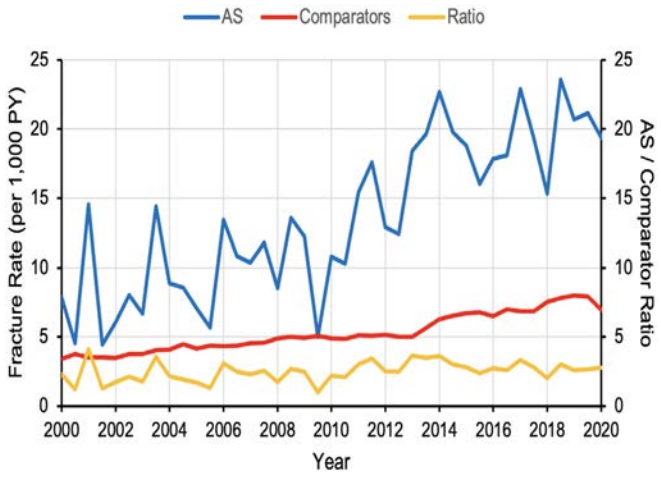


Figure 2. Fracture rates for ankylosing spondylitis (AS) and comparators. The yellow line represents the ratio of fracture rates for AS versus non-AS comparators. PY = person-year.

Primary results. Among AS patients, the fracture incidence rate was 7.9/1,000 person-years in the first half of 2000 compared to 3.4/1,000 person-years in the comparator cohort. The fracture rates were increased to 21.6 and 7.2 fractures/1,000 person-years among AS and comparators, respectively, at the end of the study in December 2020. Standardized fracture rates are shown in Supplementary Table 2 (available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25166). The ratio of fracture rates in the AS group compared to the comparison group ranged from 2.3 in 2000 to 3 in 2020 with a range from 1 to 4.1 (Figure 2).

ITS analysis. For the AS group in the pre-TNFi era from 2000 to 2002, the fracture rate increased by 0.04 fractures/1,000 person-years per each 6-month period after January to June 2000. Following the interruption period in 2003, the fracture incidence rate in the TNFi era increased by 0.19 fractures/1,000 person-years for each 6-month period

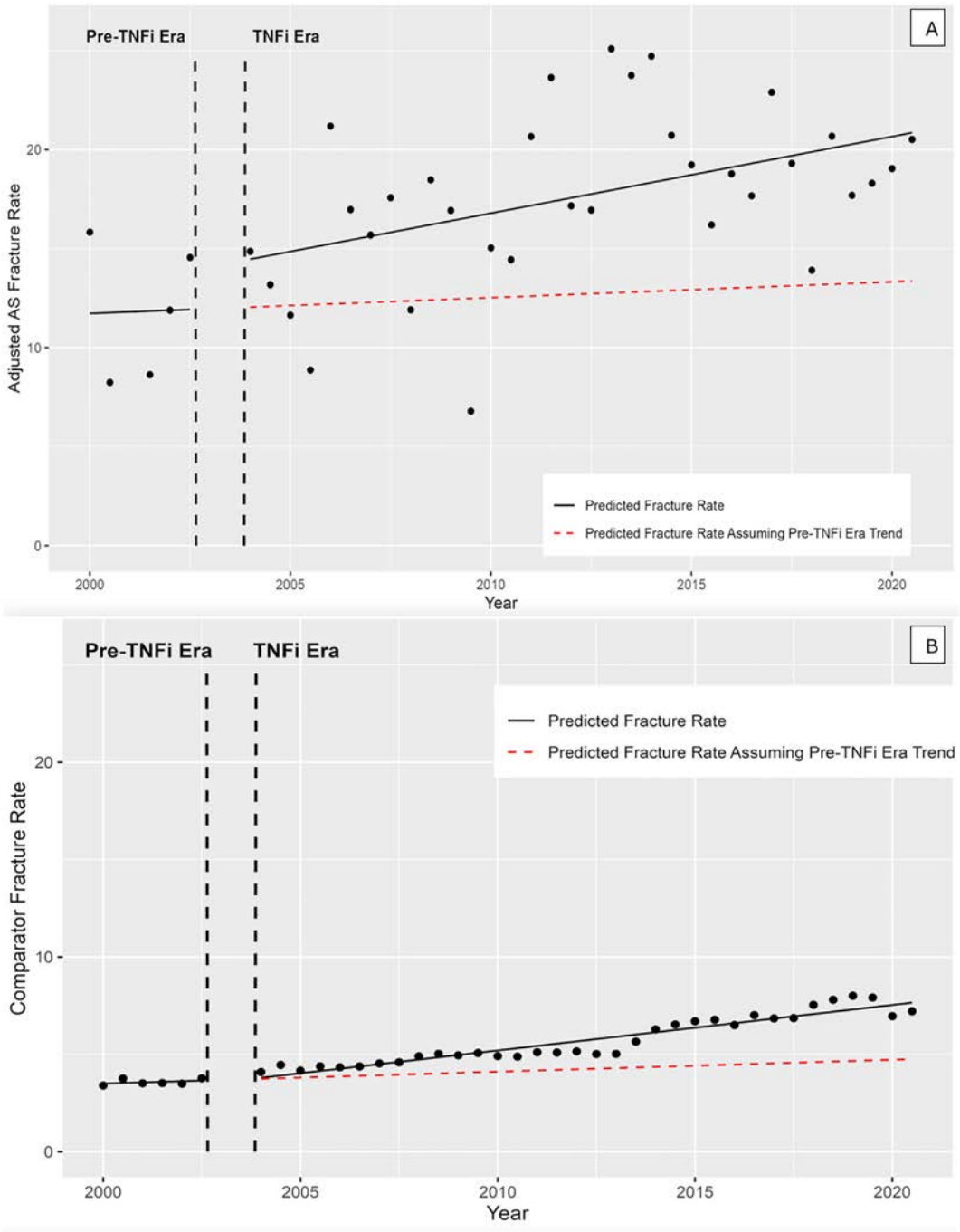


Figure 3. Results of the interrupted time series analysis for fracture rates in ankylosing spondylitis (AS) (**A**) versus comparators (**B**). The interruption period is the year 2003. The black line represents the trend in the fractures. The red dotted line represents the predicted fracture rate assuming that pre–tumor necrosis factor inhibitor (TNFi) era trends continued. Each data point represents fractures per 1,000 person-years. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25166/abstract.

(Figure 3A). The difference in fracture rates between the pre-TNFi and TNFi eras was 0.15 fractures/1,000 person-years ($P = 0.87$).

In the comparator group, the fracture rate in the pre-TNFi era increased by 0.03 fractures/1,000 person-years per 6-month period (Figure 3B). In the TNFi era, the incidence rate of fractures increased by 0.11 fractures/1,000 person-years per 6-month period. A difference of 0.08 fractures/1,000 person-years ($P = 0.35$) was observed between pre-TNFi and TNFi eras. The ITS analysis results were further detailed (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25166).

Sensitivity analysis. We assessed the impact of using different interruption periods on fracture rate trends between the

pre- and post-TNFi periods (see Supplementary Figures 1–3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25166). Overall, we saw similar trends compared to our main results.

The annual incidence rates of vertebral, femoral, and radius fracture for both AS and non-AS comparators are shown in Supplementary Figure 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25166). Vertebral fractures accounted for the majority of fractures in AS group (3 times more frequent than nonvertebral fractures). Overall, trends in fractures were increasing over time for all fracture types.

DISCUSSION

To our knowledge, this study is one of the largest US studies to assess fracture rates in individuals with AS over 2 decades, while considering the impact of biologic treatment. We observed an increase in the fracture incidence rate over time for both AS and non-AS comparator groups. These findings suggest a broader trend toward increased fracture occurrence for the general population overall, possibly through enhanced detection of fractures with more frequent or better resolution imaging, or in more accurate documentation of fractures within the medical records.

For the AS group, the ITS analysis demonstrated a nonsignificant increase in fracture rates in the TNFi era compared to the pre-TNFi era. Our findings were unable to demonstrate that the introduction and uptake of TNFi for the treatment of AS reduced the rate of fractures. These results highlight the need for clinical vigilance in the management of bone health in AS patients due to the known increased fracture risk associated with this condition.

The increased incidence of fractures in individuals with AS has been well demonstrated in multiple studies in the last decade, although most have focused on vertebral fractures (4,20–23). Vertebral compression fracture has been recognized as a common leading cause for hospitalization in AS patients and its incidence has been steadily increasing (24). Using the National Inpatient Sample, Wysham and colleagues found that cervical spine fracture was the leading cause of in-hospital mortality (6). Consistent with our results, a recent population-based study of 2,321 AS patients showed that there was increased vertebral fracture risk in AS compared to matched comparators without history of rheumatic disease; results remained unchanged after adjusting for multiple confounders including osteoporosis (21).

A study of 758 AS patients from a UK general practitioner database concluded that hip and forearm fractures were not significantly increased in AS compared to patients without AS (23). Similar results regarding hip fracture risk in AS were seen in 2 meta-analyses published in 2017. Zhang and colleagues found no statistically significant association between AS with the risk of any fracture or of hip fracture, despite finding an association between AS and increased risk of vertebral fractures (22). The hip fracture risk in these meta-analyses may have been underestimated due to the study design, the definition of fractures, and small sample sizes. Our results showed overall increased fracture rates for both vertebral and nonvertebral fractures.

Disease activity due to inflammation is associated with radiographic progression in the spine in AS patients (25,26). It has been hypothesized that treatment with antiinflammatory medications such as TNFi, via the reduction of disease activity and inflammation, slows or prevents radiographic progression in AS (27,28). Bone metabolism in AS is altered by systemic inflammation through multiple cytokine pathways (9). Studies have shown that TNF impairs bone turnover by activating osteoclasts, increasing osteoblast apoptosis, and decreasing osteoblast proliferation (13). Furthermore, studies found a beneficial effect of TNFi on bone mineral density measures in AS patients (13,29). Therefore, we expect that TNFi might reduce fracture risk through reduction of inflammation and disease activity.

There is a paucity of literature on whether the increased use of TNFi is protective against fractures. To our knowledge, only 2 earlier studies have evaluated fracture risk in AS in relation to TNFi use (21,30). An Australian population-based observational study looked at 2,321 patients with AS and found significantly increased risk of vertebral fracture in AS compared to matched non-AS comparators, but the risk did not change following the introduction of TNFi (21). A Swedish longitudinal cohort spanning 22 years found that the proportion of spinal fractures in hospitalized AS patients increased from 0.82% in 1987 up to 11.3% in 2008; however, this study could not account for predisposing factors for fractures such as age, sex, or osteoporosis (30). The increased fracture rates in AS patients in the Swedish study could be explained by greater clinical awareness of fractures, and the availability of advanced imaging modalities and resources. In our study, we attempted to address the secular trends toward increased fractures over time by reporting fracture rates in AS compared to those in comparators during the same calendar periods.

Certain limitations of our study must be acknowledged. We cannot fully exclude the possibility of misclassification of AS cases or fracture outcomes determined from ICD and CPT codes alone, although these definitions have been previously validated (18,19,21,30–32). Further, we were unable to account for the magnitude of bias from changes in diagnostic coding accuracy over time. Although we accounted for potential confounding through direct standardization for age, sex, race, and year, when possible, residual confounding may still affect the interpretation of our results. Subjects with more severe AS are more likely to be prescribed TNFi, and their severe disease would put them at risk for fracture, which may offset any potentially protective effect of treatment. Additionally, low fracture event rates, especially for AS, may have affected our study results and we may lack the precision to see a statistically significant difference. Finally, while the approval of etanercept and other TNFi for AS occurred at discrete

timepoints, use of TNFi in clinical practice may have been gradual and difficult to capture by applying a year-long interruption period in 2003 in our main ITS analysis.

The strengths of this study include the large population-based sample, strict inclusion and exclusion criteria, and longitudinal examination of trends in fracture rates over 2 decades. The interrupted time series is a commonly utilized and an established quasi-experimental design for evaluating the longitudinal effects of time-delimited interventions, such is the case for the introduction of TNFi for the treatment of AS in 2003 (33). The availability of longitudinal data allowed assessment of long-term effects of TNFi on AS fracture rates.

In conclusion, using a large cohort of AS patients and non-AS comparators from national VA data, we saw increased fracture rates over time for both AS and comparators. The interrupted time series analysis showed no decrease in the fracture incidence rate in AS patients with the introduction of TNFi. Although the evidence is strong behind using TNFi in AS for controlling disease activity, our results do not support the use of TNFi alone to decrease fracture risk. Further work should explore other potential interventions beyond screening and treating traditional risk factors for fractures among people with AS.

REFERENCES

1. Moltó A, Etcheto A, van der Heijde D, et al. Prevalence of comorbidities and evaluation of their screening in spondyloarthritis: results of the international cross-sectional ASAS-COMOSPA study. Ann Rheum Dis 2016;75:1016–23.
2. Nikiphorou E, Ramiro S, van der Heijde D, et al. Association of comorbidities in spondyloarthritis with poor function, work disability, and quality of life: results from the Assessment of SpondyloArthritis International Society Comorbidities in Spondyloarthritis Study. Arthritis Care Res (Hoboken) 2018;70:1257–62.
3. Fitzgerald G, Gallagher P, O'Shea FD. Multimorbidity in axial spondyloarthropathy and its association with disease outcomes: results from the Ankylosing Spondylitis Registry of Ireland cohort. J Rheumatol 2020;47:218–26.
4. Pray C, Feroz NI, Nigil Haroon N. Bone mineral density and fracture risk in ankylosing spondylitis: a meta-analysis. Calcif Tissue Int 2017; 101:182–92.
5. Walsh JA, Song X, Kim G, et al. Evaluation of the comorbidity burden in patients with ankylosing spondylitis using a large US administrative claims data set. Clin Rheumatol 2018;37:1869–78.
6. Wysham KD, Murray SG, Hills N, et al. Cervical spinal fracture and other diagnoses associated with mortality in hospitalized ankylosing spondylitis patients. Arthritis Care Res (Hoboken) 2017;69:271–7.
7. Davey-Ranasinghe N, Deodhar A. Osteoporosis and vertebral fractures in ankylosing spondylitis [review]. Curr Opin Rheumatol 2013; 25:509–16.
8. Appel H, Maier R, Wu P, et al. Analysis of IL-17(+) cells in facet joints of patients with spondyloarthritis suggests that the innate immune pathway might be of greater relevance than the Th17-mediated adaptive immune response Arthritis Res Ther 2011;13:R95.
9. Lories RJ, Haroon N. Bone formation in axial spondyloarthritis [review]. Best Pract Res Clin Rheumatol 2014;28:765–77.
10. Osta B, Benedetti G, Miossec P. Classical and paradoxical effects of TNF-α on bone homeostasis [review]. Front Immunol 2014;5:48.
11. Simone D, Al Mossawi MH, Bowness P. Progress in our understanding of the pathogenesis of ankylosing spondylitis [review]. Rheumatology (Oxford) 2018;57 Suppl 6:vi4–9.
12. Hakimian S, Kheder J, Arum S, et al. Re-evaluating osteoporosis and fracture risk in Crohn's disease patients in the era of TNF-alpha inhibitors. Scand J Gastroenterol 2018;53:168–72.
13. Haroon NN, Sriganthan J, Al Ghanim N, et al. Effect of TNF-alpha inhibitor treatment on bone mineral density in patients with ankylosing spondylitis: a systematic review and meta-analysis [review]. Semin Arthritis Rheum 2014;44:155–61.
14. United States Department of Veterans Affairs. National Center for Veterans Analysis and Statistics: quick facts. 2021. URL: https://www.va.gov/vetdata/Quick_Facts.asp.
15. Sauer BC, Teng CC, Tang D, 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.
16. Helget LN, England BR, Roul P, et al. Cause-specific mortality in patients with gout in the US Veterans Health Administration: a matched cohort study. Arthritis Care Res (Hoboken) 2023;75: 808–16.
17. Ebel AV, Lutt G, Poole JA, et al. Association of agricultural, occupational, and military inhalants with autoantibodies and disease features in US veterans with rheumatoid arthritis. Arthritis Rheumatol 2021;73: 392–400.
18. Singh JA, Holmgren AR, Krug H, et al. Accuracy of the diagnoses of spondylarthritides in Veterans Affairs medical center databases. Arthritis Care Res (Hoboken) 2007;57:648–55.
19. Dubreuil M, Peloquin C, Zhang Y, et al. Validity of ankylosing spondylitis diagnoses in The Health Improvement Network. Pharmacoepidemiol Drug Saf 2016;25:399–404.
20. Weiss RJ, Wick MC, Ackermann PW, et al. Increased fracture risk in patients with rheumatic disorders and other inflammatory diseases – a case-control study with 53,108 patients with fracture. J Rheumatol 2010;37:2247–50.
21. Ognjenovic M, Raymond WD, Inderjeeth CA, et al. The risk and consequences of vertebral fracture in patients with ankylosing spondylitis: a population-based data linkage study. J Rheumatol 2020;47: 1629–36.
22. Zhang M, Li XM, Wang GS, et al. The association between ankylosing spondylitis and the risk of any, hip, or vertebral fracture: a meta-analysis [review]. Medicine (Baltimore) 2017;96:e8458.
23. Vosse D, Landewé R, van der Heijde D, et al. Ankylosing spondylitis and the risk of fracture: results from a large primary care-based nested case-control study. Ann Rheum Dis 2009;68: 1839–42.
24. Dhital R, Oke I, Donato A, et al. Trends in hospitalizations for vertebral compression fracture in ankylosing spondylitis: data from the National Inpatient Sample 2000–2014. Clin Rheumatol 2021; 40:4927–32.
25. Maksymowych WP. MRI and X-ray in axial spondyloarthritis: the relationship between inflammatory and structural changes [review]. Arthritis Res Ther 2012;14:207.
26. Huang JX, Chung HY, Chui ET, et al. Intensity of spinal inflammation is associated with radiological structural damage in patients with active axial spondyloarthritis. Rheumatol Adv Pract 2019;4: rkz049.
27. Haroon N, Inman RD, Learch TJ, et al. The impact of tumor necrosis factor α inhibitors on radiographic progression in ankylosing spondylitis. Arthritis Rheum 2013;65:2645–54.
28. Karmacharya P, Duarte-Garcia A, Dubreuil M, et al. Effect of therapy on radiographic progression in axial spondyloarthritis: a

systematic review and meta-analysis. Arthritis Rheumatol 2020; 72:733–49.

29. Beek KJ, Rusman T, van der Weijden MA, et al. Long-term treatment with TNF-Alpha inhibitors improves bone mineral density but not vertebral fracture progression in ankylosing spondylitis. J Bone Miner Res 2019;34:1041–8.
30. Robinson Y, Sandén B, Olerud C. Increased occurrence of spinal fractures related to ankylosing spondylitis: a prospective 22-year cohort study in 17,764 patients from a national registry in Sweden. Patient Saf Surg 2013;7:2.
31. Wright NC, Daigle SG, Melton ME, et al. The design and validation of a new algorithm to identify incident fractures in administrative claims data. J Bone Miner Res 2019;34: 1798–807.
32. Quan H, Parsons GA, Ghali WA. Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data. Med Care 2004;42:801–9.
33. Kontopantelis E, Doran T, Springate DA, et al. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. BMJ 2015;350:h2750.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2489–2500
DOI 10.1002/acr.25158
© 2023 American College of Rheumatology.

American College of Rheumatology
Empowering Rheumatology Professionals

Burden of Knee Osteoarthritis in 204 Countries and Territories, 1990–2019: Results From the Global Burden of Disease Study 2019

Guangmin Yang,[1] Jue Wang,[2] Yun Liu,[3] Haojie Lu,[3] Liu He,[2] Changsheng Ma,[2] and Zhe Zhao[1]

Objective. To report the global, regional, and national estimates of knee osteoarthritis (OA) burden and associated risk factors (high body mass index [BMI]) by age, sex, and sociodemographic index (SDI) for 204 countries from 1990 to 2019.

Methods. We analyzed the prevalence, incidence, years lived with disability (YLDs), and age-standardized rates of knee OA using data from the Global Burden of Diseases, Injuries, and Risk Factors Study 2019. Estimates of the knee OA burden were derived from data modeled using a Bayesian meta-regression analytical tool (DisMod-MR 2.1).

Results. The global prevalence of knee OA in 2019 was ∼364.6 million (95% uncertainty interval [95% UI] 315.3 to 417.4). The age-standardized prevalence in 2019 was 4,376.0 per 100,000 (95% UI 3,793.0 to 5,004.9), an increase of 7.5% between 1990 and 2019. There were ∼29.5 million incident cases of knee OA in 2019 (95% UI 25.6 to 33.7), with an age-standardized incidence of 350.3 per 100,000 (95% UI 303.4 to 398.9). The global age-standardized YLD resulting from knee OA was 138.2 (95% UI 68.5 to 281.3) per 100,000 population in 2019, an increase of 7.8% (95% UI 7.1 to 8.4) from 1990. Globally in 2019, 22.4% (95% UI 12.1 to 34.2) of YLD resulting from knee OA was attributable to high BMI, an increase of 40.5% since 1990.

Conclusion. The prevalence, incidence, YLDs, and age-standardized rates of knee OA increased substantially in most countries and regions from 1990 to 2019. Continuous monitoring of this burden is important for establishing appropriate public prevention policies and raising public awareness, especially in high- and high–middle SDI regions.

INTRODUCTION

Osteoarthritis (OA), a heterogeneous, multifactorial chronic disease affecting multiple joints throughout the body, progresses differently between individuals over time (1,2). The knee joint is the most common site of OA (3,4). The structural lesions of knee OA that cause pain and loss of function are mainly involved in cartilage degeneration, bone remodeling, osteophyte formation, and chronic inflammation throughout the joint (5). Physical therapy, weight control, and pain relievers are the leading management modalities for knee OA in daily life (1,2). Advanced-stage knee OA can require surgical intervention, such as knee replacement, which can impose a high health and social burden on individuals and health systems.

Several studies have reported the prevalence of knee OA in a few nations, but its large-scale distribution and variability across countries and regions remain unclear (6). Using data from the Global Burden of Disease Study (GBD) 2010 and 2017, previous studies have reported the burden of musculoskeletal disease and arthritis in different regions and countries using different classification methods, but no study has yet specifically analyzed the global burden of knee OA (4–7). According to a study using the Dutch national public health database, the number of knee OA cases in the Netherlands was predicted to rise 41% between 2015 and 2040 (4,7). Given the increasing burden of knee OA, continuous disease surveillance is vital for planning and managing the health care needs of the population. On the basis of the GBD 2019 study, we analyzed prevalence, incidence, years lived with disability (YLDs), and associated risk factors (including high body mass index [BMI]) for knee OA by sex, age, and sociodemographic index (SDI) in 204 countries and regions worldwide.

[1]Guangmin Yang, MD, Zhe Zhao, MD: School of Clinical Medicine, Tsinghua University and Department of Orthopedics, Beijing Tsinghua Changgung Hospital, Tsinghua University, Beijing, China; [2]Jue Wang, MD, Liu He, MD, Changsheng Ma, MD: Department of Cardiology, Beijing Anzhen Hospital, Capital Medical University and National Clinical Research Center for Cardiovascular Diseases, Beijing, China; [3]Yun Liu, MD, Haojie Lu, MD: School of Clinical Medicine, Tsinghua University, Beijing, China.

Drs. Yang and Wang contributed equally to this work.

Author disclosures are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25158.

Address correspondence via email to Changsheng Ma, MD, at chshma@vip.sina.com and Zhe Zhao, MD, at zza00076@btch.edu.cn.

Submitted for publication February 2, 2023; accepted in revised form May 16, 2023.

SIGNIFICANCE & INNOVATIONS

- The Global Burden of Disease Study 2010 and 2017 reported heavy disease burden of osteoarthritis (OA) in the past decades; however, no studies have systematically analyzed the global burden of knee OA.
- This study reported the prevalence, incidence, years lived with disability, age and sex patterns, and related risk factors for knee OA at the global, regional, and national levels.
- Continuous burden monitoring and early health care planning can help mitigate the impacts of knee OA.

MATERIALS AND METHODS

Overview. The GBD is a tool administered by the Institute for Health Metrics and Evaluation (IHME) to analyze epidemiologic data worldwide since 1990. The GBD study estimates prevalence, incidence, mortality, YLDs, years of life lost, and disability-adjusted life years (DALYs) for each country and territory by age and sex. The GBD 2019 studied 369 diseases and injuries, rate of change, and 87 attributable risk factors in 204 countries and territories. Detailed information and methodology for GBD 2019 have been reported (8,9), and all GBD 2019 data are available online (https://vizhub.healthdata.org/gbd-compare/ and https://vizhub.healthdata.org/gbd-results/). The GBD study complies with the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) statement (10). The University of Washington Institutional Review Board approved GBD 2019 and waived informed consent because of its use of deidentified data.

Case definition. In the GBD 2019 study, knee OA is defined as symptomatic OA around the knee joint radiologically confirmed as Kellgren/Lawrence grades 2–4 (11). Grade 2 is characterized by 1 clear osteophyte in the knee joint and pain for at least 1 month of the last 12 months. Grades 3–4 are characterized by osteophytes and joint space narrowing in the affected joint, with deformity present for grade 4 and pain for at least 1 month of the last 12 months. The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code for knee OA is M17. The ICD-9 code for OA is 715, without specific codes for various sites.

Data sources and processing. In the GBD 2019, all existing sources for the prevalence and incidence of knee OA used in the GBD 2017 study were re-reviewed, and a broad systematic review was performed to gather more input data in 2019 for data on the knee joint. Detailed search strategies for prevalence and incidence have been described elsewhere (8). Exclusion criteria were as follows: 1) subpopulations not representative of the national population; 2) not a population-based study; 3) low sample size (<150); and 4) review rather than an original study. Details of the data search strategy were described in the previous literature (8). For knee OA, the GBD group identified studies reporting based on radiographs only, self-reported OA with pain, and self-reported OA with no information on pain. The analysis did not include hospital inpatient data because it may not represent true prevalence. Results of the systematic review showed that the numbers of countries reporting knee OA prevalence and incidence estimates were 26 and 4, respectively.

Modeling. Estimates of incidence and prevalence of knee OA were generated using the software DisMod-MR 2.1, a Bayesian mixed-effects meta-regression tool. DisMod-MR 2.1 imposes coherence between heterogeneous epidemiologic data to make consistent prevalence and incidence estimates. In the GBD 2019, the modeling strategy for knee OA was adjusted on remission. Bounds were set between 0 and 0.05 to account for knee replacement compared with the GBD 2017. The disability weight (DW) assessments of knee OA were based on describing significant functional consequences and symptoms. Graded descriptions and disability weights for OA severity were based on the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), with scores 0–5 taken as mild (DW 0.023 [95% uncertainty interval (95% UI) 0.013 to 0.037]), 6–13 as moderate (DW 0.079 [95% UI 0.054 to 0.110]), and ≥14 as severe (DW 0.165 [95% UI 0.112 to 0.232]) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158). The proportion of each severity of knee OA was determined from 5 studies, including 4 studies representing the high-income South Asia, Southeast Asia, East Asia, and Oceania super regions defined by the GBD, and 1 from the US Osteoarthritis Initiative study (8,12). The pooled percentages were 47.0% (95% UI 42.2 to 51.9) for mild, 35.9% (95% UI 31.3 to 40.7) for moderate, and 17.1% (95% UI 12.9 to 21.6) for severe for low- and middle-income countries, and 74.3% (95% UI 64.8 to 82.7) for mild, 24.3% (95% UI 16.4 to 33.1) for moderate, and 1.1% (95% UI 0.6 to 1.7) for severe for high-income countries.

DALYs are a standard measure used to quantify disease burden. Because no deaths were assumed to be attributable to knee OA in the GBD-estimated cause-of-death model, DALYs for OA were considered the same as YLDs. YLDs were obtained by multiplying the prevalence of each severity by the corresponding DW. Uncertainty was incorporated by sampling 1,000 draws throughout the modeling process. Estimates were summarized as means of draws and 95% UIs (2.5th and 97.5th percentiles of ordered draws). In the GBD 2019 study, age-standardized rates for prevalence, incidence, and YLDs were calculated based on the GBD world population age standard (13). The SDI is a comprehensive index that composites average rankings of the incomes per capita, average educational attainment, and fertility rates to measure the placement of countries or territories in the

development process, assigning a score ranging from 0 to 1 (8,14). The expected relation between SDI and YLD rates was determined using the smoothing splines model of estimates for all regions from 1990 to 2019. For adults >20 years, the theoretical minimum risk exposure level for BMI (20–25 kg/m^2) was determined based on the BMI level associated with the lowest risk of all-cause death in previous studies (15,16). Therefore, high BMI in the GBD 2019 study was defined as a BMI >25kg/m^2 (9).

RESULTS

Overall burden. In 2019, the prevalence of knee OA was ~364.6 million (95% UI 315.3 to 417.4) worldwide, accounting for 4.9% (95% UI 4.2 to 5.6) of all causes of disease in the global population. The age-standardized prevalence estimate was 4,376.0 per 100,000 (95% UI 3,793.0 to 5,004.9), with an increase of 7.5% between 1990 and 2019. Moreover, the incident cases of knee OA in 2019 were ~29.5 million (95% UI 25.6 to 33.7), with an age-standardized incidence rate of 350.3 per 100,000 (95% UI 303.4 to 398.9) and a 6.2% (95% UI 5.6 to 6.7) increase from 1990 to 2019 (Table 1).

Among the 21 GBD regions, high-income Asia Pacific, East Asia, and Australasia had the highest prevalence, with age-standardized rates of 5,662.9 (95% UI 4,898.9 to 6,475.1), 5,123.3 (95% UI 4,386.3 to 5,919.0), and 4,903.7 (95% UI 4,250.3 to 5,651.3) per 100,000, respectively. Central Asia, Southeast Asia, and Central Europe had the lowest estimated prevalence, with age-standardized rates of 2,908.0 (95% UI 2,489.3 to 3,370.5), 3,317.7 (95% UI 2,851.9 to 3,827.5), and 3,413.7 (95% UI 2,929.0 to 3,938.7) per 100,000, respectively. Andean Latin America, southern Latin America, and Southeast Asia showed the largest increases in age-standardized prevalence between 1990 and 2019 (12.9% [95% UI 9.6 to 16.6], 12.7% [95% UI 9.6 to 16.8], and 12.4% [95% UI 10.8 to 14.1], respectively). In contrast, high-income North America, high-income Asia Pacific, and Central Sub-Saharan Africa showed the lowest increases (–1.8% [95% UI –5.5 to 2.1], 2.6% [95% UI 1.3 to 3.9], and 4.3% [95% UI 1.1 to 7.4], respectively) over that period (Table 1).

The highest age-standardized incidence of knee OA was in high-income Asia Pacific (449.1 [95% UI 391.6 to 511.2]), east Asia (401.9 [95% UI 347.0 to 459.6]), and Australasia (394.7 [95% UI 343.1 to 453.5]). Central Asia (248.8 [95% UI 215.0 to 287.7]), Southeast Asia (273.6 [95% UI 237.6 to 313.0]), and Central Europe (288.1 [95% UI 250.0 to 331.0]) had the lowest age-standardized incidence. Andean Latin America, Southeast Asia, and Australasia showed the largest increases in age-standardized incidence (11.1% [95% UI 8.1 to 14.5], 10.5% [95% UI 9.2 to 11.9], and 10.5% [95% UI 6.5 to 13.8], respectively). In contrast, high-income North America, high-income Asia Pacific, and Central Sub-Saharan Africa showed the lowest increases (0.3% [95% UI –3.0 to 3.9], 1.8% [95% UI 0.4 to 3.3], and 3.7% [95% UI 0.7 to 6.6], respectively) between 1990 and 2019 (Table 1).

In GBD 2019, China, India, the US, and Japan had the largest numbers of cases of knee OA, which were ~108.1 million (95% UI 91.9 to 125.8 million), 46.9 million (95% UI 40.5 to 53.7 million), 24.7 (95% UI 21.6 to 28.0 million), and 16.1 million (95% UI 14.0 to 18.3), respectively (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158). The countries with the highest age-standardized prevalence were the Republic of Korea, Brunei Darussalam, Singapore, and Japan, which were 6,211.1 (95% UI 5,360.4 to 7,095.2), 5,954.0 (95% UI 5,127.2 to 6,827.8), 5,850.4 (95% UI 5,071.7 to 6,682.9), and 5,462.2 (95% UI 4,726.4 to 6,267.4) per 100,000, respectively (Figure 1 and Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158). The countries with the lowest age-standardized prevalence per 100,000 were Tajikistan (2,624.6 [95% UI 2,240.2 to 3,037.6]), Canada (2,752.8 [95% UI 2,356.6 to 3,186.7]), and Mongolia (2,790.7 [95% UI 2,392.2 to 3,238.5]). Oman, Thailand, and Equatorial Guinea showed the most significant change in the age-standardized prevalence rate from 1990 to 2019, which were 18.8% (95% UI 14.0% to 23.7%), 18.6% (95% UI 12.8% to 25.3%), and 17.2% (95% UI 12.2% to 22.3%), respectively. The US showed negative growth in age-standardized prevalence, at –2.1% (95% UI –6.0 to 2.0) from 1990 to 2019. The 2 other lowest-growth countries over that period were the Republic of Korea and Japan, which were 0.6% (95% UI –3.3 to 5.0) and 1.1% (95% UI 0.1 to 2.1), respectively. Incidence trends among countries were similar to prevalence results (see Supplementary Figure 1 and Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158).

Knee OA–related YLDs. In 2019 globally, the number of YLDs as a result of knee OA was 11.5 million (95% UI 5.7 to 23.5), accounting for 1.33% (95% UI 0.76 to 2.51) of all-cause YLDs in the population, with an age-standardized rate of 138.2 (95% UI 68.5 to 281.3) per 100,000 population. The percentage change in the age-standardized rate of YLDs increased by 7.8% (95% UI 7.1 to 8.4) between 1990 and 2019.

Regionally, high-income Asia Pacific (180.9 [95% UI 89.6 to 367.5]), East Asia (163.7 [95% UI 80.1 to 334.0]), and Australasia (154.6 [95% UI 76.8 to 315.0]) had the highest age-standardized rates of YLDs, while Central Asia (92.2 [95% UI 45.9 to 189.0]), Southeast Asia (105.2 [95% UI 51.8 to 215.0]) and Central Europe (107.8 [95% UI 53.3 to 219.2]) showed the lowest rates in 2019. Andean Latin America (12.9% [95% UI 6.7 to 9.3]), Southeast Asia (12.8% [95% UI 11.1 to 14.6]), and southern Latin America (12.6% [95% UI 8.9 to 17.1]) showed the largest percentage change in age-standardized YLDs between 1990 and 2019. High-income North America (–2.3% [95% UI –5.9 to 1.8]), high-income Asia Pacific (3.3% [95% UI 1.8 to 4.8]), and Central

Table 1. Absolute number and age-standardized rates (ASRs) per 100,000 population of prevalence, incidence, and years lived with disability (YLDs) due to knee osteoarthritis in 2019, and percentage change between 1990 and 2019 by Global Burden of Disease regions (data available at http://ghdx.healthdata.org/gbd-results-tool)*

	Prevalence			Incidence			YLDs		
Region	Absolute number (95% UI) 2019	ASR per 100,000 (95% UI) 2019	% change in ASR (95% UI) 1990–2019	Absolute number (95% UI) 2019	ASR per 100,000 (95% UI) 2019	% change in ASR (95% UI) 1990–2019	Absolute number (95% UI) 2019	ASR per 100,000 (95% UI) 2019	% change in ASR (95% UI) 1990–2019
Global	364,577,083 (315,252,454 to 417,398,451)	4,376.0 (3,793.0 to 5,004.9)	7.5 (6.8 to 8.1)	29,510,498 (25,560,399 to 33,675,194)	350.3 (303.4 to 398.9)	6.2 (5.6 to 6.7)	11,534,020 (5,719,115 to 23,489,983)	138.2 (68.5 to 281.3)	7.8 (7.1 to 8.4)
East Asia	111,798,402 (95,096,454 to 130,035,875)	5,123.3 (4,386.3 to 5,919.9)	7.0 (5.4 to 8.4)	8,700,203 (7,469,677 to 9,989,325)	401.9 (347.0 to 459.6)	6.8 (5.5 to 8.0)	3,582,773 (1,746,375 to 7,306,509)	163.7 (80.1 to 334.0)	7.3 (5.7 to 8.9)
Southeast Asia	21,653,751 (18,544,761 to 25,027,693)	3,317.7 (2,851.9 to 3,827.5)	12.4 (10.8 to 14.1)	1,942,570 (1,677,453 to 2,233,401)	273.6 (237.6 to 313.2)	10.5 (9.2 to 11.9)	691,559 (340,907 to 1,418,927)	105.2 (51.8 to 215.0)	12.8 (11.1 to 14.6)
Oceania	317,967 (271,127 to 369,641)	4,039.0 (3,484.8 to 4,666.6)	8.0 (4.5 to 11.5)	31,052 (26,623 to 35,643)	324.9 (281.9 to 370.2)	6.7 (3.4 to 10.1)	10,095 (4,975 to 20,597)	126.7 (62.2 to 258.0)	7.6 (3.7 to 11.5)
Central Asia	2,259,310 (1,921,936 to 2,639,839)	2,908.0 (2,489.4 to 3,370.5)	6.6 (4.9 to 8.6)	217,115 (185,751 to 252,184)	248.8 (215.4 to 287.7)	5.9 (4.3 to 7.7)	72,178 (35,319 to 149,261)	92.2 (45.9 to 189.0)	6.5 (4.3 to 8.7)
Central Europe	6,908,689 (5,907,629 to 7,954,696)	3,413.7 (2,929.0 to 3,938.7)	7.6 (6.3 to 9.0)	520,170 (450,485 to 598,001)	288.1 (250.0 to 331.0)	6.8 (5.6 to 8.2)	217,062 (107,174 to 439,737)	107.8 (53.3 to 219.2)	8.0 (6.5 to 9.6)
Eastern Europe	12,234,181 (10,458,732 to 14,173,527)	3,604.5 (3,094.1 to 4,162.3)	7.6 (6.3 to 9.0)	963,188 (826,724 to 1,114,823)	305.4 (263.9 to 351.6)	6.8 (5.7 to 8.0)	384,232 (190,305 to 782,855)	113.5 (56.2 to 231.2)	8.3 (6.8 to 9.9)
High-income Asia Pacific	22,234,484 (19,302,190 to 25,267,377)	5,662.9 (4,898.9 to 6,475.1)	2.6 (1.3 to 3.9)	1,419,475 (1,246,517 to 1,615,144)	449.1 (391.6 to 511.2)	1.8 (0.4 to 3.3)	703,455 (351,319 to 1,417,539)	180.9 (89.6 to 367.5)	3.3 (1.8 to 4.8)
Australasia	2,252,824 (1,946,216 to 2,577,695)	4,903.7 (4,250.3 to 5,651.3)	12.3 (7.9 to 16.2)	160,589 (139,816 to 185,025)	394.7 (343.1 to 453.5)	10.5 (6.5 to 13.8)	70,621 (35,440 to 143,155)	154.6 (76.8 to 315.0)	12.5 (8.1 to 16.8)
Western Europe	34,767,067 (29,990,024 to 39,760,414)	4,255.3 (3,663.2 to 4,893.3)	6.1 (4.7 to 7.4)	2,425,687 (2,110,509 to 2,770,405)	349.7 (303.9 to 401.0)	5.0 (3.7 to 6.3)	1,090,508 (542,397 to 2,200,480)	134.5 (66.6 to 273.4)	6.3 (4.7 to 7.8)
Southern Latin America	3,866,021 (3,335,978 to 4,436,394)	4,734.4 (4,087.4 to 5,435.2)	12.7 (9.6 to 16.8)	298,062 (258,057 to 341,518)	380.8 (330.0 to 435.4)	9.9 (6.5 to 13.6)	122,391 (60,644 to 249,056)	150.1 (74.3 to 306.1)	12.6 (8.9 to 17.1)
High-income North America	26,500,088 (23,105,047 to 30,038,720)	4,494.7 (3,921.8 to 5,097.7)	−1.8 (−5.5 to 2.1)	1,934,225 (1,694,202 to 2,208,746)	367.6 (323.4 to 417.3)	0.3 (−3.0 to 3.9)	814,571 (408,934 to 1,640,190)	139.0 (69.6 to 282.7)	−2.3 (−5.9 to 1.8)
Caribbean	2,387,604 (2,052,648 to 2,745,394)	4,588.6 (3,944.2 to 5,274.4)	9.4 (7.0 to 11.8)	190,960 (165,413 to 218,987)	368.6 (319.9 to 422.9)	7.5 (5.6 to 9.7)	75,572 (37,433 to 154,216)	145.2 (71.9 to 296.5)	9.0 (6.3 to 11.6)

(Continued)

Table 1. (*Cont'd*)

Region	Prevalence			Incidence			YLDs		
	Absolute number (95% UI) 2019	ASR per 100,000 (95% UI) 2019	% change in ASR (95% UI) 1990–2019	Absolute number (95% UI) 2019	ASR per 100,000 (95% UI) 2019	% change in ASR (95% UI) 1990–2019	Absolute number (95% UI) 2019	ASR per 100,000 (95% UI) 2019	% change in ASR (95% UI) 1990–2019
Andean Latin America	2,640,755 (2,291,812 to 3,016,238)	4,644.5 (4,019.0 to 5,299.0)	12.9 (9.6 to 16.6)	222,815 (194,170 to 252,635)	375.0 (326.3 to 425.1)	11.1 (8.1 to 14.5)	84,039 (41,499 to 170,311)	147.5 (73.5 to 299.1)	12.9 (9.3 to 16.7)
Central Latin America	11,066,354 (9,539,183 to 12,746,973)	4,597.5 (3,972.8 to 5,286.1)	8.9 (7.8 to 10.2)	942,508 (816,894 to 1,081,914)	375.8 (326.3 to 429.6)	7.9 (6.7 to 9.1)	349,776 (172,205 to 710,492)	145.0 (71.5 to 293.7)	9.3 (8.0 to 10.8)
Tropical Latin America	11,269,828 (9,697,657 to 13,019,198)	455.9 (3,929.1 to 5,257.4)	10.7 (9.1 to 12.5)	953,081 (823,458 to 1,092,840)	374.0 (324.3 to 428.2)	9.3 (7.6 to 10.8)	356,340 (175,504 to 724,446)	143.8 (70.9 to 293.2)	11.6 (9.7 to 13.5)
North Africa and Middle East	17,751,396 (15,270,039 to 20,418,318)	3,863.7 (3,321.6 to 4,441.9)	11.4 (9.8 to 13.3)	1,694,582 (1,461,455 to 1,945,115)	320.2 (276.7 to 365.6)	9.8 (8.3 to 11.3)	564,658 (275,059 to 1,150,685)	122.0 (60.2 to 247.3)	11.3 (9.4 to 13.2)
South Asia	56,437,708 (48,780,278 to 64,865,938)	3,871.8 (3,341.3 to 4,441.9)	11.4 (10.3 to 12.4)	5,075,834 (4,404,911 to 5,836,161)	320.3 (278.3 to 366.1)	9.5 (8.7 to 10.4)	1,763,550 (862,365 to 3,599,801)	120.2 (59.0 to 244.9)	11.8 (10.6 to 13.0)
Central Sub-Saharan Africa	2,004,462 (1,706,971 to 2,334,843)	3,508.8 (3,019.4 to 4,073.5)	4.3 (1.1 to 7.4)	204,690 (174,686 to 237,292)	293.3 (252.1 to 338.1)	3.7 (0.7 to 6.6)	63,765 (31,213 to 130,340)	110.6 (54.9 to 225.8)	4.9 (1.2 to 8.3)
Eastern Sub-Saharan Africa	6,046,975 (5,171,018 to 6,960,361)	3,502.2 (3,013.8 to 4,039.8)	6.3 (5.0 to 7.5)	614,556 (529,086 to 708,678)	294.4 (254.3 to 337.5)	5.5 (4.4 to 6.7)	192,735 (95,179 to 395,042)	110.7 (55.0 to 225.2)	6.8 (5.3 to 8.2)
Southern Sub-Saharan Africa	2,342,562 (2,007,807 to 2,706,954)	4,001.0 (3,440.2 to 4,618.4)	7.1 (5.4 to 8.7)	217,434 (187,486 to 250,574)	333.8 (287.7 to 383.3)	6.4 (4.8 to 8.0)	73,661 (36,428 to 150,438)	125.2 (62.0 to 253.9)	6.5 (4.7 to 8.2)
Western Sub-Saharan Africa	7,836,655 (6,677,243 to 9,029,951)	3,877.7 (3,329.8 to 4,468.4)	7.1 (6.3 to 8.1)	781,700 (672,894 to 900,303)	321.7 (277.1 to 369.6)	6.5 (5.8 to 7.4)	250,479 (123,436 to 510,594)	122.9 (60.8 to 248.9)	7.5 (6.5 to 8.7)

* 95% UI = 95% uncertainty interval.

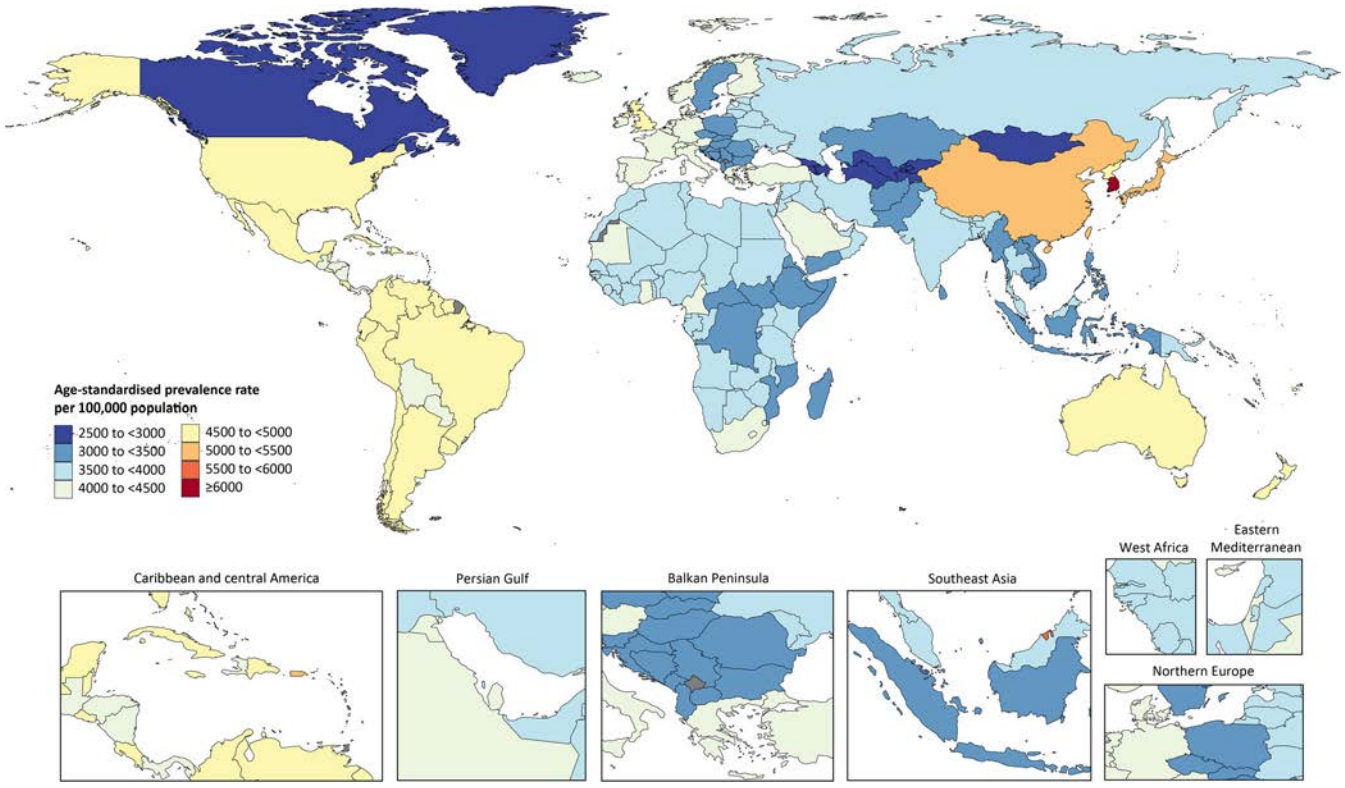


Figure 1. Global map of age-standardized prevalence rate for knee osteoarthritis in 2019 (data available at http://ghdx.healthdata.org/gbd-results-tool).

Sub-Saharan Africa (4.9% [95% UI 1.2 to 8.3]) showed the lowest trends, including decreases, in the percentage change in age-standardized YLDs (Table 1).

The Republic of Korea, Brunei Darussalam, Singapore, and Japan showed the largest age-standardized YLD rates, which were 198.0 (95% UI 98.6 to 402.8), 187.7 (95% UI 92.2 to 382.5), 187.1 (95% UI 92.6 to 378.6), and 174.6 (95% UI 86.3 to 353.5), respectively, per 100,000 (see Supplementary Figure 2 and Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158). Tajikistan, Canada, and Mongolia had the lowest age-standardized YLD rates, which were 83.5 (95% UI 40.7 to 171.6), 87.2 (95% UI 43.2 to 176.5), and 88.8 (95% UI 43.7 to 181.3) per 100,000. Thailand, Oman, and Equatorial Guinea had the largest changes in YLD rates from 1990 to 2019, which were 19.7% (95% UI 12.8 to 26.8), 18.8% (95% UI 13.1 to 24.7), and 18.3% (95% UI 12.8 to 24.3), respectively. In addition, the US, Burundi, and the Republic of Korea showed the lowest changes in YLD rates from 1990 to 2019, which were –2.7% (95% UI –6.6 to 1.7), 1.4% (95% UI –3.6 to 6.9), and 1.6% (95% UI –2.8 to 6.4) respectively (see Supplementary Figure 3 and Supplementary Table 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158).

Burden by SDI. In 2019, the prevalence rate attributable to knee OA was higher in the high SDI and high–middle SDI quintiles than in the middle, low-middle, and low SDI quintiles (Figure 2). The high–middle SDI quintiles showed the most notable changes, with age-standardized prevalence rates increasing by 11.9% (95% UI 10.5 to 13.3) between 1990 and 2019. The high SDI quintile showed the lowest change in age-standardized prevalence rates among quintiles, at 4.3% (95% UI 2.9 to 5.7) between 1990 and 2019. The incidence and YLD rates showed similar trends in prevalence between 1990 and 2019 (see Supplementary Figures 3 and 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158).

Generally, the age-standardized YLDs of knee OA increased gradually with increasing SDI between 1990 and 2019 (Figure 3). Between 1990 and 2019, high-income Asia Pacific and East Asia showed much higher age-standardized YLD rates than expected based on SDI level, although these trends have declined in recent years. Although SDI increased between 1990 and 2019, Central Asia, Southeast Asia, Central Europe, and Eastern Europe showed much lower YLDs than other regions. Age-standardized YLD rates for knee OA fluctuated in high-income North America between 1990 and 2019, with an overall downward trend (Figure 3).

Age and sex patterns. Globally in 2019, knee OA prevalence increased with age, with a peak in both sexes within the age range of 55–69 years. The incidence of knee OA also increased with age, peaking in both sexes at ages 45–59 years.

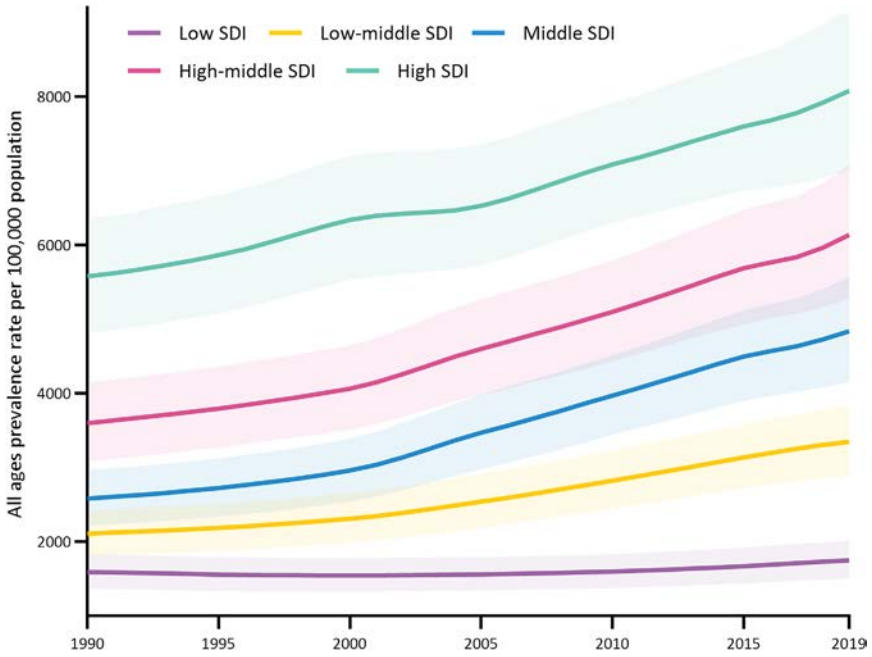


Figure 2. All ages prevalence per 100,000 population by 5 common sociodemographic index (SDI) quintiles for knee osteoarthritis, 1990–2019. Shaded areas show 95% uncertainty intervals (data available at http://ghdx.healthdata.org/gbd-results-tool). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158/abstract.

Female patients showed higher prevalence estimates than male patients across all age groups (Figure 4). The global age-standardized prevalence rate was 5,161.4 (95% UI 4,470.9 to 5,889.7) per 100,000 population in female patients and 3,510.2 (95% UI 3,032.6 to 4,037.0) in male patients, and the female-to-male ratio was 1.47. East Asia, high-income Asia Pacific, and

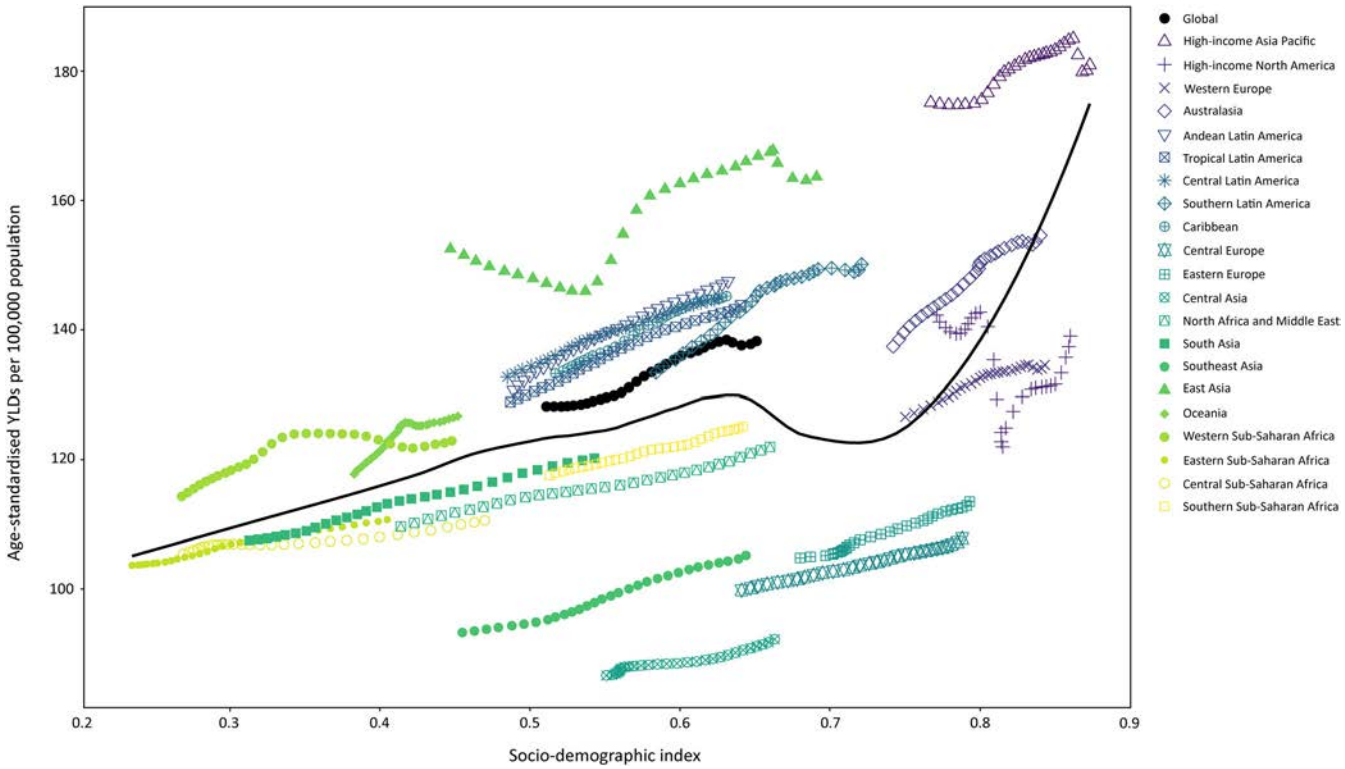


Figure 3. Age-standardized rate of years lived with disability (YLDs) for knee osteoarthritis for 21 world regions, 1990–2019. Solid black line indicates expected values based on Socio-demographic Index and YLD rates of all regions (data available at http://ghdx.healthdata.org/gbd-results-tool). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158/abstract.

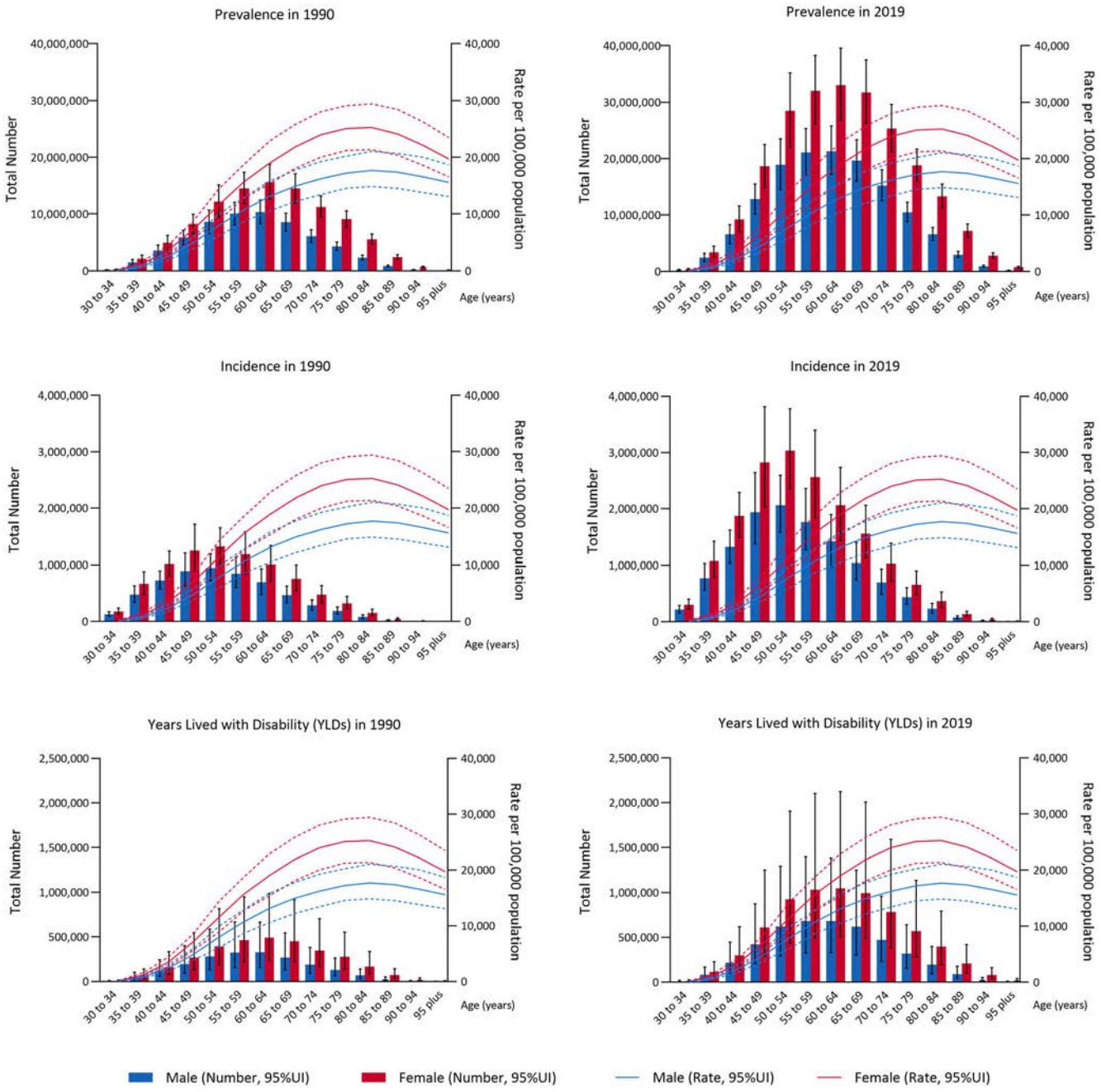


Figure 4. Global number of prevalent and incident cases, years lived with disability (YLDs), and estimates per 100,000 population of osteoarthritis by age and sex, 1990 and 2019. Broken lines represent the upper and lower 95% uncertainty intervals (95% UIs) (data available at http://ghdx.healthdata.org/gbd-results-tool). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158/abstract.

Oceania had the largest female-to-male ratios of all regions at 1.72, 1.69, and 1.55, respectively. The Republic of Korea, Brunei Darussalam, China, and Singapore had the highest female-to-male ratios at 1.77, 1.72, 1.72, and 1.72, respectively. Globally, the female-to-male ratios in age-standardized incidence rates and YLDs were 1.22 and 1.46, respectively (see Supplementary Tables 2–4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158). The age-standardized incidence and YLDs among regions and countries showed trends similar to those of prevalence (see Supplementary Tables 3 and 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158).

Risk factors. Globally in 2019, 22.4% (95% UI 12.1 to 34.2) of YLDs resulting from knee OA were attributable to high BMI, an increase of 40.5% between 1990 and 2019. Regionally, high-income North America, Eastern Europe, and Central Europe showed the highest percentages of age-standardized YLDs due to high BMI, which were 37.6% (95% UI 21.8 to 53.3), 35.7%

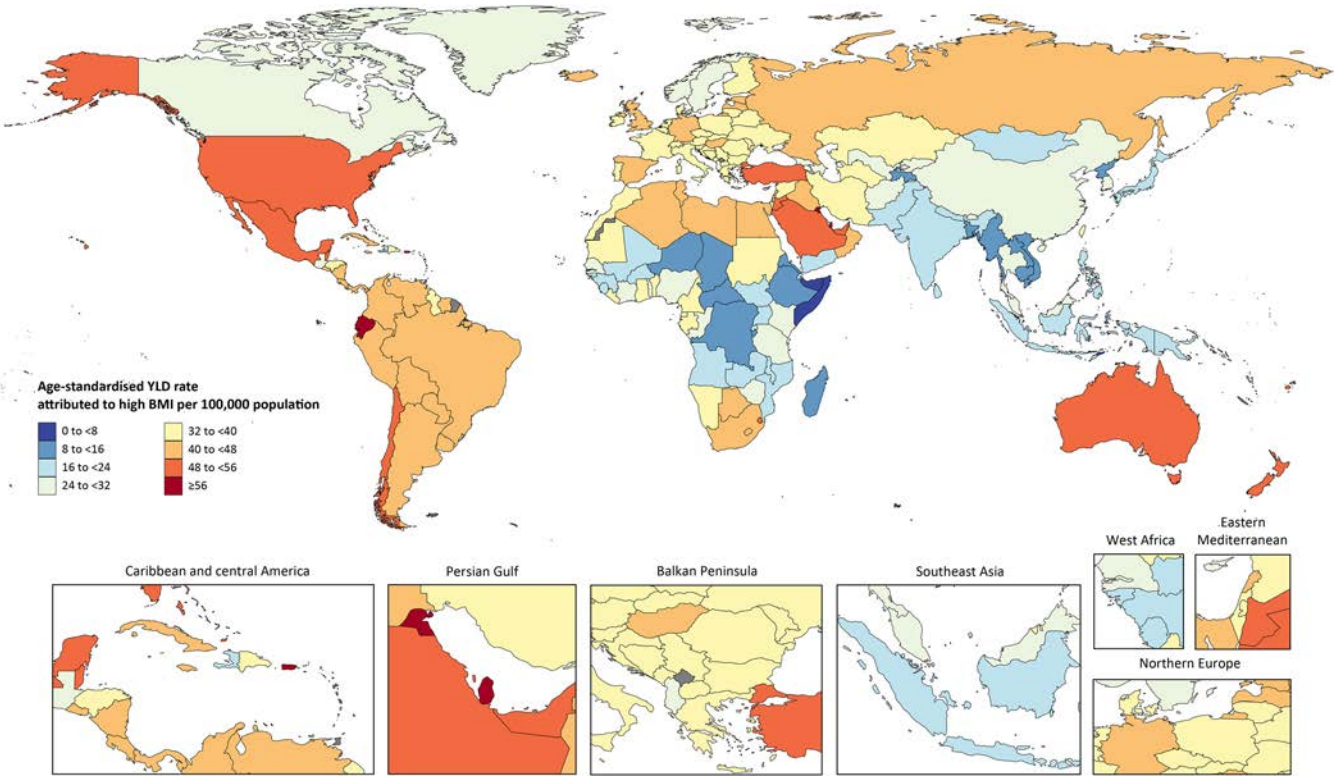


Figure 5. Global map of age-standardized years lived with disability (YLDs) rates for knee osteoarthritis attributable to high body mass index (BMI) for both sexes in 2019 (data available at http://ghdx.healthdata.org/gbd-results-tool). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25158/abstract.

(95% UI 20.8 to 51.2), and 35.7% (95% UI 21.1 to 51.1), respectively. Nationally, Qatar, the United Arab Emirates, and Kuwait showed the highest percentages of age-standardized YLD attributed to high BMI, which were 46.0% (95% UI 28.4 to 61.9), 45.2% (95% UI 28.7 to 60.4), and 44.4% (95% UI 28.0 to 59.4), respectively (Figure 5).

DISCUSSION

This study reports the updated data of prevalent cases, incident cases, YLDs, and associated age-standardized rates of knee OA in 204 countries and territories from 1990 to 2019. Globally, there were >364.6 million prevalent cases of knee OA and 29.5 million incident cases in 2019, resulting in ~11.5 million YLDs. These updated data reveal that knee OA affects a large number of patients worldwide and has become a serious public health condition threatening human health and mobility.

The GBD 2017 study reported ~303 million cases of hip OA and knee OA worldwide, with an age-standardized prevalence estimate of 3,754 per 100,000 (17,18). Long et al reported the prevalence cases of OA categorized by site from 1990 to 2019, with results consistent with our findings (3). In the current study, the worldwide prevalence, incidence, YLDs, sex differences, and risk factors of knee OA have been comprehensively analyzed (19,20). A recent systematic review estimated that the global knee OA prevalence exceeds 22% in the population >40 years of age, with an overall incidence of 203 per 100,000 person-years >20 years of age (6). The GBD study defined symptomatic knee pain and imaging-diagnosed or self-reported OA in the collection of knee OA cases. Although previous studies were not fully comparable to the 2019 GBD study in terms of methodology and the number of countries collected, these studies, together with the present study, confirm the increasing trends of knee OA in both prevalence rate and absolute number (9).

Regionally, high-income Asia Pacific, East Asia, and Australasia have the highest age-standardized prevalence rates of knee OA. In 2019, China, the US, and India had the largest numbers of knee OA patients, with >108.1 million, 46.9 million, and 24.7 million, respectively. These countries also have the highest total medical costs resulting from knee OA. In 2019, >370,000 patients underwent knee replacement surgery in China, increasing by ~20% annually (21). In the US, >700,000 total knee replacements are performed each year, >95% of which are due to knee OA, with direct medical costs exceeding $15,000 per patient (2,22,23). Although the severity of knee OA may progress with patient age, strategies for OA management among younger patients, such as the use of pain relievers, neuromuscular training, and maintenance of healthy BMI and physical activity, as a means of managing symptoms and potentially delaying disease progression and primary knee replacement merit consideration by health

care providers (5,24,25). In addition to direct medical costs, the consequences of absenteeism and activity limitations caused by severe knee OA are substantial but often neglected (22). While some may assume that no death is directly attributable to knee OA, previous studies have found that all-cause mortality in knee OA patients is ~20–50% higher than in those without knee OA, which may be partly attributable to reduced physical activity and comorbid conditions (26,27).

Knee OA is one of the fastest growing noncommunicable diseases, after type 2 diabetes mellitus and opioid use disorders in the causes of YLDs, with a 7.8% increase annually between 1990 and 2019 (8,28). The high-income Asia Pacific, East Asia, and Australasia regions had the highest YLDs due to knee OA. Andean Latin America, Southeast Asia, and southern Latin America showed the highest increase rates of YLD between 1990 and 2019. Although high-income North America and high-income Asia Pacific exhibit higher YLD rates, the percentage change of the YLDs was the lowest among all regions and has been trending downward in recent years after a rapid increase. These changes in YLDs in recent years may be attributable to the implementation of public health strategies aimed at prevention and management of knee OA in these regions and the increasing public demand for exercise control and pain management of OA (29–32).

The present study found that the age-standardized prevalence rate was nearly 1.5 times higher in female than in male patients, with the ratio in Asia higher still, at ~1.7. The sex discrepancy for the knee OA burden was consistent with that reported by other population-based studies (6,33). In addition, female patients have been reported to have more severe symptoms and lower preoperative functional scores after admission, indicating that female patients may be at a later stage of the disease than male patients when receiving interventions (34,35). Encouraging earlier introduction of knee OA strategies in women patients may delay disease progression and achieve more pronounced benefits (24,36).

The results of this study also show that the disease burden of knee OA generally increases along with sociodemographic level. From 1990 to 2019, the high–middle SDI region had the highest increase in age-standardized YLD rates at 12.5%. The disease burden in high and high–middle SDI regions is higher than in the other quintiles, which may be associated with the increasing life expectancy in these regions.

The age pattern in this study was consistent with that in previous findings (28). The peak of incident cases of knee OA in 2019 appears in the age range between 45 and 59 years, earlier than the prevalence peak between 55 and 69 years. As age increases, the prevalence and incidence rate per 100,000 population gradually increases, peaking at 80–84 years. This study provided more age categories of the knee OA population, and the results may indicate the potential cost-effectiveness of early screening programs and health education for specific age groups in high-burden regions (25,37).

Identifying and reducing risk factors are effective strategies for prevention. Obesity is considered to be the primary risk factor for knee OA (38,39). In 2019 globally, high BMI accounted for ~22% of the YLDs attributable to knee OA. High BMI showed the most significant impact on YLDs in high-income North America, Eastern Europe, and Central Europe, exceeding 30% in these regions, indicating the importance of obesity control programs. Compared with those regions, the contribution of high BMI to the YLD burden in high-income Asia Pacific and East Asia was relatively low. Other factors, such as characteristic life habits or occupational risks, may need to be assessed to identify additional risk factors (40).

This study had several limitations. First, the disease burden estimates provided by the GBD study were derived from the DISMOD-MR 2.1 model. The original input data for calculating prevalence and incidence rates for the 204 countries and territories were based on data collected from a few countries, not all countries, as described in the Methods section. In addition, the data from these countries do not cover all periods and regions (41). Therefore, researchers should be aware of the incompleteness of included data when applying the findings. Second, YLDs were based on the estimated distribution of severity and disability weights for the knee OA patient population. The GBD 2019 study set a 5% remission rate for severe knee OA patients undergoing joint replacements. Therefore, calculation of YLDs caused by knee OA may have been affected by differing research methodologies. Third, in the GBD study, high BMI was identified as a significant contributor to YLDs in knee OA. Potential risk factors for knee OA may include physical work, traumatic knee injuries, movement patterns, and nutrition (38,42,43). However, no quantitative tool is currently available to assess these factors in the GBD study. The impact of these factors should be considered when collecting basic information and developing prevention strategies.

In conclusion, knee OA is a highly prevalent condition worldwide. The prevalence, incidence, YLDs, and age-standardized rates of knee OA increased substantially in most countries and regions from 1990 to 2019. The burden of knee OA appears to rise with increasing SDI and is higher in female than in male patients. High BMI is a significant risk factor for knee OA. Continuous monitoring of the burden of knee OA is of great significance for establishing appropriate public prevention policies and raising public awareness, especially in high- and high–middle SDI regions.

ACKNOWLEDGMENTS

We thank the IHME for their excellent work on the data visualization and statistics.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Yang, Wang, Ma, Zhao.
Acquisition of data. Wang, Liu, Lu, He.
Analysis and interpretation of data. Yang, Wang, He.

REFERENCES

1. Sharma L. Osteoarthritis of the knee. N Engl J Med 2021;384:51–9.
2. Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis: a review [review]. JAMA 2021;325:568–78.
3. Long H, Liu Q, Yin H, et al. Prevalence trends of site-specific osteoarthritis from 1990 to 2019: findings from the Global Burden of Disease Study 2019. Arthritis Rheumatol 2022;74:1172–83.
4. Arslan IG, Damen J, de Wilde M, et al. Incidence and prevalence of knee osteoarthritis using codified and narrative data from electronic health records: a population-based study. Arthritis Care Res (Hoboken) 2022;74:937–44.
5. Hurley M, Dickson K, Hallett R, et al. Exercise interventions and patient beliefs for people with hip, knee or hip and knee osteoarthritis: a mixed methods review [review]. Cochrane Database Syst Rev 2018;4:CD010842.
6. Cui A, Li H, Wang D, et al. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. EClinicalMedicine 2020;29–30:100587.
7. Verschuuren M, Hilderink HB, Vonk RA. The Dutch Public Health Foresight Study 2018: an example of a comprehensive foresight exercise. Eur J Public Health 2020;30:30–5.
8. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020;396:1204–22.
9. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020;396:1223–49.
10. Stevens GA, Alkema L, Black RE, et al. Guidelines for Accurate and Transparent Health Estimates Reporting: the GATHER statement [review]. Lancet 2016;388:e19–23.
11. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
12. Collins NJ, Misra D, Felson DT, et al. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS) [review]. Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S208–28.
13. GBD 2019 Demographics Collaborators. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. Lancet 2020;396:1160–203.
14. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2020;395:709–33.
15. Global BMI Mortality Collaboration, Di Angelantonio E, Bhupathiraju Sh N, et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. Lancet 2016;388:776–86.
16. Afshin A, Forouzanfar MH, Reitsma MB, et al. Health effects of overweight and obesity in 195 countries over 25 years. N Engl J Med 2017;377:13–27.
17. Safiri S, Kolahi AA, Smith E, et al. Global, regional and national burden of osteoarthritis 1990–2017: a systematic analysis of the Global Burden of Disease Study 2017. Ann Rheum Dis 2020;79:819–28.
18. Cross M, Smith E, Hoy D, 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.
19. Nguyen US, Zhang Y, Zhu Y, et al. Increasing prevalence of knee pain and symptomatic knee osteoarthritis: survey and cohort data. Ann Intern Med 2011;155:725–32.
20. Hunter DJ, March L, Chew M. Osteoarthritis in 2020 and beyond: a Lancet Commission. Lancet 2020;396:1711–2.
21. Feng B, Zhu W, Bian YY, et al. China artificial joint annual data report. Chin Med J (Engl) 2020;134:752–3.
22. Losina E, Paltiel AD, Weinstein AM, et al. Lifetime medical costs of knee osteoarthritis management in the United States: impact of extending indications for total knee arthroplasty. Arthritis Care Res (Hoboken) 2015;67:203–15.
23. Price AJ, Alvand A, Troelsen A, et al. Knee replacement [review]. Lancet 2018;392:1672–82.
24. De Vos BC, Landsmeer ML, van Middelkoop M, et al. Long-term effects of a lifestyle intervention and oral glucosamine sulphate in primary care on incident knee OA in overweight women. Rheumatology (Oxford) 2017;56:1326–34.
25. Mazzei DR, Ademola A, Abbott JH, et al. Are education, exercise and diet interventions a cost-effective treatment to manage hip and knee osteoarthritis? A systematic review. Osteoarthritis Cartilage 2021;29:456–70.
26. Hawker GA. Osteoarthritis is a serious disease [review]. Clin Exp Rheumatol 2019;37 Suppl 120:3–6.
27. Osteoarthritis Research Society International. Osteoarthritis: a serious disease, submitted to the U.S Food and Drug Administration. 2016. URL: https://research-information.bris.ac.uk/en/publications/osteoarthritis-a-serious-disease-submitted-to-the-us-food-and-dru.
28. Glyn-Jones S, Palmer AJ, Agricola R, et al. Osteoarthritis [review]. Lancet 2015;386:376–87.
29. Cohen SA, Brophy RH, Chen AF, et al. Public interest in hyaluronic acid injections for knee osteoarthritis in the United States and Europe: an international Google trends analysis. Arthroplast Today 2022;18:157–62.
30. Theis KA, Murphy L, Hootman JM, et al. Social participation restriction among US adults with arthritis: a population-based study using the International Classification of Functioning, Disability and Health. Arthritis Care Res (Hoboken) 2013;65:1059–69.
31. Piercy KL, Troiano RP, Ballard RM, et al. The Physical Activity Guidelines for Americans. JAMA 2018;320:2020–8.
32. Ebata-Kogure N, Murakami A, Nozawa K, et al. Treatment and healthcare cost among patients with hip or knee osteoarthritis: a cross-sectional study using a real-world claims database in Japan between 2013 and 2019. Clin Drug Investig 2020;40:1071–84.
33. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis [review]. Lancet 2019;393:1745–59.
34. Ferre IM, Roof MA, Anoushiravani AA, et al. Understanding the observed sex discrepancy in the prevalence of osteoarthritis [review]. JBJS Rev 2019;7:e8.

35. Parsley BS, Bertolusso R, Harrington M, et al. Influence of gender on age of treatment with TKA and functional outcome. Clin Orthop Relat Res 2010;468:1759–64.
36. Runhaar J, van Middelkoop M, Reijman M, et al. Prevention of knee osteoarthritis in overweight females: the first preventive randomized controlled trial in osteoarthritis. Am J Med 2015;128:888–95.e4.
37. Jin X, Ackerman IN, Ademi Z. Loss of productivity-adjusted life-years in working-age Australians due to knee osteoarthritis: a life-table modeling approach. Arthritis Care Res (Hoboken) 2023;75:482–90.
38. Reyes C, Leyland KM, Peat G, et al. Association between overweight and obesity and risk of clinically diagnosed knee, hip, and hand osteoarthritis: a population-based cohort study. Arthritis Rheumatol 2016; 68:1869–75.
39. Sasaki E, Ota S, Chiba D, et al. Early knee osteoarthritis prevalence is highest among middle-aged adult females with obesity based on new set of diagnostic criteria from a large sample cohort study in the Japanese general population. Knee Surg Sports Traumatol Arthrosc 2020;28:984–94.
40. Newberry SJ, FitzGerald J, SooHoo NF, et al. Treatment of osteoarthritis of the knee: an update review [review]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2017.
41. Maher C, Ferreira G. Time to reconsider what Global Burden of Disease studies really tell us about low back pain. Ann Rheum Dis 2022;81:306–8.
42. Whittaker JL, Losciale JM, Juhl CB, et al. Risk factors for knee osteoarthritis after traumatic knee injury: a systematic review and meta-analysis of randomised controlled trials and cohort studies for the OPTIKNEE Consensus [review]. Br J Sports Med 2022;56: 1406–21.
43. Werner DM, Golightly YM, Tao M, et al. Environmental risk factors for osteoarthritis: the impact on individuals with knee joint injury [review]. Rheum Dis Clin North Am 2022;48:907–30.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2501–2507
DOI 10.1002/acr.25176
© 2023 The Authors. *Arthritis Care & Research* published by Wiley Periodicals LLC on behalf of American College of Rheumatology.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

American College of Rheumatology
Empowering Rheumatology Professionals

Risk of Malnutrition in Patients With Systemic Sclerosis-Associated Interstitial Lung Disease Treated With Nintedanib in the Randomized, Placebo-Controlled SENSCIS Trial

Elizabeth R. Volkmann,[1] Zsuzsanna H. McMahan,[2] Vanessa Smith,[3] Stéphane Jouneau,[4] Corinna Miede,[5] Margarida Alves,[6] and Ariane L. Herrick,[7] on behalf of the SENSCIS Trial Investigators

Objective. To assess adverse events (AEs) in relation to baseline body mass index (BMI) and the risk of malnutrition in patients with systemic sclerosis–associated interstitial lung disease (SSc–ILD) treated with nintedanib.

Methods. Among patients with SSc–ILD randomized to receive nintedanib or placebo in the SENSCIS trial, we assessed AEs in subgroups by baseline BMI ≤20 kg/m^2 and BMI >20 kg/m^2, and the risk of malnutrition using a modified version of the Malnutrition Universal Screening Tool (MUST), over 52 weeks.

Results. The AE profile of nintedanib was similar between subgroups with a baseline BMI ≤20 kg/m^2 (n = 61) and a baseline BMI >20 kg/m^2 (n = 515). In these subgroups, respectively, AEs led to treatment discontinuation in 16.7% and 15.9% of the nintedanib group and 13.5% and 8.0% of the placebo group, respectively. Based on the modified MUST, the proportions of patients who had a low risk of malnutrition at baseline and at their last assessment were 74.0% in the nintedanib group and 78.1% in the placebo group, while the proportions who were classified as at low risk at baseline but at high risk by their last assessment were 4.5% in the nintedanib group and 1.0% in the placebo group.

Conclusion. In the SENSCIS trial, most patients with SSc–ILD remained at low risk of malnutrition over 52 weeks, but the proportion at high risk was higher in patients who received treatment with nintedanib compared to those who received placebo. Management of disease manifestations and AEs that may be associated with weight loss is important to reduce the risk of malnutrition in patients with SSc–ILD.

INTRODUCTION

Systemic sclerosis (SSc) is a complex and heterogeneous autoimmune disease characterized by immune dysregulation and progressive fibrosis of the skin and internal organs (1). Gastrointestinal involvement is common in patients with SSc and can lead to a myriad of symptoms, including reflux, nausea, bloating, diarrhea, and/or constipation (2–5). Among 402 patients with SSc at a UK hospital, 94% reported upper gastrointestinal symptoms and 79% reported lower gastrointestinal symptoms (2). Gastrointestinal complications and increased disease severity are associated with an increased risk of weight loss and malnutrition (6–9). Malnutrition has also been associated with increased mortality in patients with SSc (10–12), but it is unclear to what extent this reflects a direct impact of malnutrition on the risk of death compared to the higher prevalence of malnutrition in patients with greater disease severity.

In addition to the underlying SSc, some of the drugs used to treat SSc or SSc-associated interstitial lung disease (SSc–ILD) are associated with gastrointestinal adverse events (AEs) (13–16). The AE profile of nintedanib, which is licensed for the treatment of SSc–ILD, as well as for idiopathic pulmonary fibrosis (IPF) and progressive fibrosing ILDs of any etiology, is characterized mainly by gastrointestinal AEs, particularly diarrhea (17–19). In the randomized, placebo-controlled SENSCIS trial of nintedanib in patients with SSc–ILD, a greater proportion of patients who received treatment with nintedanib reported diarrhea over

ClinicalTrials.gov identifier: NCT02597933.
Supported by Boehringer Ingelheim International GmbH.
[1]Elizabeth R. Volkmann, MD, MS: University of California, David Geffen School of Medicine, Los Angeles; [2]Zsuzsanna H. McMahan, MD, MHS: Johns Hopkins University, Baltimore, Maryland; [3]Vanessa Smith, MD, PhD: Ghent University Hospital, Ghent University, and VIB Inflammation Research Center, Ghent, Belgium; [4]Stéphane Jouneau, MD, PhD: Competences Centre for Rare Pulmonary Diseases, and Univ Rennes, CHU Rennes, Inserm, EHESP, IRSET, Rennes, France; [5]Corinna Miede, MSc: mainanalytics GmbH, Sulzbach (Taunus), Germany; [6]Margarida Alves, MD: Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany; [7]Ariane L. Herrick, MD, FRCP: The University of Manchester and NIHR Manchester Biomedical Research Centre, Manchester, UK.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25176.

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

Submitted for publication February 23, 2023; accepted in revised form June 20, 2023.

SIGNIFICANCE & INNOVATIONS

- Nintedanib is a licensed treatment for systemic sclerosis–associated interstitial lung disease (SSc–ILD) that may be associated with gastrointestinal adverse events (AEs).
- In patients with SSc–ILD, the AE profile of nintedanib was similar between subgroups by body mass index (BMI) ≤20 kg/m^2 and BMI >20 kg/m^2 at baseline.
- Based on a modified version of the Malnutrition Universal Screening Tool, the proportion of patients classified as having a high risk of malnutrition over 52 weeks was small but was greater in patients who received treatment with nintedanib compared to those who received placebo.
- These findings highlight the importance of managing disease manifestations and gastrointestinal AEs that may be associated with weight loss in patients with SSc–ILD.

52 weeks compared to those who received placebo (76% versus 32%) (15). Most cases of diarrhea were of mild or moderate intensity and did not lead to permanent discontinuation of nintedanib (18). It remains unclear whether treatment with nintedanib is associated with an increased risk of malnutrition in patients with SSc–ILD. Thus, we performed a post hoc analysis of data from the SENSCIS trial to evaluate the risk of malnutrition over 52 weeks of treatment using a screening tool and to assess AEs in subgroups according to body mass index (BMI) at baseline.

PATIENTS AND METHODS

The design of the SENSCIS trial has been described, and the protocol is publicly available (15). Briefly, eligible patients had SSc with a first non–Raynaud's symptom in the prior ≤7 years, extent of fibrotic ILD (assessed in the whole lung) of ≥10% on high-resolution computed tomography, forced vital capacity ≥40% predicted, and diffusing capacity of the lung for carbon monoxide 30–89% predicted. Patients receiving prednisone ≤10 mg/day or equivalent and/or receiving stable therapy with mycophenolate or methotrexate for ≥6 months were allowed to participate. Patients were randomized in a 1:1 ratio (stratified by the presence of anti–topoisomerase I antibody) to receive nintedanib 150 mg twice daily or placebo until the last patient had reached week 52, but for ≤100 weeks. Treatment interruptions (≤4 weeks for AEs considered related to trial medication or ≤8 weeks for other AEs) and dose reductions to 100 mg twice daily were allowed to manage AEs. After resolution of the AE, treatment could be reintroduced, or the dose could be increased back to 150 mg twice daily. For diarrhea with an increase of <4 stools per day, antidiarrheal medicines were recommended; for diarrhea with an increase of 4–6 stools per day that persisted despite symptomatic care, or with an increase of ≥7 stools per day, incontinence, or life-threatening consequences, treatment interruption and/or dose adjustment was recommended (in addition to symptomatic care) (18).

AEs were reported by the investigators irrespective of causality and coded according to the Medical Dictionary for Regulatory Activities version 21.1. Weight was measured at baseline and at weeks 2, 4, 6, 12, 24, 36, and 52. A modified version of the Malnutrition Universal Screening Tool (MUST), a tool developed to identify adults at risk of malnutrition (20), was used to assess the risk of malnutrition at baseline and weeks 12, 24, 36, and 52. The MUST, which takes account of BMI, unplanned weight loss, and acute disease likely to affect nutritional intake, has been used to assess the risk of malnutrition in several studies in patients with SSc (6,9,11,21) and has been recommended for this purpose by expert groups (22,23). In the modified MUST, we calculated scores using BMI, weight loss, and a surrogate for acute disease effect (any serious AE that led to hospitalization between weight assessments and for which the patient received medication from the World Health Organization classification code "solutions for parenteral nutrition" for ≥5 days) (Figure 1). At baseline, the modified MUST score was based solely on BMI, as no data were available to assess weight loss and acute disease effect. The MUST score ranged from 0 to 6. As in the original MUST, we regarded scores of 0, 1, and ≥2 as indicating a low, medium, and high risk of malnutrition, respectively.

To investigate whether AEs were reported more frequently in patients with a low BMI at baseline, we assessed AEs reported over 52 weeks in subgroups by baseline BMI ≤20 kg/m^2 and BMI >20 kg/m^2. We assessed mean MUST scores; the proportions of patients at low, medium, and high risk of malnutrition based on MUST scores at baseline and at weeks 12, 24, 36, and 52; and the risk of malnutrition based on MUST scores at baseline and at last assessment. All analyses were descriptive and performed in patients who received ≥1 dose of trial drug.

The SENSCIS trial was carried out in compliance with the principles of the Declaration of Helsinki and the harmonized tripartite guideline for good clinical practice of the International Conference on Harmonization. The trial was performed at 194 sites in 32 countries and was approved by an independent ethics committee or institutional review board at every site. The sites are listed in the supplementary appendix to the primary manuscript on the trial results (15). All patients provided written informed consent before trial entry.

Data availability. To ensure independent interpretation of clinical study results and enable authors to fulfil their role and obligations under the International Committee of Medical Journal Editors criteria, Boehringer Ingelheim grants all external authors access to relevant clinical study data. In adherence with the Boehringer Ingelheim Policy on Transparency and Publication of Clinical Study Data, scientific and medical researchers can request access to clinical study data after publication of the primary manuscript in a

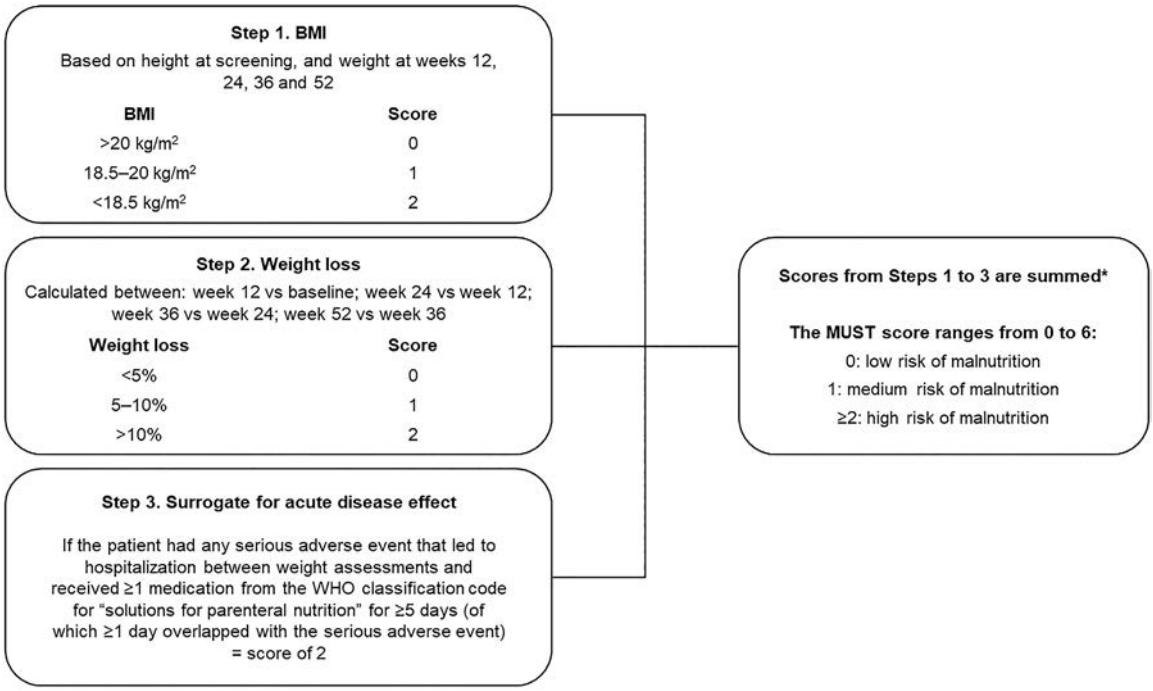


Figure 1. The modified Malnutrition Universal Screening Tool (MUST). BMI = body mass index; WHO = World Health Organization.

peer-reviewed journal, regulatory activities are complete, and other criteria are met. Researchers should access https://vivli.org/ to request access to study data and visit https://www.mystudywindow.com/msw/datasharing for further information.

RESULTS

Characteristics of subgroups by BMI at baseline. Among 576 patients, 61 patients (10.6%) had a BMI ≤20 kg/m^2 at baseline. Compared with patients with a BMI >20 kg/m^2 at baseline, those with a BMI ≤20 kg/m^2 had a lower mean age (48.6 versus 54.6 years) and a higher (worse) mean modified Rodnan skin score (16.1 versus 10.5). A greater proportion of patients with a baseline BMI ≤20 kg/m^2 compared to a BMI >20 kg/m^2 were female (83.6% versus 74.2%) and had diffuse cutaneous SSc (dcSSc) (62.3% versus 50.7%), while a lower proportion were receiving mycophenolate (39.3% versus 49.5%) (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25176/abstract). The proportions of patients positive for anti–topoisomerase I, anticentromere, antinuclear, or anti–RNA polymerase III antibodies were similar between the subgroups (see Supplementary Table 1). The proportions of patients with esophageal or stomach involvement, constipation, diarrhea, or hypertension at screening were lower or similar in those with a BMI ≤20 kg/m^2 compared to those with a BMI >20 kg/m^2 (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25176/abstract). Greater proportions of patients with a BMI ≤20 kg/m^2 compared to a BMI >20 kg/m^2 had pulmonary hypertension, joint contractures, digital ulcers, friction rubs, and atrophy (Supplementary Table 2).

AEs in subgroups by BMI at baseline. The AE profile of nintedanib was similar between subgroups according to BMI ≤20 kg/m^2 and BMI >20 kg/m^2 at baseline (Table 1). Diarrhea was the most common AE, reported in 79.2% of patients with a BMI ≤20 kg/m^2 and 75.4% of patients with a BMI >20 kg/m^2 at baseline. The frequencies of nausea and vomiting were also similar between the subgroups according to BMI. Weight loss AEs were less frequent in patients with a BMI ≤20 kg/m^2 compared to those with a BMI >20 kg/m^2 (4.2% versus 12.5%), but abdominal pain was more frequent in patients with a BMI ≤20 kg/m^2 compared to those with a BMI >20 kg/m^2 (20.8% versus 10.6%). Serious AEs were more frequent in patients with a BMI ≤20 kg/m^2 compared to those with a BMI >20 kg/m^2 (33.3% versus 23.1%). The frequencies of AEs leading to dose reduction, and of AEs leading to discontinuation of nintedanib, were similar in patients with a BMI ≤20 kg/m^2 and those with a BMI >20 kg/m^2 at baseline.

Modified MUST scores. In the nintedanib group, the mean ± SD MUST score increased (worsened) slightly from 0.3 ± 0.6 at weeks 12 and 24 to 0.4 ± 0.7 at weeks 36 and 52.

Table 1. Adverse events (AEs) in subgroups by body mass index (BMI) at baseline in the SENSCIS trial*

	BMI ≤20 kg/m^2		BMI >20 kg/m^2	
	Nintedanib (n = 24)	Placebo (n = 37)	Nintedanib (n = 264)	Placebo (n = 251)
Any AE(s)	24 (100)	34 (91.9)	259 (98.1)	242 (96.4)
Most frequent AEs†				
Diarrhea	19 (79.2)	8 (21.6)	199 (75.4)	83 (33.1)
Nausea	6 (25.0)	1 (2.7)	85 (32.2)	38 (15.1)
Vomiting	5 (20.8)	2 (5.4)	66 (25.0)	28 (11.2)
Skin ulcer	10 (41.7)	13 (35.1)	43 (16.3)	37 (14.7)
Cough	1 (4.2)	5 (13.5)	33 (12.5)	47 (18.7)
Nasopharyngitis	3 (12.5)	3 (8.1)	33 (12.5)	46 (18.3)
Upper respiratory tract infection	3 (12.5)	6 (16.2)	30 (11.4)	29 (11.6)
Abdominal pain	5 (20.8)	2 (5.4)	28 (10.6)	19 (7.6)
Fatigue	1 (4.2)	0 (0.0)	30 (11.4)	20 (8.0)
Weight decrease	1 (4.2)	0 (0.0)	33 (12.5)	12 (4.8)
AE(s) leading to treatment discontinuation	4 (16.7)	5 (13.5)	42 (15.9)	20 (8.0)
AE(s) leading to dose reduction	9 (37.5)	2 (5.4)	89 (33.7)	8 (3.2)
Serious AE(s)‡	8 (33.3)	8 (21.6)	61 (23.1)	54 (21.5)
Fatal AE(s)	2 (8.3)	2 (5.4)	3 (1.1)	2 (0.8)

* Values are number (%) of patients with ≥1 such AE reported over 52 weeks (or until 28 days after last trial drug intake in patients who discontinued the trial drug before week 52).
† AEs were coded according to preferred terms in the Medical Dictionary for Regulatory Activities, and AEs reported in >10% of patients in either treatment group in the overall population are shown.
‡ Serious AE indicates an event that resulted in death, was life threatening, resulted in hospitalization or prolonged hospitalization, resulted in persistent or clinically significant disability or incapacity, was a congenital anomaly or birth defect, or was deemed serious for any other reason.

In the placebo group, the mean ± SD MUST score was 0.2 ± 0.5 at weeks 12, 24, and 36 and 0.2 ± 0.6 at week 52.

At baseline, the proportions of patients at low, medium, and high risk of malnutrition based on MUST score were 91.7%, 5.9%, and 2.4% in the nintedanib group, and 87.2%, 8.0%, and 4.9%, respectively, in the placebo group (Figure 2). Between week 12 and week 52, the proportions of patients classified as at low risk of malnutrition based on MUST score ranged from 72.9% to 81.8% in the nintedanib group and from 80.8% to 88.3% in the placebo group. Over the same period, the proportions of patients classified as at high risk of malnutrition ranged from 5.6% to 9.6% in the nintedanib group and from 4.3% to 5.4% in the placebo group (Figure 2).

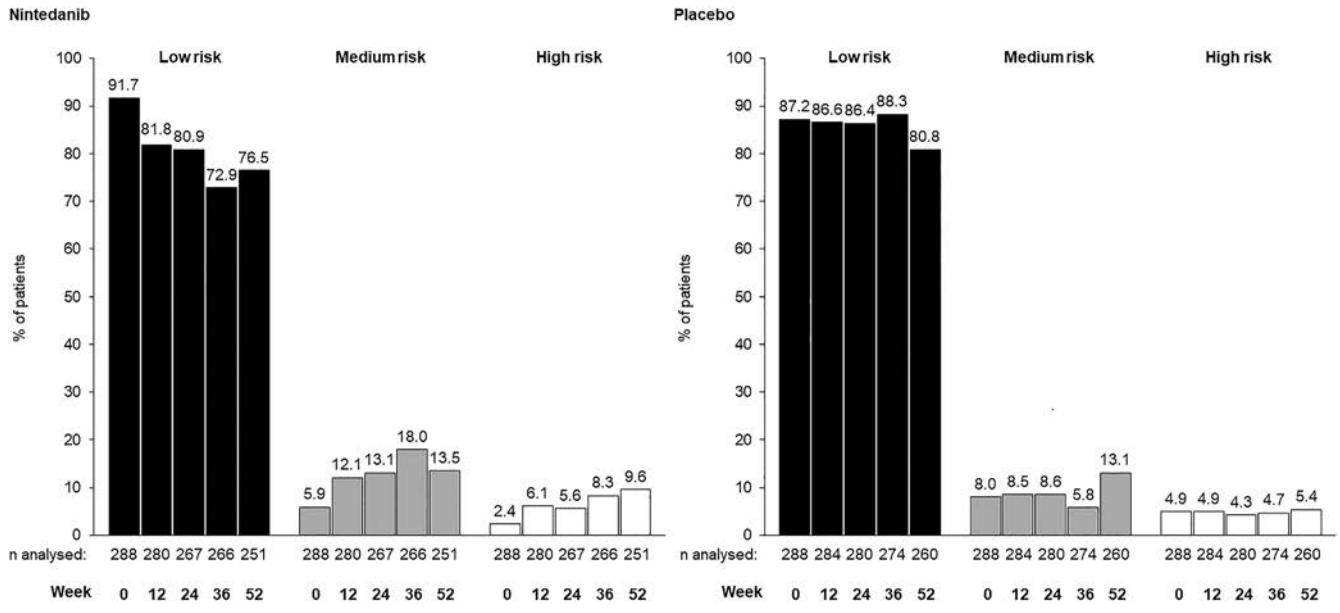


Figure 2. Risk of malnutrition based on a modified Malnutrition Universal Screening Tool (MUST) at baseline and at weeks 12, 24, 36, and 52 in the SENSCIS trial.

MUST scores suggested that the proportion of patients who remained at low risk of malnutrition between baseline and their last measurement was numerically lower in the nintedanib group compared to the proportion in the placebo group (74.0% versus 78.1%) (Table 2). The proportion of patients who were at low risk of malnutrition at baseline and were at high risk of malnutrition by their last measurement was higher in the nintedanib group compared to the placebo group (4.5% versus 1.0%) (Table 2).

DISCUSSION

In these post hoc analyses of data from the SENSCIS trial, the AE profile of nintedanib, including the frequency of AEs leading to treatment discontinuation, were similar between patients with a low BMI (≤20 kg/m^2) and those with a higher BMI at baseline. Nintedanib was associated with an increased risk of gastrointestinal AEs and weight loss compared to placebo, but patients with a low BMI at baseline did not appear to be at a greater risk of experiencing these events. Based on a modified version of the MUST, most patients had a low risk of malnutrition at baseline. However, the proportion of patients who had a low risk of malnutrition at baseline and remained at low risk at their last assessment over 52 weeks was lower in patients who received treatment with nintedanib compared to those who received placebo (74.0% versus 78.1%). The proportion of patients who were classified as at high risk of malnutrition at week 52 was greater in patients who received nintedanib compared to those who received placebo (9.6% versus 5.4%).

There is some evidence to suggest that the risks of various types of gastrointestinal involvement, and their severity, vary between patients with SSc with different characteristics related to sex (24,25), autoantibody profile (25–29), duration of SSc (25,27), SSc subtype (dcSSc versus limited cutaneous SSc) (25,29), manifestations of SSc such as myopathy (24,25), and medication use (29,30). A number of factors have been associated with malnutrition in patients with SSc, such as a greater number of gastrointestinal symptoms (6), the presence of oral aperture or microstomia (6,31), and greater disease severity (6,7,9). Nintedanib is an intracellular inhibitor of tyrosine kinases that inhibits processes such as fibroblast proliferation, migration and activation, and the deposition of extracellular matrix (32,33). The exact mechanism or mechanisms by which nintedanib causes gastrointestinal side effects is unknown, but it may be that inhibition of the vascular endothelial growth factor receptor causes morphometric changes in the bowel mucosa, altering motility (34). At present, it is not possible to predict gastrointestinal side effects, or their severity, in an individual patient who receives treatment with nintedanib.

Our findings illustrate the importance of monitoring for gastrointestinal problems, weight loss, and malnutrition in patients with SSc–ILD who receive treatment with nintedanib and ensuring that patients receive nutritional counselling when needed. Indeed, monitoring weight and nutritional status should be part of the care of all patients with SSc (22,23,35,36). An expert panel recommended that all patients with SSc be screened for malnutrition using a tool such as the MUST, combined with laboratory tests and detailed questioning of the patient regarding gastrointestinal problems, and that patients with SSc should weigh themselves monthly (22). Patients with SSc–ILD who receive treatment with nintedanib should be informed about the risk of gastrointestinal side effects and how these should be managed through dose adjustment, treatment interruption, and/or the use of therapies to relieve symptoms. Involvement of a gastroenterology team may be helpful, particularly when it is unclear whether gastrointestinal symptoms are due to the underlying SSc, comorbidities, or medication use (37).

Previous analyses of data from the SENSCIS trial suggested that gastrointestinal AEs associated with nintedanib were not more frequent in patients with a predisposition to gastrointestinal problems based on medical history and/or the presence of certain gastrointestinal problems at baseline (18). Further, although mycophenolate may be associated with gastrointestinal side effects, the proportion of patients with gastrointestinal AEs, and the proportion who prematurely discontinued nintedanib, were

Table 2. Risk of malnutrition based on a modified Malnutrition Universal Screening Tool (MUST) at baseline and at the last assessment of risk over 52 weeks in the SENSCIS trial*

	Last assessment of risk				
Baseline risk	Low	Medium	High	Missing	Total
Nintedanib					
Low	213 (74.0	31 (10.8)	13 (4.5)	7 (2.4)	264 (91.7)
Medium	1 (0.3)	8 (2.8)	8 (2.8)	0 (0.0)	17 (5.9)
High	0 (0.0)	0 (0.0)	7 (2.4)	0 (0.0)	7 (2.4)
Total	214 (74.3)	39 (13.5)	28 (9.7)	7 (2.4)	288 (100)
Placebo					
Low	225 (78.1)	20 (6.9)	3 (1.0)	3 (1.0)	251 (87.2)
Medium	7 (2.4)	14 (4.9)	2 (0.7)	0 (0.0)	23 (8.0)
High	0 (0.0)	4 (1.4)	10 (3.5)	0 (0.0)	14 (4.9)
Total	232 (80.6)	38 (13.2)	15 (5.2)	3 (1.0)	288 (100)

* Values are number (%) of patients.

similar between patients receiving mycophenolate and those not receiving mycophenolate at baseline (16). Analyses of pooled data from clinical trials of nintedanib in patients with a variety of ILDs have indicated that its AE profile is generally similar between male and female patients, but that nausea, vomiting, and hepatic AEs, and the use of dose reductions and treatment interruptions to manage AEs are more frequent in female patients (38).

Weight loss associated with nintedanib therapy does not appear to be a greater problem in patients with SSc–ILD compared to patients with other ILDs. The proportion of nintedanib-treated patients who experienced weight loss AEs over 52 weeks of the SENSCIS trial (11.8%) (15) was similar to the proportion observed in the INPULSIS trials in patients with IPF (9.7%) (17) and the INBUILD trial in patients with progressive fibrosing ILDs other than IPF (12.3%) (39). Data from the open-label extension of the SENSCIS trial, SENSCIS-ON, suggest that the safety and tolerability profile of nintedanib, including the risk of weight loss, is similar over longer-term use (40,41).

Strengths of our analyses include the large cohort of patients included and the standardization of data collection in the setting of a clinical trial. Limitations of our analyses include that they were post hoc and that the follow-up period was only 52 weeks, so the long-term consequences of weight loss or malnutrition could not be assessed. Information on weight loss and nutritional status prior to inclusion in the trial were not available. The number of patients with a BMI ≤20 kg/m^2 at baseline was quite small (n = 61). The MUST was not developed to evaluate the risk of malnutrition in patients with SSc. Tools developed specifically for patients with SSc, such as the PREdictor of MAlnutrition in Systemic Sclerosis score (42), may be valuable for future research.

In conclusion, in the SENSCIS trial in patients with SSc–ILD, the AE profile of nintedanib was similar between subgroups by BMI ≤20 kg/m^2 and BMI >20 kg/m^2 at baseline. Scores based on a modified MUST indicated that most patients remained at low risk of malnutrition over 52 weeks of treatment, but the proportion of patients who were classified as at high risk of malnutrition was higher among patients receiving treatment with nintedanib compared to those receiving placebo. Management of disease manifestations and gastrointestinal AEs that may be associated with weight loss is important to reduce the risk of malnutrition in patients with SSc–ILD treated with nintedanib.

ACKNOWLEDGMENTS

We would like to thank the patients who participated in the SENSCIS trial. The authors meet criteria for authorship as recommended by the International Committee of Medical Journal Editors. We would also like to thank Elizabeth Ng and Wendy Morris of FleishmanHillard, London, UK, who provided writing assistance, which was contracted and funded by Boehringer Ingelheim.

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. Volkmann 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. Volkmann, Miede, Alves.

Acquisition of data. Smith, Jouneau, Herrick.

Analysis and interpretation of data. Volkmann, McMahan, Smith, Jouneau, Miede Alves, Herrick.

ROLE OF THE STUDY SPONSOR

Boehringer Ingelheim International GmbH participated in the study design, data collection, statistical analyses, data interpretation, and the writing of the report. Boehringer Ingelheim was given the opportunity to review the manuscript for medical and scientific accuracy as well as intellectual property considerations. Publication of this article was contingent upon approval by Boehringer Ingelheim.

REFERENCES

1. Van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/-European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
2. Thoua NM, Bunce C, Brough G, et al. Assessment of gastrointestinal symptoms in patients with systemic sclerosis in a UK tertiary referral centre. Rheumatology (Oxford) 2010;49:1770–5.
3. Schmeiser T, Saar P, Jin D, et al. Profile of gastrointestinal involvement in patients with systemic sclerosis. Rheumatol Int 2012;32:2471–8.
4. Jaeger VK, Wirz EG, Allanore Y, et al. Incidences and risk factors of organ manifestations in the early course of systemic sclerosis: a longitudinal EUSTAR study. PloS One 2016;11:e0163894.
5. Volkmann ER, McMahan Z. Gastrointestinal involvement in systemic sclerosis: pathogenesis, assessment and treatment [review]. Curr Opin Rheumatol 2022;34:328–36.
6. Baron M, Hudson M, Steele R, et al. Malnutrition is common in systemic sclerosis: results from the Canadian Scleroderma Research Group database. J Rheumatol 2009;36:2737–43.
7. Caimmi C, Caramaschi P, Venturini A, et al. Malnutrition and sarcopenia in a large cohort of patients with systemic sclerosis. Clin Rheumatol 2018;37:987–97.
8. Hughes M, Heal C, Siegert E, et al. Significant weight loss in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. Ann Rheum Dis 2020;79:1123–5.
9. Hvas CL, Harrison E, Eriksen MK, et al. Nutritional status and predictors of weight loss in patients with systemic sclerosis. Clin Nutr ESPEN 2020;40:164–70.
10. Krause L, Becker MO, Brueckner CS, et al. Nutritional status as marker for disease activity and severity predicting mortality in patients with systemic sclerosis. Ann Rheum Dis 2010;69:1951–7.
11. Cereda E, Codullo V, Klersy C, et al. Disease-related nutritional risk and mortality in systemic sclerosis. Clin Nutr 2014;33:558–61.
12. Cruz-Dominguez MP, Garcia-Collinot G, Saavedra MA, et al. Malnutrition is an independent risk factor for mortality in Mexican patients with systemic sclerosis: a cohort study. Rheumatol Int 2017;37:1101–9.
13. Omair MA, Alahmadi A, Johnson SR. Safety and effectiveness of mycophenolate in systemic sclerosis. A systematic review [review]. PLoS One 2015;10:e0124205.
14. Barnes H, Holland AE, Westall GP, et al. Cyclophosphamide for connective tissue disease-associated interstitial lung disease [review]. Cochrane Database Syst Rev 2018;1:CD010908.
15. Distler O, Highland KB, Gahlemann M, et al. Nintedanib for systemic sclerosis-associated interstitial lung disease. N Engl J Med 2019; 380:2518–28.

16. Highland KB, Distler O, Kuwana M, et al. Efficacy and safety of nintedanib in patients with systemic sclerosis-associated interstitial lung disease treated with mycophenolate: a subgroup analysis of the SENSCIS trial. Lancet Respir Med. 2021;9:96–106.
17. Corte T, Bonella F, Crestani B, et al. Safety, tolerability and appropriate use of nintedanib in idiopathic pulmonary fibrosis. Respir Res 2015;16:116.
18. Seibold JR, Maher TM, Highland KB, et al. Safety and tolerability of nintedanib in patients with systemic sclerosis-associated interstitial lung disease: data from the SENSCIS trial. Ann Rheum Dis 2020;79:1478–84.
19. Cottin V, Martinez FJ, Jenkins RG, et al. Safety and tolerability of nintedanib in patients with progressive fibrosing interstitial lung diseases: data from the randomized controlled INBUILD trial. Respir Res 2022;23:85.
20. British Association for Parenteral and Enteral Nutrition. The 'MUST' toolkit. 2020. URL: https://www.bapen.org.uk/screening-and-must/must/must-toolkit/the-must-itself.
21. Yalcinkaya Y, Erturk Z, Unal AU, et al. The assessment of malnutrition and severity of gastrointestinal disease by using symptom-based questionnaires in systemic sclerosis: is it related to severe organ involvement or capillary rarefaction at microcirculation? Clin Exp Rheumatol 2020;38 Suppl 125:127–31.
22. Baron M, Bernier P, Côté LF, et al. Screening and therapy for malnutrition and related gastro-intestinal disorders in systemic sclerosis: recommendations of a North American expert panel. Clin Exp Rheumatol 2010;28 Suppl 58:S42–6.
23. Hansi N, Thoua N, Carulli M, et al. Consensus best practice pathway of the UK Scleroderma Study Group: gastrointestinal manifestations of systemic sclerosis [review]. Clin Exp Rheumatol 2014;32 Suppl 86:S214–21.
24. McMahan ZH, Paik JJ, Wigley FM, et al. Determining the risk factors and clinical features associated with severe gastrointestinal dysmotility in systemic sclerosis. Arthritis Care Res (Hoboken) 2018;70:1385–92.
25. Dein E, Kuo PL, Hong YS, et al. Evaluation of risk factors for pseudo-obstruction in systemic sclerosis. Semin Arthritis Rheum 2019;49:405–10.
26. Lazzaroni MG, Marasco E, Campochiaro C, et al. The clinical phenotype of systemic sclerosis patients with anti-PM/Scl antibodies: results from the EUSTAR cohort. Rheumatology (Oxford) 2021;60:5028–41.
27. Tauber M, Avouac J, Benahmed A, et al. Prevalence and predictors of small intestinal bacterial overgrowth in systemic sclerosis patients with gastrointestinal symptoms. Clin Exp Rheumatol 2014;32 Suppl 86:S82–7.
28. Ahmed F, Maclean RH, Nihtyanova SI, et al. Autoantibody predictors of gastrointestinal symptoms in systemic sclerosis. Rheumatology (Oxford) 2022;61:781–6.
29. Van Leeuwen NM, Boonstra M, Fretheim H, et al. Gastrointestinal symptom severity and progression in systemic sclerosis. Rheumatology (Oxford) 2022;61:4024–34.
30. Maclean RH, Ahmed F, Ong VH, et al. A phenome-wide association study of drugs and comorbidities associated with gastrointestinal dysfunction in systemic sclerosis. J Rheumatol 2023:jrheum.220990.
31. Türk İ, Cüzdan N, Çiftçi V, et al. Malnutrition, associated clinical factors, and depression in systemic sclerosis: a cross-sectional study. Clin Rheumatol 2020;39:57–67.
32. Wollin L, Distler JH, Denton CP, et al. Rationale for the evaluation of nintedanib as a treatment for systemic sclerosis-associated interstitial lung disease [review]. J Scleroderma Relat Disord 2019;4:212–8.
33. Wollin L, Distler JH, Redente EF, et al. Potential of nintedanib in treatment of progressive fibrosing interstitial lung diseases [review]. Eur Respir J 2019;54:1900161.
34. Bowen JM. Mechanisms of TKI-induced diarrhea in cancer patients [review]. Curr Opin Support Palliat Care 2013;7:162–7.
35. Saketkoo LA, Frech T, Varjú C, et al. A comprehensive framework for navigating patient care in systemic sclerosis: a global response to the need for improving the practice of diagnostic and preventive strategies in SSc [review]. Best Pract Res Clin Rheumatol 2021;35:101707.
36. Burlui AM, Cardoneanu A, Macovei LA, et al. Diet in scleroderma: is there a need for intervention? Diagnostics (Basel) 2021;11:2118.
37. Chatterjee S, Perelas A, Yadav R, et al. Viewpoint: a multidisciplinary approach to the assessment of patients with systemic sclerosis-associated interstitial lung disease [review]. Clin Rheumatol 2023;42:653–61.
38. Hoffmann-Vold AM, Volkmann ER, Allanore Y, et al. Safety and tolerability of nintedanib in patients with interstitial lung diseases in subgroups by sex: a post-hoc analysis of pooled data from four randomised controlled trials. Lancet Rheumatol 2022;4:e679–87.
39. Flaherty KR, Wells AU, Cottin V, et al. Nintedanib in progressive fibrosing interstitial lung diseases. N Engl J Med 2019;381:1718–27.
40. Allanore Y, Vonk MC, Distler O, et al. Continued treatment with nintedanib in patients with systemic sclerosis-associated interstitial lung disease: data from SENSCIS-ON. Ann Rheum Dis 2022;81:1722–9.
41. Allanore Y, Vonk MC, Distler O, et al. Continued treatment with nintedanib in patients with systemic sclerosis-associated interstitial lung disease (SSc-ILD): three-year data from SENSCIS-ON [abstract]. Arthritis Rheumatol 2022;74 Suppl 9. URL: https://acrabstracts.org/abstract/continued-treatment-with-nintedanib-in-patients-with-systemic-sclerosis-associated-interstitial-lung-disease-ssc-ild-three-year-data-from-senscis-on/.
42. Bagnato G, Pigatto E, Bitto A, et al. The PREdictor of MAlnutrition in Systemic Sclerosis (PREMASS) Score: a combined index to predict 12 months onset of malnutrition in systemic sclerosis. Front Med (Lausanne) 2021;8:651748.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2508–2518
DOI 10.1002/acr.25172
© 2023 American College of Rheumatology.

Understanding Stakeholders' Perspectives to Increase COVID-19 Vaccine and Booster Uptake Among Black Individuals With Rheumatic Conditions

Nnenna Ezeh,[1] Greta Sirek,[1] Sciaska N. Ulysse,[1] Jessica N. Williams,[2] Mia T. Chandler,[3] Bisola O. Ojikutu,[4] Michael York,[5] Monica Crespo-Bosque,[5] Muriel Jean-Jacques,[6] Tonya Roberson,[7] Karen Mancera-Cuevas,[8] Holly Milaeger,[9] Elena Losina,[10] Amar Dhand,[1] Mary Beth Son,[3] Rosalind Ramsey-Goldman,[9] and Candace H. Feldman[10]

Objective. Disparities in COVID-19 vaccine and booster uptake persist. This study aimed to obtain perspectives from community and physician stakeholders on COVID-19 vaccine and booster hesitancy and strategies to promote vaccine uptake among Black individuals with rheumatic and musculoskeletal conditions.

Methods. We invited community leaders and physicians in greater Boston and Chicago to participate in semi-structured interviews using a moderator guide developed a priori. Participants were queried about how to best address vaccine hesitancy, strategies to target high-risk populations, and factors to identify future community leaders. Interviews were audio recorded, transcribed verbatim, and analyzed thematically using Dedoose.

Results. A total of 8 physicians and 12 community leaders participated in this study between November 2021 and October 2022. Qualitative analyses revealed misinformation/mixed messaging and mistrust, with subthemes including conspiracy theories, concerns regarding vaccine development and function, racism and historical injustices, and general mistrust of health care systems as the top cited reasons for COVID-19 vaccine hesitancy. Participants also shared demographic-specific differences, such as race, ethnicity, age, and gender that influenced the identified themes, with emphasis on COVID-19 vaccine access and apathy. Strategies for community-based vaccine-related information dissemination included personal storytelling with an iterative and empathetic approach, while recognizing the importance of protecting community leader well-being.

Conclusion. To increase vaccine uptake among Black individuals with rheumatic conditions, strategies should acknowledge and respond to racial/ethnic and socioeconomic injustices that engender vaccine hesitancy. Messaging should be compassionate, individually tailored, and recognize heterogeneity in experiences and opinions. Results from these analyses will inform a planned community-based intervention in Boston and Chicago.

INTRODUCTION

While nationwide efforts have helped reduce COVID-19 vaccine hesitancy and promote uptake among individuals from historically marginalized populations (1), inequities in vaccine series completion and booster uptake persist (2). As of May 10, 2023, the percentage of Black individuals (45.0%) who have completed the primary COVID-19 vaccine series remains below that of White individuals (51.9%) and Hispanic/Latinx

The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Drs. Ramsey-Goldman and Feldman's work was supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant R01-AR-080089). Dr. Feldman's work was supported by the Rheumatology Research Foundation (R Bridge award).

[1]Nnenna Ezeh, MD, Greta Sirek, BA, Sciaska N. Ulysse, BA, Amar Dhand, MD, DPhil: Brigham and Women's Hospital, Boston, Massachusetts; [2]Jessica N. Williams, MD, MPH: Emory School of Medicine, Atlanta, Georgia; [3]Mia T. Chandler, MD, MPH, Mary Beth Son, MD: Boston Children's Hospital and Harvard Medical School, Boston, Massachusetts; [4]Bisola O. Ojikutu, MD, MPH, FIDSA: Harvard Medical School, Boston Public Health Commission, and Brigham and Women's Hospital, Boston, Massachusetts; [5]Michael York, MD, Monica Crespo-Bosque, MD: Boston Medical Center, Boston, Massachusetts; [6]Muriel Jean-Jacques, MD, MPH: Northwestern Medicine, Chicago, Illinois; [7]Tonya Roberson, PhD, MPH, DTR: Governors State University, University Park, Illinois; [8]Karen Mancera-Cuevas, DrPH, MS, MPH, CHES: Illinois Department of Public Health, Springfield, Illinois; [9]Holly Milaeger, MPH, Rosalind Ramsey-Goldman, MD, DrPH: Northwestern Medicine/Feinberg School of Medicine, Chicago, Illinois; [10]Elena Losina, PhD, Candace H. Feldman, MD, MPH, ScD: Harvard Medical School and Brigham and Women's Hospital, Boston, Massachusetts.

Dr. Ezeh, Ms. Sirek, and Ms. Ulysse contributed equally to this work.

Drs. Ramsey-Goldman and Feldman contributed equally to this work.

Author disclosures and a graphical abstract are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25172.

Address correspondence via email to Candace H. Feldman, MD, MPH, ScD, at cfeldman@bwh.harvard.edu.

Submitted for publication January 10, 2023; accepted in revised form June 8, 2023.

SIGNIFICANCE & INNOVATIONS

- This study uniquely leverages qualitative interviews from both community members and physicians from 2 US cities to understand COVID-19 vaccine and booster hesitancy and to develop strategies to promote COVID-19 vaccine uptake, specifically among Black individuals with rheumatic conditions.
- Misinformation and mistrust stemming from racism and historical injustices were uncovered as central themes that need to be addressed by vaccine/booster hesitancy programs to facilitate uptake and reduce racial disparities.
- These interviews with key community and academic stakeholders will inform a robust, tailored, community-based intervention to promote vaccine and booster uptake among individuals with rheumatic and musculoskeletal conditions, with a focus on Black and African American communities in Chicago and Boston.

individuals (57.3%) (3). Further, fewer Black individuals (9.5%) and Hispanic/Latinx individuals (9.1%) have received the bivalent Omicron booster compared to White individuals (16.7%) (3). Vaccine hesitancy among Black and Hispanic individuals may stem from inequitable care and structural racism in the US (4–6). Structural racism refers to the ways in which societies foster racial discrimination through mutually reinforcing macro-level systems (e.g., housing, education, health care, criminal justice) (7). Additionally, the ever-changing COVID-19 vaccination public health campaign has led to confusion, perpetuating vaccine hesitancy (8–10).

Among individuals with rheumatic conditions, studies have demonstrated a disproportionate burden of COVID-19 disease, adverse outcomes, and vaccine hesitancy among historically marginalized individuals. In a global study, 51% (n = 131 of 256) of Black patients (odds ratio [OR] 3.18 [95% confidence interval (95% CI) 2.31–4.36]) and 37% (n = 103 of 279) of Latinx patients (OR 2.00 [95% CI 1.46–2.75]) with rheumatic conditions were hospitalized with COVID-19–related complications compared to 29% of White (n = 187 of 639) patients (11). Odds were modestly attenuated yet remained statistically significant after adjusting for comorbidities, medication use, and disease activity (11). Immunocompromised individuals often have a less robust vaccine response, increasing the importance of vaccinating their social networks for a "cocooning effect" to reduce transmission from close contacts (12). Despite high rates of serious infections among individuals with systemic lupus erythematosus (SLE) and rheumatoid arthritis, vaccination rates are low overall, with disparities by race, ethnicity, and insurance status (13–15).

As of May 10, 2023, only 17.0% of the US population had received an updated bivalent booster dose (3). Exploring strategies to increase vaccination and booster rates are urgently needed among historically marginalized populations with rheumatic conditions. Partnering with community-based leaders to disseminate accurate, culturally tailored information is essential. Popular Opinion Leaders (POLs), or trusted community members, can be trained to disseminate information about risk-reducing health behaviors through their social networks (16–18). The POL model can increase community knowledge, reduce stigma, and prompt behavioral changes and has been applied to improve access to care and clinical trial enrollment among individuals with SLE (16–18).

We aimed to interview community and physician stakeholders in 2 US cities regarding community and patient perspectives on COVID-19 vaccines and boosters, barriers, and strategies to promote vaccine uptake. This qualitative data will inform a future intervention that leverages the POL model to improve COVID-19 vaccine and booster uptake among Black individuals with rheumatic conditions.

PATIENTS AND METHODS

Participants and data collection

We recruited physicians and community organization leaders and/or active members we had established partnerships with, several of whom were previously trained POLs, ages ≥18 years, to participate in virtual, key informant, semi-structured individual interviews. We developed moderator guides a priori with separate versions for physicians and community leaders (see Supplementary Material 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25172/abstract). Study staff (SNU, CHF), who identify as Black and White, respectively, together led physician and community leader semi-structured interviews. Following verbal consent, interviews (~40 minutes) were conducted virtually via Zoom from November 2021 to October 2022, and were audio recorded, deidentified, and transcribed verbatim. We concluded recruitment once thematic saturation was reached. This study was approved by the Mass General Brigham Institutional Review Board.

Qualitative analyses

Transcripts were analyzed using the constant comparative method, a data coding process used for categorizing and comparing qualitative data (19). Study team members (NE, GS, SNU, CHF) individually reviewed transcripts and independently identified themes. The team met, reviewed preliminary themes, adjudicated discrepancies, and defined final themes. The team developed a coding system that reflected themes, subthemes, and categories for future intervention design. Initial codes, definitions, and examples were collated (GS) and agreed upon by the team (NE, SNU, CHF). Demographic-specific reasons for vaccine

hesitancy, and the ways these factors permeated identified themes, were emphasized in the coding system.

We used an iterative approach to develop a standard coding system. Two researchers (NE, GS) individually coded Transcript 1, adjudicated differences in codes, changed the coding system accordingly, and finalized the coding system after parallel coding was implemented for Transcript 2. The remaining transcripts were then coded separately by 2 reviewers (NE, GS). Kappa coefficients for inter-coder reliability were 0.79 (NE for re-review of GS transcripts) and 0.72 (GS for re-review of NE transcripts); 75% (286 of 380) of excerpts were coded the same. Five of the differences were excluded because of lack of context/incomplete quotes, 28 had the same theme but different subthemes (e.g., mistrust with conspiracy theories versus mistrust with politicized skepticism) or used the same theme but 1 added demographic-specific code (e.g., strategies versus strategies, demographic specific), and 61 had a difference in theme. A third coder (SNU) reviewed and adjudicated all discrepancies. Eight interviewees also reviewed the themes, descriptions, and selected quotes as presented in this article for their critical input. We stratified qualitative analyses by participant role (physician or community leader/POL). While primary analyses focused on themes, we also quantified the number of quotes by theme/subtheme stratified by role. Dedoose software was used for analyses (20).

RESULTS

Twenty individuals (ages ≥18 years) participated in 18 semi-structured interviews. Eight were physicians in adult and pediatric rheumatology, infectious disease, and primary care. Twelve were community leaders affiliated with the Center for Community Health Education Research, Sportsmen's Tennis & Enrichment Center, Mission Hill Health Movement, AllianceChicago, Women of Courage, and the Labalaba Foundation. A total of 6 community leaders were previously trained as POLs through SLE-related projects (16,18). A total of 5 participants were from greater Chicago, IL, and 15 participants were from greater Boston, MA. A total of 9 participants identified as Black, 4 identified as White, 3 identified as "other," 1 identified as Asian, 1 identified as Black and other, and 2 did not respond. A total of 15 participants identified as female, 3 identified as male, and 2 did not respond. We were unable to transcribe and analyze 1 community leader interview due to poor sound quality, but detailed notes during the interview informed the themes. Three community representatives from the same organization joined 1 interview, which was coded as 1 community-leader participant, since the individuals spoke collectively, providing complementary answers.

We defined and identified overarching themes and subthemes (Table 1). Representative quotes by theme were extracted from physician (Table 2) and community leader/POL

Table 1. Description of themes and subthemes*

Theme and subthemes†	Definition
Mistrust	Doubt in and suspicion of people, organizations, and systems
Conspiracy theories	Belief that the vaccine and pandemic is a result of a secret plot
Health care systems	General suspicion of pharmaceutical companies, physicians, and health care institutions
Politicized skepticism	Politically driven vaccine hesitancy
Racism and historical injustices	Racism or historical injustices in health care
Vaccine development and function	Concerns and misunderstanding regarding vaccine development and function
Access	Material, psychological, and systematic barriers that inhibit vaccination
Apathy	Notions that fuel vaccine-related indifference
Booster-specific hesitancy	Assessment of personal risk to benefit ratio that engenders delayed booster vaccine uptake
Disruptions	Disturbances and barriers to essential daily tasks and schedules that impact booster vaccination status
Messaging	Ineffective information dissemination that fuels booster-specific confusion and hesitancy
Categorization by degree of hesitancy	Classification of individuals by degree of vaccine hesitancy
Misinformation, and changing and inconsistent messaging	Vaccine information sources and messaging that fuel confusion, misinformation, and fear
Role of religion and cultural beliefs	Influence of religion and cultural beliefs on vaccination status
Safety	Profile of well-being and side effect concerns
Motivating factors	Reasons to seek out vaccination
Role of social network and sources	Who, where, and how an individual receives vaccination information impacts vaccine uptake
Strategies	Effective communication, psychological, and material means to disseminate vaccine-related information to community members
Categories for POL intervention design	
POL traits	Traits and characteristics that make an effective POL
Physician knowledge of patient social network	Knowledge and description of patient's social network
Concerns about being a POL	Concern and worries about being a person who disseminates information

* POL = Popular Opinion Leader.
† Demographic factor subtheme was included in coding system for the following themes: Mistrust, Access, Apathy, Safety, Motivating Factors, and Strategies.

Table 2. Reasons for COVID-19 vaccine hesitancy—selected quotes from physicians

Reason	Transcript
Mistrust	"There are many people who have no specific thing that they will cite on questioning. They just don't trust it." (Transcript 11)
Conspiracy theories	"I have had a couple people who do talk about it being implanting a microchip and being followed. I was surprised because they were not people that I would've thought would endorse that belief." (Transcript 11)
Demographic specific	"I work predominantly in [city], [city], and [city], and so obviously, predominant African American population… the younger demographic within that racial demographic, 18–35-year-olds who are often last to be targeted with the vaccine and not have a lot of direct, creative messaging that makes them compelled in any way to do it, and particularly at-risk young adults, so those who are gang involved and may not be getting information from school, may not get it from their employer, may not be getting it from other trusted sources." (Transcript 12)
Health care systems	"Trust in the pharmaceutical industry, I had patients saying right off before the vaccines even came out, 'I'm not taking anything from [name]. How can you trust them after the [name], or whatever, incident?'" (Transcript 11)
Politicized skepticism	"I get this sense that they think it's political and the democrats have pushed it forward and therefore it is inherently not to be trusted. Like inherently. No matter what I said or did." (Transcript 1)
Racism and historical injustices	"…there are young Black people, educated middle-class folks in my generation right now who are saying, 'You know, health care has not done well by us. I'm sure that this vaccine hasn't been studied in us. I'm not about to get a vaccine that is not made for us by us.'" (Transcript 6)
Vaccine development and function	"Initially I think there was a lot of hesitance from patients in general, particularly around maybe the newness of it or how quick they thought it had come out and maybe a general sense of things being experimental as opposed to established." (Transcript 4)
Access	"I would add ongoing issues with access… You have to put that in the context of people's lives. The chaos that people may be living in, the stress, the strain, the other worries, and think about, no, it's not gonna happen unless you bring that booster to their door and knock on it, and say, 'Here it is.'" (Transcript 15)
Demographic specific	"Then my seniors, many of them just don't like the online thing at all and completely are turned off by that. I have a few patients who are homebound, so that was another thing, and a couple who truly cannot leave, so identifying resources to go to their home to administer the vaccine, especially twice, is—and then, again, with the booster is a big difficulty." (Transcript 11)
Apathy	"Those numbers just don't seem to move people, like '3,000 people died today from COVID. X-number of people have had this.' It just doesn't seem to move." (Transcript 11)
Demographic-specific	"I also have a group of older patients particularly where the messaging is more like 'Why would I even get it, it's sort've if I get it and I die, I die.'" (Transcript 4). "We started first with food access work. That's what people really cared about when the pandemic started. They're like, 'Screw this vaccine. I don't care. I need food. I need basic services that aren't available to me right now.'" (Transcript 12)
Booster-specific hesitancy Disruptions	"I think, for people who got sick the first two times and missed work, that was a problem. They don't mind the idea of getting boosted. They just want a time where they feel like they can be sick for the next day or two." (Transcript 10)
Messaging	"The messaging has been poor in terms of boosters… Many people believe that boosters are an option, and just the term 'booster' doesn't equate to essential—doesn't equate to severity of illness and death if I don't get it. It just means that it's just an option." (Transcript 15)
Categorization by degree of hesitancy	"It's what everyone says about vaccine hesitancy. There are some people who are just no's, and they are super hard to move and there are people in the middle. And I usually know within the first couple of minutes of talking to them where they are going to be…" (Transcript 1)
Misinformation, and changing and inconsistent messaging	"People get really turned off when the message changes too much… the [name] vaccine was great, and you recommended me to take it. Now, it says that it's not good. Now, I trust you less." (Transcript 10)
Role of religion and cultural beliefs	"When you think about what are the norms within somebody's, not only somebody's personal norms, but their family norms, their social network norms, there are norms that are out there that people just don't necessarily feel like this is necessary. That their body will handle this." (Transcript 15)
Safety	"Heart side effects, especially, I hear, 'Heart disease is in our family. Why would we take something that can cause heart problems?'" (Transcript 11)

transcripts (Table 3). Previously trained POLs and community leaders were combined to maintain anonymity. The number of participants who mentioned each theme (Figure 1) and total number of times each theme was mentioned (Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25172/abstract) were counted and stratified by physician and community leader/POL. We stratified our figures based on whether the theme was mentioned by a physician or community leader/POL; however, we present thematic analyses by commonality of themes, since physicians often reflected on discussions with their patients.

Barriers to COVID-19 vaccine uptake

Mistrust. We defined mistrust as doubt in and suspicion of people, organizations, and systems that engendered vaccine

Table 3. Reasons for COVID-19 vaccine hesitancy—selected quotes from community leaders

Reason	Transcript
Mistrust	
Conspiracy theories	"That they wanna kill all the Black people, that they're tryin' to put a chip in so they can follow everyone. Oh, it was a lot of different ones. Some of them I couldn't even believe." (Transcript 18)
Politicized skepticism	"Then if you're looking at the society where we live, I think a lot of it has come from the fact that COVID and vaccines and all these things has really found its way very early on into the political arena. That skewed a lot of things. We're not talking about health and science anymore." (Transcript 7)
Racism and historical injustices	"First his comment was, 'I'm pretty sure there's something wrong with the vaccine, or'—and he said the reason being was because he was trying to figure out how it was that they were able to get access to the vaccines in the jail 'cause, as you know, we were prioritizing people who are incarcerated when his people on the outside couldn't get it. We explained that, well, actually we prioritize you, and so that's why you're able to get access to the vaccine. He paused for a moment, and then he looked, and he said, 'Well, I'm just wondering, because they never prioritized me before. That's why I'm here because nobody ever prioritized me. Why are they prioritizing me now? There must be something wrong with this vaccine.'" (Transcript 17)
Vaccine development and function	"You would hear community members sayin', 'Oh, how can they have a cure for COVID-19 and they don't have one for breast cancer." (Transcript 18)
Access	"Well, I mean there are always barriers, depending on where a person lives, but then I see so many other locations opening up. To me personally, there's no reason why you can't get the shot in terms of availability." (Transcript 8)
Demographic specific	"They were focused on the older Hispanic/Latinx population… a lot of concerns around documentation and insurance and what would be needed." (Transcript 16)
Apathy, demographic specific	"If you meet teenagers, a Black teenage boy who lives in a low-income area. I've met them and they've said, 'Oh,' a lot of them already have the colors for their funeral planned. It's like no expectation of getting old. Or I've met men who are 27, 28 who suddenly have to figure out how am I living my life? I never thought I would get this old. If you don't expect to live a long time, then maybe that's a reason not to get vaccinated. I don't know their perception. You know how boys are—teenagers. Their perception of risk is totally different than most people." (Transcript 14)
Booster-specific hesitancy	"I think people just are—they don't know exactly why they need it especially since there seems to be an opening up of everything… Then there's been a lot of people who had no problem with the first two shots and had severe reactions with the third one. That also makes 'em reluctant to get a fourth one because they don't know what the impact is. I think a lot of people—I think there's still a lot of question about what is [an] RNA vaccine. People are either in why we need to get these constant booster."(Transcript 14)
Messaging	"I think there's a carelessness with how words are used in media about talking about vaccines. I think because after talking with my doctor, I found out there's a difference between boosters and third doses, and I've heard them use interchangeably in public spaces, in the media and in radio and TV and that stuff." (Transcript 9)
Categorization by degree of hesitancy	"I know a lot of people had adopted the 'wait and see' attitude. Let the first group of people get shot up and we'll see if they're still alive in six months. Then we'll go ourselves." (Transcript 14)
Misinformation, and changing and inconsistent messaging	"I think there's confusion. I think people don't know where to go and get them. Someone just said to me the other day, 'now people over 65 or older can get the vaccine' and I'm like 'no, no no, no, no. People 65 and older are eligible, have been eligible'… However, when it just came out this week, it wasn't clear that it was for the general population. I had three people actually mention that to me because they were shocked I got the vaccine. I'm like 'what do you mean? I was cleared.'" (Transcript 3)
Role of religion and cultural beliefs	"I have some family members who are not—close family members who are not vaccinated, do not plan to get vaccinated, and it's been a real struggle to figure out how to overcome the objection 'cause I think some of it is real personal, religious-based. It just feels like it's really deeply ingrained, and almost as though nothing you could say could change their mind or spark their curiosity, or help them see what would, what some of the benefits could be on the other's side of the vaccination." (Transcript 9)

hesitancy. Mistrust and its subcategories (conspiracy theories, health care systems, demographic-specific mistrust, politicized skepticism, racism and historical injustices, and vaccine development/function) were the most cited barrier with >70 references by participants (see Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25172/abstract). "Conspiracy theories" was the most common subtheme, without explicit mention by the facilitators.

Conspiracy theories. With 12 participants discussing conspiracy theories (Figure 1), many felt the various beliefs were rooted in mistrust. Specific theories included microchip implantation, foreign government interference, money-making schemes by the government or pharmaceutical companies, or genocide concerns. One participant described, "I hear a lot of, 'They say.' My favorite thing to say, 'Who are they? Who are these they that are always saying, 'They say,' they who?' You always have these

Figure 1. Number of participants who mentioned theme or subtheme, stratified by physician and community leader title. Demographic characteristic–specific subthemes for access, apathy, mistrust, motivating factors, safety, and strategies were combined with the parent theme. POL = Popular Opinion Leader.

conspiracy theories of everything from death in a few years. They're going to die. They're trying to clean out people to all these things..." (Transcript 7, community leader).

Health care systems. Six participants mentioned 13 times that mistrust in health care systems created hesitancy. Many mentioned generalized mistrust of the medical system, direct challenges to science and scientific leaders, and frustration that the health care system functions primarily in a capitalistic role with respect to pharmaceutical companies developing the vaccine. One participant reflected, "Trust in the pharmaceutical industry, I had patients saying right off before the vaccines even came out, 'I'm not taking anything from [name]. How can you trust them after the [name] incident?'" (Transcript 11, physician).

Politicized skepticism. Political tension that engendered vaccine hesitancy was cited 7 times by 4 participants, noting that vaccine communication was impacted by the messengers, specifically those in political offices. This led to polarizing viewpoints, and concern for political motivations for vaccination support. One physician stated, "I get this sense that they think its political and the Democrats have pushed it forward and therefore it is inherently not to be trusted" (Transcript 1, physician).

Racism and historical injustices. Seven participants cited the role of historical injustices committed against marginalized communities and medical racism impacting hesitancy. They noted that previous racist acts were mentioned by those who were hesitant. Others stated that the vaccine felt like a tool for systematic oppression, while several identified concerns about lack of racial representation in clinical trials. For some, historical injustice played a direct role in mistrust; one participant described a person experiencing incarceration, noting, "First his comment was, 'I'm pretty sure there's something wrong with the vaccine'… He was trying to figure out how it was that they were able to get access to the vaccines in the jail. 'Well, I'm just wondering, because they never prioritized me before. That's why I'm [in jail] because nobody ever prioritized me. Why are they prioritizing me now? There must be something wrong with this vaccine'" (Transcript 17, community leader).

Vaccine development and function. Ten participants acknowledged concerns regarding vaccine development describing rushed processes or concern the messenger RNA vaccines altered DNA. One noted, "Initially I think there was a lot of hesitance from patients in general, particularly around maybe the newness of it or how quick they thought it had come out and maybe a general sense of things being experimental as opposed to established" (Transcript 4, physician).

Access. We defined the theme of access as systematic, psychological, and material barriers that inhibited vaccination. Nine participants mentioned access concerns 18 times. Barriers included limited internet access and proficiency, distances to vaccination clinics, availability of vaccines, and poor integration of vaccination efforts into existing health clinics. Some observed that access efforts were tied to trust; for example, "There are so many layers to trust and to access. I've heard from enumerable patients where they would not feel comfortable going to, say, a pop-up

vaccine [place] in a parking lot, but maybe they feel comfortable going to their local [pharmacy]…There's some that would never get it at their local [place]…[who] only feel comfortable getting it at their doctor's office" (Transcript 11, physician). Other participants felt otherwise: "I see so many other locations opening up. To me personally, there's no reason why you can't get the shot in terms of availability" (Transcript 8, community leader).

Apathy. Nine participants mentioned apathy, or vaccine indifference, as a barrier to uptake, noting that young, healthy individuals' altruism to protect the more vulnerable began to wane over time. One physician noted that vaccination was a lower priority for socially disadvantaged individuals who had to focus on basic needs like access to food. Others noted that some teenagers and elderly individuals expressed apathy regarding dying. For teenagers, the theme of fatalistic apathy was expressed: "If you meet teenagers, a Black teenage boy who lives in a low-income area… They've said, 'Oh,' a lot of them already have the colors for their funeral planned. It's like no expectation of getting old. Or I've met men who are 27, 28 who suddenly have to figure out how am living my life? I never thought I would get this old. If you don't expect to live a long time, then maybe that's a reason not to get vaccinated… Their perception of risk is totally different than most people" (Transcript 14, community leader).

Booster-specific hesitancy. We defined booster-specific hesitancy as assessment of personal risk/benefit ratio that limited uptake. Participants mentioned inconsistent messaging, waning precautions, fatigue, fear of side effects, and the perception of boosters as nonessential, 17 times. A physician noted, "This period of time where we're looking at second boosters and more boosters—I know that somebody asked me this who's not medical, 'Will this drain my immune system? You keep revving it up or boosting it'" (Transcript 15, physician). People adopted a "wait and see approach:" "I think people just are—they don't know exactly why they need it especially since there seems to be an opening up of everything… I think if there's another surge… people might start thinking about it… They'll just wait until the flu season starts and get both of them at once" (Transcript 14, community leader).

Categorization by degree of hesitancy. Seven participants acknowledged that some specific concerns of individuals varied by degree of hesitancy. Many reflected on individuals who would ultimately become vaccinated but wanted others to go first. Others refused vaccination outright, stating, "There's some people who know it's good and maybe take a pause before, weren't the first in line to sign up for it, but eventually did" (Transcript 9, community leader). A physician noted the importance of gauging this sense in clinic to tailor their messaging: "There are some people who are just no's and they are super hard to move and there are people in the middle. And I usually know within the first couple of minutes of talking to them where they are going to be" (Transcript 1, physician).

Misinformation and changing and inconsistent messaging. We defined misinformation and changing and inconsistent messaging as vaccine information sources and messaging that fueled confusion, misinformation, and fear. Thirteen participants reflected on the levels of misinformation impacting vaccine perception compounded by changing and inconsistent messaging as the pandemic progressed. The internet and word of mouth were avenues of misinformation as opinions were conflated with facts. Participants identified other factors including language barriers, policy implementation lags, and strong opinions. "People get really turned off when… the message changes too much… The [name] vaccine was great, and you recommended me to take it. Now, it says that it's not good. Now, I trust you less" (Transcript 10, physician). A community member compared information passing in communities to the telephone game, stating, "I get it from news sources… It gets handed down and everybody else adds. It's like the telephone game. Everybody else adds a little 2 inches more to the story, makes it 15 times worse than what it was" (Transcript 8, community leader).

Role of religion and cultural beliefs. A total of 3 community leaders and 1 physician identified the influence of religion and cultural beliefs on vaccination status. Some mentioned that it can be a cultural belief to question any vaccine. Many felt the incorporation of religion or spiritual beliefs made an individual even more resistant: "I think some of it is real personal, religious-based. It just feels like it's really deeply ingrained, and almost as though nothing you could say could change their mind or spark their curiosity, or help them see… Some of the benefits could be on the other's side of the vaccination" (Transcript 9, community leader). On the contrary, others noted how religion helped support vaccination, stating, "I pray about it, really get wisdom and discernment not only from the people who're the scientists but also just, 'God, is this a prudent thing to do?'… [They were] feeling confident that God will work through physicians again. Yes, we'll get vaccinated" (Transcript 17, community leader).

Safety. Safety concerns, notably side effects, were described by 8 individuals (6 physicians and 2 community leaders). Short- and long-term effects were mentioned, including risk of rheumatic disease flares, severity of vaccine-related adverse effects, and concern regarding interaction with treatment drugs. Others noted that while parents were vaccinated, safety concerns precluded them from vaccinating their children: "'Oh, no. I'm not vaccinating my kids. I heard it can affect your fertility.' That's been a really hard one to bust" (Transcript 11, physician).

Motivating factors and strategies for vaccine uptake

Motivating factors. We defined this theme as reasons to seek out vaccination. Nine participants described motivating factors >24 times, citing safety, desire to travel or spend time with loved ones, decreased social isolation, engagement in something

controllable, minimizing risk due to immunocompromised status, and fear of death/adverse outcomes (Table 4). One participant mentioned in their social circle specific reasons like "To be able to see their loved ones. That was the number one reason. To have a peace of mind… We want to have some stability of what we had before and not live in this isolation…" (Transcript 3, community leader).

Strategies. We defined this theme as effective communication, psychological, and material means to disseminate vaccine-related information to community members. Strategies ranged from personal storytelling, open-ended questioning, empathetic acknowledgement of concerns, iterative conversations, and an optimistic approach with a targeted integration of data (Table 4). Many spoke of managing expectations and having a goal. Physicians identified the importance of roles like POLs, with one mentioning, "How do you convince someone who's totally dug in?… If it came from someone outside the medical field…someone they knew in another sphere and trusted, they certainly have a way better shot than I did" (Transcript 1, physician).

Considerations for future intervention design

Participants were asked to identify strategies that would aid in designing a future vaccine-related intervention using the POL model (Table 5). Discussion points included ideal POL traits and the ability of physicians to identify patients to be trained as POLs. While not explicitly asked, 2 community leaders expressed potential concerns, which were explored as a theme.

POL traits. Eleven participants identified effective POL traits as humility, empathy, warmth, honesty, and sensitivity: "They need to be humble…approachable and honest. If you break down telling your story, it's okay. If you say I don't know the answer, I'll get back to you, it's totally okay" (Transcript 3, community leader). Others

Table 4. Motivating factors and strategies for vaccine uptake—selected quotes from physicians and community leaders

Variable	Transcript
Motivating factors	"The patients who have told me 'I'm hesitant to get this vaccine but I'm going to do it anyway because I want to protect my grandbaby.' I've heard people say things like that and that can be a powerful motivator for people." (Transcript 2, physician)
	"To be able to see their loved ones. That was the number one reason. To have a peace of mind, at least in my family that was a whole thing. Some of my friends were like, 'you can't see me until I get vaccinated or until you get vaccinated.' It became, we want to have—and I don't want to call it a normal lifestyle—but we want to have some stability of what we had before and not live in this isolation that we were all placed in." (Transcript 3, community leader)
Demographic specific	"I would say patients that are highly engaged with their health are more likely to get the vaccine. For example, if a patient in our clinic is frequently missing their appointments or they have medication adherence issues, they are the ones I'm finding to be less likely to be receptive to the vaccine. The patients who are coming to every appointment and participating in Lupus support groups and research studies, they tend to be the patients who are more likely to be vaccinated from my experience." (Transcript 2, physician)
	"But I think a lot of—almost everybody, almost every Black person I know knows somebody who died from COVID-19. People are more willing to get vaccinated because they know people who have died. I think it has a lot to do with what are people's expectations of how long they expect to live." (Transcript 14, community leader)
Strategies	"I think if you can incorporate real world examples, I think that will be really powerful for patients. I've had patients where I've said 'Look, I've gotten this vaccine, my family members—I've encouraged them all to get it. We've all done fine' and they say 'Ok, I'll get it. Promise me I'm not going to die.' I think people appreciate that, me sharing personal stories with that. I think people also really appreciate you letting them know that you care deeply about them and their outcomes. Some people are kind of angry when you bring up the vaccines and I say 'the only reason I'm talking to you about this is because I think it's really important. I care deeply about you and your health and I do not want to see you end up in a bad situation.'" (Transcript 2, physician)
	"They're still not inclined to be vaccinated. We stay in conversation. That's really the thing I can say. I think it becomes even less urgent as time goes by. I tell you the thing that was most important is that people felt respected, whatever their decision was, and we—that was really, really important to us. We are not going to isolate people and dog them if you didn't." (Transcript 17, community leader)
	"How do you convince someone who's totally dug in, right? But if it came from someone outside the medical field and someone they knew in another sphere and trusted, they certainly have a way better shot than I did." (Transcript 1, physician)
	"I think the benefit of getting the vaccine and learning about the vaccine. What are the pros and cons of this vaccine? And make it so they can understand it, not make it high tech, high scientific studies. Make it real." (Transcript 3, community leader)
Demographic specific	"I typically bring up the fact we are vaccinating patients with their particular condition in our clinics and those patients are doing fine and they're not having any problems. For example, if it's a patient with lupus I'll say, 'We've had many patients who have lupus who are seen in our clinic who have gotten the vaccine and they've done great and there's not been any problem relating to their lupus or any adverse effect from the vaccine.'" (Transcript 2, physician)
	"My impression is that when people are vaccine-hesitant or vaccine opposed because they see it as a dominating—another example of systemic oppression that allowing—respecting their opinion or showing respect and acknowledging some concerns—not validating things that are clearly not based in fact, but validating concerns like, 'I understand your concern about the pharmaceutical industry based on this that and the other. I understand that vaccines have been used in various studies against Black individuals unethically.' I think that acknowledging those things is important." (Transcript 11, physician)

Table 5. Considerations for POL intervention design—selected quotes from physicians and community leaders*

Considerations	Transcript
POL traits	"I think in this particular sphere, being a great leader is really important. In order to be successful in changing people's minds, you have to be really able to listen to what people are saying and interpret that. I would say having an outgoing personality but a sensitive approach. Meaning you're not bombarding a patient with a bunch of information and you're approaching it softly is what I would say and great people skills. Of course, that is part of being a community leader as well." (Transcript 2, physician) "I think they need to be humble. They need to be approachable and honest. If you break down telling your story, it's okay. If you say I don't know the answer, I'll get back to you, it's totally okay. If you give example[s] of what you've gone through, people can relate. So they can say 'I know somebody, or I've been through that.'" (Transcript 3, community leader)
Physician knowledge of patient social networks	"I think I would have to say I get a sense if they are isolated or not, but I don't know the details of their social network. I just know who I don't know." (Transcript 1, physician) "Some of my patients I know a lot about their families because they tell me, and they're more open about it and maybe I've seen them for longer and we just have that relationship. Other patients I don't know much about their life outside of their health." (Transcript 2, physician)
Concerns about being a POL	"I think for me, the hump that I have to get over is it's tiring, it becomes very tiring. It's depressive. It puts me in a state, and I think for me what I have to do is change my mindset because I'm at this point I can't do it anymore cause I'm so tired of the stupidity. It's wearing, it takes its toll on you." (Transcript 5, community leader)

POL = Popular Opinion Leader.

voiced the importance of being knowledgeable about vaccine development, clinical trial data, and side effects. Many noted that community leaders could play an important role engaging target populations.

Physician knowledge of patient social networks. Identifying the social networks of potential POLs is integral to creating an outreach social network for future vaccine dissemination efforts. When asked, physicians noted this knowledge was obtained through direct questioning, especially for pediatric patients, through patient-volunteered information based on patient–health care worker rapport, or by proxy of perceived coping mechanisms/supports. Many acknowledged that this knowledge is not consistently understood, with one stating, "How often are [physicians] digging into their social situation or their life outside their health?… Some physicians do that more so than others" (Transcript 2, physician).

POL concerns. While not specifically asked, 2 community leaders offered a perspective regarding potential concerns about being a POL promoting vaccine uptake including repeated exposure to unvaccinated individuals while being immunocompromised, fear of losing or damaging personal relationships, and ethics considerations about promoting vaccination to staunch opponents to vaccination. One shared, "You would think that the relationship with this person that says, 'I will get sick if you do not do this,' would be compelling, but that's also a very traumatizing thing to say…especially if that doesn't have an outcome that's positive 'cause…that person doesn't care" (Transcript 16, community leader). Considering the well-being of the POLs was highlighted as central to plans for a future intervention.

DISCUSSION

With the disproportionate burden of COVID-19 infection, adverse outcomes, and vaccine hesitancy among people of color with rheumatic diseases, this qualitative study sought to better elucidate perceived barriers and facilitators to vaccine uptake in Black communities in Chicago and Boston (11). Both physician- and community-based participants identified sources of vaccine hesitancy including mistrust, apathy, inaccessibility, safety concerns, inconsistent messaging, and misinformation. Mistrust was noted to be a strong barrier buoyed by conspiracy theories, politicized skepticism, and mistrust of health care systems and vaccine development and function, along with historical injustices and racism. Approaches for increasing vaccine uptake centered on information dissemination from trusted and empathetic social network members with an emphasis on lifestyle benefits gained from vaccination.

Our findings regarding barriers to vaccination among Black individuals with rheumatic conditions are congruent with previous studies among the general Black population (21–25). A recent qualitative study of Black churchgoers from a single congregation in Boston found lack of trust, rushed development, fear of side effects, history of medical mistreatment, and a perception of low risk as reasons for vaccine hesitancy in this group (26). Furthermore, a qualitative study assessed vaccine hesitancy factors for 70 members of racial and ethnic minority communities at high risk for COVID-19 and found misinformation, politicization, apprehension based on historical inequities, access barriers, and a need for trusted messengers to be shared sentiments (27). Additionally, studies in patients with SLE and other rheumatic conditions highlight concerns regarding vaccine safety, side effects, and impact on underlying rheumatic diseases in patients with vaccine hesitancy, similar to our findings (28–34). Our study additionally explored COVID-19 vaccine booster hesitancy, highlighting inconsistent messaging, waning precaution vigilance, fatigue, fear of prior or more severe side effects, along with the perception of boosters as nonessential.

This qualitative study explored perceived barriers by physicians, community leaders, and established POLs. Limitations

include a small sample size, although thematic saturation was reached, with the largest percentage of participants being female and from Boston. Black communities are not monolithic, and one set of perceived barriers or strategies will not be generalizable to all individuals identifying as Black, as evidenced by the heterogeneity reflected in participant responses from 2 US cities. While our interviewees were made aware of the overarching study aims (to understand and increase COVID-19 vaccinations/boosters overall and specifically among Black individuals with rheumatic conditions), and individuals identifying as Black/African American were oversampled (10 of 18 interviews), perspectives presented may not only represent views of individuals in these communities, and at the same time, may not be broadly applicable both within and outside of Black communities. We specifically included leaders of organizations and physicians serving Black communities as well as community leaders who identified with various Black cultural backgrounds to incorporate diverse life experiences.

We aimed to represent the community of individuals with rheumatic/musculoskeletal conditions (12 of 18 interviews), but also to have a broader understanding of perspectives from individuals who interact with this community. Therefore, some quotes may be representative of more general opinions beyond those specific to individuals with these conditions. We did not choose to ask participants if they were fully vaccinated, though many community leaders spoke of their experiences receiving the COVID-19 vaccine, and physicians were required per workplace mandates. Additionally, the interviews were conducted virtually from November 2021 to October 2022 and thus the state of vaccination efforts, messaging, and access varied during this time.

There are strengths and limitations inherent to our use of qualitative over quantitative methods. Quantitative studies can reveal population-level data on vaccine uptake and potentially reasons for vaccine hesitancy. However, they do not allow for in-depth, nuanced data regarding reasons behind vaccine hesitancy. While surveys may have provided information from a larger number of participants, the in-depth exploration of experiences to facilitate a future intervention would not have been possible. While we considered holding focus groups instead of individual interviews, we chose the latter to elevate the unique perspective of each individual on a sensitive topic, and to allow us to accommodate individuals working tirelessly in the midst of ongoing COVID-19 waves who would have been overlooked with a longer focus group meeting at a designated time (35).

Generalized mistrust, conspiracy theories, mistrust of health care systems and vaccine development and function, politicized skepticism, racism, and historical injustices were the most cited barriers to vaccine uptake. A national survey showed the importance of trust in the vaccine development process; vaccine intent was 75% higher among those with high trust versus those with low trust (36). Furthermore, another national analysis highlighted the role of structural racism, as Black respondents were more likely to have vaccine hesitancy while controlling for personal traits like stress, conspiracy thinking, and medical trust (37). Addressing structural racism through informed interventions that acknowledge past and current racism and mistreatment could help improve vaccine uptake in Black communities. Our findings suggest that trust restoration must be at the forefront of strategies to improve racial equity in vaccine uptake through community-driven interventions, meaningful investment, and the building of community power through direct engagement with Black communities (38,39).

The experiences and perspectives of participants will be used to develop a curriculum to train POLs on COVID-19 vaccine and booster safety and hesitancy in communities of Black individuals with rheumatic conditions with a racial justice lens. We hope these perspectives will help efforts to reduce mistrust and fear. In building this intervention, our key informants highlighted that POLs should engage in personalized, iterative conversations, employing humility, empathy, and curiosity. Care must be taken to avoid creating adverse effects related to moral injury for trusted messengers. This qualitative work will directly inform a planned randomized controlled trial to assess whether this community–academic partnership intervention rooted in racial justice with acknowledgement of structural racism will improve COVID-19 vaccine and booster uptake among Black individuals with rheumatic diseases.

ACKNOWLEDGMENTS

We would like to first acknowledge the community leaders and physicians for their interview participation. We greatly appreciate their reflections and insights. We would also like to thank Gina Curry, MBA, MPH, Richard Joseph, MD, MBA, and Elmer Freeman, MSW for their invaluable input. We acknowledge our community-based partnership organizations for their help recruiting participants and connecting us with other organizations including the Center for Community Health Education Research, Sportsmen's Tennis & Enrichment Center, Mission Hill Health Movement, AllianceChicago, Women of Courage, The Labalaba Foundation, and the Boston Public Health Commission.

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

Study conception and design. Williams, Ojikutu, York, Crespo-Bosque, Jean-Jacques, Mancera-Cuevas, Milaeger, Losina, Dhand, Son, Ramsey-Goldman, Feldman.

Acquisition of data. Sirek, Ulysse, Feldman.

Analysis and interpretation of data. Ezeh, Sirek, Ulysse, Williams, Chandler, Roberson, Ramsey-Goldman, Feldman.

REFERENCES

1. Kriss JL, Hung MC, Srivastav A, et al. COVID-19 vaccination coverage, by race and ethnicity: national immunization survey adult COVID

module, United States, December 2020–November 2021. MMWR Morb Mortal Wkly Rep 2022;71:757–63.
2. Fast HE, Murthy BP, Zell E, et al. Booster COVID-19 vaccinations among persons aged ≥5 years and second booster COVID-19 vaccinations among persons aged ≥50 years: United States, August 13, 2021–August 5, 2022. MMWR Morb Mortal Wkly Rep 2022;71:1121–5.
3. Centers for Disease Control and Prevention. COVID data tracker: trends in number of COVID-19 vaccinations in the US. 2023. URL: https://covid.cdc.gov/covid-data-tracker/#vaccination-trends.
4. Dong L, Bogart LM, Gandhi P, et al. A qualitative study of COVID-19 vaccine intentions and mistrust in Black Americans: recommendations for vaccine dissemination and uptake. PLoS One 2022;17:e0268020.
5. Scharff DP, Mathews KJ, Jackson P, et al. More than Tuskegee: understanding mistrust about research participation. J Health Care Poor Underserved 2010;21:879–97.
6. Merschel M, American Heart Association. For many Hispanic people, vaccination worries are a matter of trust. 2021. URL: https://www.heart.org/en/news/2021/09/15/for-many-hispanic-people-vaccination-worries-are-a-matter-of-trust.
7. Bailey ZD, Krieger N, Agénor M, et al. Structural racism and health inequities in the USA: evidence and interventions [review]. Lancet 2017;389:1453–63.
8. Thunström L, Ashworth M, Finnoff D, et al. Hesitancy toward a COVID-19 vaccine. Ecohealth 2021;18:44–60.
9. Van Scoy LJ, Snyder B, Miller EL, et al. Public anxiety and distrust due to perceived politicization and media sensationalism during early COVID-19 media messaging. J Commun Healthc 2021;14:193–205.
10. Larson H. Blocking information on COVID-19 can fuel the spread of misinformation. Nature 2020;580:306.
11. Gianfrancesco MA, Leykina LA, Izadi Z, et al. Association of race and ethnicity with COVID-19 outcomes in rheumatic disease: data from the COVID-19 Global Rheumatology Alliance Physician Registry. Arthritis Rheumatol 2021;73:374–80.
12. Curtis JR, Johnson SR, Anthony DD, et al. American College of Rheumatology guidance for COVID-19 vaccination in patients with rheumatic and musculoskeletal diseases: version 3. Arthritis Rheumatol 2021;73:e60–75.
13. Figueroa-Parra G, Gilbert EL, Valenzuela-Almada MO, et al. Risk of severe COVID-19 outcomes associated with rheumatoid arthritis and phenotypic subgroups: a retrospective, comparative, multicentre cohort study. Lancet Rheumatol 2022;4:e765–74.
14. Felten R, Dubois M, Ugarte-Gil MF, et al. Vaccination against COVID-19: expectations and concerns of patients with autoimmune and rheumatic diseases. Lancet Rheumatol 2021;3:e243–5.
15. Furer V, Rondaan C, Heijstek M, et al. Incidence and prevalence of vaccine preventable infections in adult patients with autoimmune inflammatory rheumatic diseases (AIIRD): a systemic literature review informing the 2019 update of the EULAR recommendations for vaccination in adult patients with AIIRD. RMD Open 2019;5:e001041.
16. Arneson LC, Taber KA, Williams JN, et al. Use of popular opinion leader models to disseminate information about clinical trials to Black individuals with lupus in two US cities. Arthritis Care Res (Hoboken) 2023;75:44–52.
17. Kelly JA, St Lawrence JS, Diaz YE, et al. HIV risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. Am J Public Health 1991;81:168–71.
18. Phillip CR, Mancera-Cuevas K, Leatherwood C, et al. Implementation and dissemination of an African American popular opinion model to improve lupus awareness: an academic-community partnership. Lupus 2019;28:1441–51.
19. Glaser BG. The constant comparative method of qualitative analysis. Soc Probl 1965;12:436–45.
20. Dedoose. Great research made easy. URL: www.dedoose.com.
21. Huang W, Dove-Medows E, Shealey J, et al. COVID-19 vaccine attitudes among a majority Black sample in the Southern US: public health implications from a qualitative study. BMC Public Health 2023;23:88.
22. Warren AM, Bennett MM, da Graca B, et al. Intentions to receive COVID-19 vaccines in the United States: sociodemographic factors and personal experiences with COVID-19. Health Psychol 2023;42:531–40.
23. Callaghan T, Moghtaderi A, Lueck JA, et al. Correlates and disparities of intention to vaccinate against COVID-19. Soc Sci Med 2021;272:113638.
24. Latkin CA, Dayton L, Yi G, et al. Trust in a COVID-19 vaccine in the U.S.: a social-ecological perspective. Soc Sci Med 2021;270:113684.
25. Bogart LM, Ojikutu BO, Tyagi K, et al. COVID-19 related medical mistrust, health impacts, and potential vaccine hesitancy among Black Americans living with HIV. J Acquir Immune Defic Syndr 2021;86:200–7.
26. Sekimitsu S, Simon J, Lindsley MM, et al. Exploring COVID-19 vaccine hesitancy amongst Black Americans: contributing factors and motivators. Am J Health Promot 2022;36:1304–15.
27. Carson SL, Casillas A, Castellon-Lopez Y, et al. COVID-19 vaccine decision-making factors in racial and ethnic minority communities in Los Angeles, California. JAMA Netw Open 2021;4:e2127582.
28. McMaster C, Liew DF, Lester S, et al. COVID-19 vaccine hesitancy in inflammatory arthritis patients: serial surveys from a large longitudinal national Australian cohort. Rheumatology (Oxford) 2023;62:1460–6.
29. Mohanasundaram K, Santhanam S, Natarajan R, et al. Covid-19 vaccination in autoimmune rheumatic diseases: a multi-center survey from southern India. Int J Rheum Dis 2022;25:1046–52.
30. Li YK, Lui MP, Yam LL, et al. COVID-19 vaccination in patients with rheumatic diseases: vaccination rates, patient perspectives, and side effects. Immun Inflamm Dis 2022;10:e589.
31. Boekel L, Hooijberg F, van Kempen ZL, et al. Perspective of patients with autoimmune diseases on COVID-19 vaccination. Lancet Rheumatol 2021;3:e241–3.
32. Priori R, Pellegrino G, Colafrancesco S, et al. SARS-CoV-2 vaccine hesitancy among patients with rheumatic and musculoskeletal diseases: a message for rheumatologists. Ann Rheum Dis 2021;80:953–4.
33. Tang Q, Li F, Tian J, et al. Attitudes towards and safety of the SARS-CoV-2 inactivated vaccines in 188 patients with systemic lupus erythematosus: a post-vaccination cross-sectional survey. Clin Exp Med 2023;23:457–63.
34. Shaw YP, Hustek S, Nguyen N, et al. Rheumatic disease patient decision-making about COVID-19 vaccination: a qualitative analysis. BMC Rheumatol 2022;6:76.
35. RAND Corporation. Harrell MC, Bradley MA. Data collection methods: semi-structured interviews and focus groups. 2009. URL: https://www.rand.org/pubs/technical_reports/TR718.html.
36. Savoia E, Piltch-Loeb R, Goldberg B, et al. Predictors of COVID-19 vaccine hesitancy: socio-demographics, co-morbidity, and past experience of racial discrimination. Vaccines (Basel) 2021;9:767.
37. Stoler J, Enders AM, Klofstad CA, et al. The limits of medical trust in mitigating COVID-19 vaccine hesitancy among Black Americans. J Gen Intern Med 2021;36:3629–31.
38. Ojikutu BO, Stephenson KE, Mayer KH, et al. Building trust in COVID-19 vaccines and beyond through authentic community investment. Am J Public Health 2021;111:366–8.
39. Warren RC, Forrow L, Hodge DA Sr, et al. Trustworthiness before trust: Covid-19 vaccine trials and the Black community. N Engl J Med 2020;383:e121.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2519–2528
DOI 10.1002/acr.25164
© 2023 American College of Rheumatology. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

American College of Rheumatology
Empowering Rheumatology Professionals

Life Events, Caregiving, and Risk of Autoimmune Rheumatic Diseases in the Women's Health Initiative Observational Study

Christine G. Parks,[1] Mary Pettinger,[2] Anneclaire J. de Roos,[3] Hilary A. Tindle,[4] Brian T. Walitt,[5] and Barbara V. Howard[6]

Objective. Growing evidence suggests psychosocial stressors may increase risk of developing autoimmune disease. We examined stressful life events and caregiving in relation to incident rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) in the Women's Health Initiative Observational Study cohort.

Methods. The sample of postmenopausal women included 211 incident RA or SLE cases reported within 3 years after enrollment, confirmed by use of disease-modifying antirheumatic drugs (i.e., probable RA/SLE), and 76,648 noncases. Baseline questionnaires asked about life events in the past year, caregiving, and social support. We used Cox regression models to calculate hazard ratios (HR) and 95% confidence intervals (95% CIs), adjusting for age, race/ethnicity, occupational class, education, pack-years of smoking and BMI.

Results. Incident RA/SLE was associated with reporting 3 or more life events (e.g., age-adjusted HR 1.70 [95% CI 1.14, 2.53]; *P* for trend = 0.0026). Elevated HRs were noted for physical (HR 2.48 [95% CI 1.02, 6.04]) and verbal (HR 1.34 [0.89, 2.02]) abuse (*P* for trend = 0.0614), 2 or more interpersonal events (HR 1.23 [95% CI 0.87, 1.73]; *P* for trend = 0.2403), financial stress (HR 1.22 [95% CI 0.90, 1.64]), and caregiving 3 or more days per week (HR 1.25 [95% CI 0.87, 1.81]; *P* for trend = 0.2571). Results were similar, excluding women with baseline symptoms of depression or moderate-to-severe joint pain in the absence of diagnosed arthritis.

Conclusion. Our findings support the idea that diverse stressors may increase risk of developing probable RA or SLE in postmenopausal women, supporting the need for further studies in autoimmune rheumatic diseases, including childhood adverse events, life event trajectories, and modifying psychosocial and socioeconomic factors.

INTRODUCTION

Autoimmune rheumatic diseases, including the 2 most common systemic autoimmune diseases, rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE), affect as many as 1.6 million adults in the U.S. (1). Sometimes co-occurring, RA and SLE share risk factors including female sex and family history of RA or SLE, while RA is increased in older women and SLE predominately affects reproductive age women (2–5). Characterized by complex etiologies involving environmental and genetic factors (6–8), only a few modifiable risk factors have been identified besides smoking.

A broad literature supports the idea that stress may play a role in worsening RA and SLE symptoms and outcomes, and growing evidence suggests stress may trigger a variety of autoimmune diseases (9, 10). Recent studies show that history of trauma or posttraumatic stress disorder is associated with risk of developing RA or SLE (11–18). Given the known effects of stress on the immune system, such as immunosuppression and inflammation (19–21), these findings support a broader hypothesis that other stressors

Supported in part by the intramural program of the NIH/National Institute of Environmental Health Sciences (grant Z01-ES-049028) and the Division of Intramural Research NIH/National Institute of Neurological Diseases and Stroke. The Women's Health Initiative program is supported by the NIH/National Heart, Lung, and Blood Institute and the US Department of Health and Human Services (contracts N01-WH-22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-32, and 44221).

[1]Christine G. Parks, PhD: NIH/National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina; [2]Mary Pettinger, MS: Fred Hutchinson Cancer Center, Seattle, Washington; [3]Anneclaire J. de Roos, PhD: Drexel University Dornsife School of Public Health, Philadelphia, Pennsylvania; [4]Hilary A. Tindle, MD, MPH: Vanderbilt University Medical Center and Veterans Affairs Tennessee Valley Healthcare System, Nashville, Tennessee; [5]Brian T. Walitt, MD, MPH: NIH/National Institute of Nursing Research, Bethesda, Maryland; [6]Barbara V. Howard, PhD: Georgetown-Howard Universities Center for Clinical and Translational Science, Washington, DC, and MedStar Health Research Institute, Hyattsville, Maryland.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25164.

Address correspondence via email to Christine G. Parks, PhD, at parks1@mail.nih.gov.

Submitted for publication July 15, 2022; accepted in revised form May 23, 2023.

SIGNIFICANCE & INNOVATIONS

- In a well-characterized cohort of postmenopausal women, those reporting 3 or more major life events in the past year at baseline had a 70% increased risk of developing rheumatoid arthritis (RA) or systemic lupus erythematosus (SLE) in the subsequent 3 years.
- Associations were not confounded by socio-demographic and lifestyle factors and were robust to sensitivity analyses excluding women with baseline depressive symptoms or unexplained moderate-to-severe joint pain.
- Our findings for diverse stressors, ranging from interpersonal factors to financial stress to abuse, add to a growing literature on the role of psychosocial stressors in the development of RA and SLE.

may increase risk of RA/SLE. Stressful life events can include acute experiences such as death of a partner or a serious accident, abuse, interpersonal conflicts such as marital separation, or financial stress, with some experiences reflecting traumatic stressors, patterns of longer-term experiences and chronic stress, or compounding trajectories of stressors across the lifespan (22). Research on stressful life events and RA/SLE is limited; one large case-control study in Sweden reported stressful life events in the 5 years prior to diagnosis were associated with an increased odds of developing RA (23), while a smaller study of SLE in southern Sweden showed no associations with life events in the past year (24).

Individuals experience external stressors within a broad socioeconomic and psychological, and social context. The Reserve Capacity model posits that lower socioeconomic status (SES) modifies the impact of stressors, undermining tangible, interpersonal and intrapersonal resources for coping, such as social support, amplifying over time the potential for adverse effects on health (25, 26). Past research in the Women's Health Initiative (WHI) found that having more negative life events was related to lower education, non-white race, lower social support, and adverse health behaviors (e.g., smoking, BMI) (27, 28). While the latter may have direct physiological impacts, the others may reflect differential reserve and resources, which may modify the impact of life events on health, including autoimmunity, inflammation, and disease (Figure 1). Race/ethnicity and age may further contextualize this relationship. In the WHI Extension Study (mean age 77 years), younger women and those of Black or African American (versus White) race/ethnicity reported higher resiliency (29).

In the WHI Observational Study (OS) cohort, we previously noted associations of SES-related covariates with risk of RA (i.e., non-professional occupation) and SLE (i.e., lower education) (30). Here we investigated whether risk of RA/SLE in the first three years of follow-up was associated with recent life events and caregiving, another potential stressor associated with depressive symptoms in the WHI (31). We hypothesized that having more stressful life events and caregiving might contribute to risk of developing RA or SLE. We also examined specific types of life events and in secondary analyses explored potential differences in associations with RA/SLE by age and indicators of Reserve Capacity (occupational class, social support).

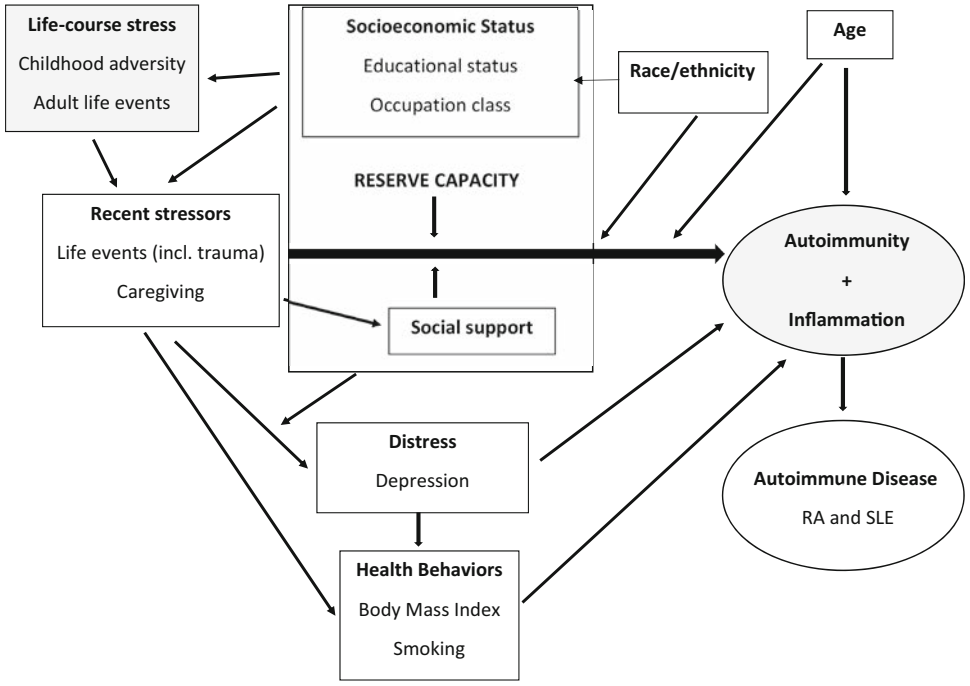


Figure 1. Life-course and recent stressors, reserve capacity, and pathways contributing to the development of autoimmune disease. *RA = rheumatoid arthritis; SLE = systemic lupus erythematosus.

MATERIALS AND METHODS

Study sample. The WHI-OS cohort includes 93,676 women (enrolled 1994-98) from 40 clinical centers across the United States (32). At baseline and year 3 of follow-up, women were asked if a doctor ever told them they had systemic lupus erythematosus (SLE), or arthritis, and if so, what type, i.e., rheumatoid arthritis (not rheumatism) or other/don't know. For the current study, eligible participants were those with complete data on RA/SLE status [exclude missing, N = 2,429 (2.6%)], medication use [exclude missing, N = 13,871 (14.8%)], the life events scale and major covariates [exclude missing, N = 842 (0.9%)]. The analysis sample excluded prevalent RA or SLE cases who used disease modifying anti-rheumatic drugs (DMARDs; N = 815, 0.9%) and potential cases with either RA or SLE without DMARDS or DMARDS without RA/SLE at baseline or follow-up (N = 2,533; 2.7%). Probable cases were identified based on a new self-reported diagnosis during the first three years of follow-up, confirmed by DMARD use at year 3 (a highly specific method for case ascertainment in the absence of medical records review or physician validation) (33). Derivation of the analysis sample and characteristics of the study sample relative to the complete OS Cohort are shown (see Supplementary Figure 1 and Supplementary Table 1, respectively, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25164). The final analysis sample of 74,897 women included 211 cases of probable incident RA (N = 176), SLE (N = 27), or both (N = 8) and 74,686. The institutional review boards of the participating institutions approved protocols and consent forms, which were signed by the women at enrollment.

Stress and covariates. Baseline questionnaires included a standardized life events scale, asking about 11 items representing external stressors in the past year, i.e., "hard things that sometimes happen to people" (34). We used the total count of events reported (0, 1, 2, 3, 4+ for descriptive frequencies, and collapsed categories to reduce the impact of small cell size in modeling and secondary analyses). We also considered different types: interpersonal (6 items: e.g., spouse was deceased, close friend or family member deceased or serious illness, divorce or breakup, close friend/family member divorce, close friend/family member lost job or retired, major conflict with children or grandchildren), financial stress (1 item: major money problems), and abuse (2 items: physically or verbally abused by a family member or close friend). Physical abuse was infrequently reported in the absence of verbal abuse (27), and for multivariable modeling we grouped these as three categories (none, verbal abuse only, and physical abuse ± verbal abuse). Other items included death of a pet, and major accidents, disasters, muggings, unwanted sexual experiences, robberies, or similar events.

Baseline questionnaires also asked about recent caregiving, a potential source of chronic stress: questions asked whether a woman was regularly providing care for an ill relative or friend, and how many days per week in the past 4 weeks, which we grouped into 3 categories (none, up to 2 days, 3 or more days per week) (35). Social support was assessed through the general social support index (including emotional, tangible, affection, and positive social interactions), which we dichotomized as lower (≤median) and higher (>median) (36). Depressive symptoms in the past 2 weeks were assessed by the modified CESD-6 scale, dichotomized using a cut-point of ≥0.06 (37). Joint pain or stiffness in the past 4 weeks was rated as none, mild, moderate, severe. Other covariate data included age, self-reported race/ethnicity, education, occupational class, pack-years of smoking, and body mass index (BMI).

Analyses. We modeled risk of developing RA/SLE using Cox proportional hazards regression to calculate hazard ratios (HR) and 95% confidence intervals (95% CIs). Initial models adjusted for age, while fully adjusted models also included race/ethnicity, occupational class, education, pack-years of smoking, and BMI. No evidence of confounding was seen in fully adjusted models, so age-adjusted results are reported in the text. Proportionality was assessed by adding a term for the interaction between the exposure and log-transformed survival time (*P* values: 0.6455 total life events, 0.1999 interpersonal events, 0.0466 financial stress, 0.6681 for abuse, and 0.6730 caregiving). Graphs of the survival function versus log-transformed survival time were also examined, showing the departure from proportionality for financial stress was based on a small number of cases at 1 extreme. Trends tests for life event and interpersonal event counts were based on linear terms in the models.

In sensitivity analyses we excluded women with symptoms that could indicate preclinical or undiagnosed RA (i.e., moderate-to-severe joint pain without doctor diagnosed arthritis or missing data on arthritis: 5,343 noncases, 35 cases). We also ran models excluding women with depressive symptoms (8,859 noncases, 44 cases), which may be a cause or effect of chronic inflammation, autoimmunity, and response to pain, and may also result from earlier life events and past trauma (38–41).

In secondary analyses, we explored potential modifiers of the relationship of RA/SLE risk with life events, including age, social support, and occupational class. Interaction *P* values were derived by comparing models including both variables with and without cross-product term; the test statistic, based on the deviance method, was tested against a chi-square distribution (degrees of freedom = difference in degrees of freedom between the 2 models) with a $P < 0.10$ for statistical significance, allowing a higher type 1 error rate to accommodate the lower power for testing interactions.

RESULTS

Table 1 shows that, across categories of the number of life events reported (0, 1, 2, 3, and 4+), women who reported more life events or stressors were slightly younger and had lower social support. They were more likely to be Black or Hispanic, have a

Table 1. Characteristics of study participants by number of stressful life events at baseline*

	0 events reported (n = 16,836)	1 event reported (n = 23,948)	2 events reported (n = 17,452)	3 events reported (n = 9,527)	≥4 events reported (n = 7,134)
Age, mean ± SD years	64.3 ± 7.2	64.1 ± 7.3	63.3 ± 7.3	62.6 ± 7.2	61.6 ± 7.2
Social support index score, mean ± SD	37.6 ± 7.2	37.0 ± 7.3	36.0 ± 7.6	34.9 ± 7.9	32.8 ± 8.3
Race/ethnicity					
White	14,863 (88.3)	20,966 (87.5)	14,849 (85.1)	7,905 (83.0)	5,477 (76.8)
Black	752 (4.5)	1,378 (5.8)	1,327 (7.6)	838 (8.8)	819 (11.5)
Hispanic	426 (2.6)	563 (2.4)	502 (2.9)	358 (3.8)	441 (6.2)
American Indian	43 (0.3)	66 (0.3)	59 (0.3)	40 (0.4)	61 (0.9)
Asian/Pacific Islander	571 (3.4)	704 (2.9)	482 (2.8)	233 (2.4)	203 (2.8)
Unknown	181 (1.1)	271 (1.1)	233 (1.3)	153 (1.6)	133 (1.9)
Education					
<High school graduate	565 (3.4)	806 (3.4)	697 (4.0)	441 (4.7)	431 (6.1)
High school/GED	2,665 (15.9)	3,761 (15.8)	2,711 (15.6)	1,470 (15.6)	1,075 (15.2)
School after high school	5,478 (32.8)	8,258 (34.7)	6,431 (37.1)	3,695 (39.1)	2,963 (41.9)
≥College degree	8,007 (47.9)	10,962 (46.1)	7,491 (43.2)	3,834 (40.6)	2,598 (36.8)
Occupation					
Managerial/professional	7,507 (46.4)	10,639 (46.2)	7,496 (44.7)	3,951 (43.2)	2,714 (40.0)
Technical/sales/admin	4,386 (27.1)	6,470 (28.1)	4,858 (28.9)	2,735 (29.9)	2,062 (30.4)
Service/labor	2,335 (14.4)	3,503 (15.2)	2,762 (16.5)	1,674 (18.3)	1,460 (21.5)
Homemaker only	1,951 (12.1)	2,439 (10.6)	1,672 (10.0)	789 (8.6)	556 (8.2)
Body mass index (kg/m^2)					
<25	7,950 (47.8)	10,520 (44.4)	7,030 (40.8)	3,638 (38.6)	2,335 (33.1)
25 to <30	5,574 (33.5)	8,058 (34.0)	5,951 (34.5)	3,252 (34.5)	2,408 (34.1)
≥30	3,106 (18.7)	5,112 (21.6)	4,258 (24.7)	2,536 (26.9)	2,310 (32.8)
Pack years of smoking					
Never-smoker	8,918 (54.6)	12,310 (53.1)	8,791 (52.1)	4,689 (51.0)	3,488 (50.8)
<5	2,320 (14.2)	3,440 (14.9)	2,500 (14.8)	1,436 (15.6)	1,092 (15.9)
5 to <20	2,288 (14.0)	3,327 (14.4)	2,451 (14.5)	1,328 (14.4)	974 (14.2)
≥20	2,807 (17.2)	4,087 (17.6)	3,140 (18.6)	1,743 (19.0)	1,310 (19.1)
Specific life events†					
Friend/family member died	— (NA)	11,560 (48.4)	11,201 (64.3)	7,069 (74.3)	5,868 (82.4)
Financial stress	— (NA)	2,534 (10.6)	5,391 (30.9)	4,839 (50.9)	5,402 (75.9)
Major conflict	— (NA)	1,961 (8.2)	3,744 (21.5)	3,727 (39.2)	4,538 (63.7)
Abuse – verbal only	— (NA)	826 (3.4)	1,812 (10.4)	1,945 (20.4)	2,658 (37.3)
Major accident	— (NA)	637 (2.7)	1,234 (7.1)	1,175 (12.3)	1,859 (26.1)
Divorce of break-up	— (NA)	117 (0.7)	327 (1.9)	440 (4.6)	1,006 (14.1)
Spouse/partner died	— (NA)	339 (1.4)	628 (3.6)	548 (5.8)	694 (9.7)
Physical (± verbal abuse)‡	— (NA)	16 (0.1)	90 (0.5)	162 (1.7)	635 (8.9)
Caregiving					
None or infrequent	11,196 (66.8)	14,788 (62.0)	10,164 (58.6)	5,110 (54.0)	3,540 (49.9)
Up to 2 times/week	3,659 (21.8)	5,880 (24.7)	4,518 (26.0)	2,764 (29.2)	2,102 (29.6)
3 or more times/week	1,902 (11.4)	3,166 (13.3)	2,674 (15.4)	1,593 (16.8)	1,450 (20.4)
Depression symptoms§	577 (3.5)	1,527 (6.5)	1,910 (11.2)	1,501 (16.1)	1,991 (28.6)
Joint pain or stiffness					
Symptom did not occur	5,915 (35.2)	7,635 (32.0)	4,816 (27.7)	2,373 (25.0)	1,423 (20.0)
Mild	7,817 (46.6)	11,260 (47.1)	8,288 (47.6)	4,361 (45.9)	3,137 (44.2)
Moderate	2,592 (15.4)	4,040 (16.9)	3,352 (19.3)	2,112 (22.3)	1,813 (25.5)
Severe	459 (2.7)	960 (4.0)	940 (5.4)	646 (6.8)	729 (10.3)

* Values are the number (%) unless indicated otherwise.
† Common or severe events listed, including include financial stressor (i.e., major money problems); spouse or partner died, divorce or break-up; major conflict (i.e., with children or grandchildren); major accident (or disaster). Other included close friend/family member divorced, close friend/family member lost job or retired; pet died.
‡ Physical abuse occurred in the absence of reported verbal abuse for 8, 33, 45, and 72 women across the 4 event categories.
§ Depression symptoms if Center for Epidemiologic Studies Depression Scale score ≥0.06.

lower educational attainment and occupational class, greater BMI, and pack-years of smoking, and were more likely to report depressive symptoms or moderate-to-severe joint pain or stiffness. The most common stressor was death of a friend/family member (range 48–82% across categories 1 to 4+), followed by financial stress (11–76%), major conflict (8–64%), verbal abuse only (3%–37% [not including physical]), major accident (3–26%), divorce/breakup (1–14%), spouse/partner died (1–10%), and physical abuse (<1% to 9% [with or without verbal abuse]). Frequent caregiving (3 or more times per week) was reported by

11% and 13% of those with 0 or 1 life events, to 20% of those with 4+ events.

Table 2 shows frequencies of life events in the past year, caregiving days per week, and symptoms of depression and joint pain in RA/SLE cases and noncases. Across all life events evaluated, 12.8% of cases reported ≥4 events and 18.5% reported 3 events. Across the different types of events, 10% reported at least 3 interpersonal life events and 24.2% reported 2 events; 28.4% reported financial stress. Only 2.4% reported physical abuse, while 12.3% reported verbal abuse only. Caregiving 3 or more days per week was reported by 17.5% of cases. At baseline, 17.3% had symptoms of depression, and 8.1% reported moderate-to-severe joint pain in the absence of diagnosed arthritis. Similar frequencies were seen in RA cases, while frequencies in SLE cases are shown in Supplementary Table 2 (available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25164). Covariate frequencies are shown in Supplementary Table 3 (available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25164.

In proportional hazard regression models (Table 3), RA/SLE risk was associated with a greater number of recent life events or stressors (e.g., age-adjusted HR = 1.70 [95%CI 1.14, 2.53] for 3 or more versus none; *P* for trend 0.0026), and with physical abuse (e.g., HR = 2.48 [95%CI 1.02, 6.03]). Elevated HRs were

Table 2. Incident RA/SLE by stressful life events, caregiving, depressive symptoms, moderate-to-severe joint pain or stiffness, and doctor diagnosed arthritis other than RA*

	Non-cases (n = 74,686)		RA/SLE (n = 211)†		RA (n = 184)	
	No.	%	No.	%	No.	%
Life events (all)						
0	16,797	22.5	39	18.5	32	17.4
1	23,884	32.0	64	30.3	59	32.1
2	17,410	23.3	42	19.9	34	18.5
3	9,488	12.7	39	18.5	36	19.6
≥4	7107	9.5	27	12.8	23	12.5
Interpersonal						
0	22,895	30.7	60	28.4	47	25.5
1	29,581	39.6	79	37.4	73	39.7
2	15,581	20.9	51	24.2	46	25.0
≥3	6,629	8.9	21	10.0	18	9.8
Financial stress						
No	56,448	75.7	151	71.6	133	72.3
Yes	18,106	24.3	60	28.4	51	27.7
Missing	132	–	–	–	–	–
Abuse						
None	66,731	89.3	180	85.3	161	87.5
Verbal only	7,215	9.7	26	12.3	20	10.9
Physical +/− verbal	740	1.0	5	2.4	3	1.6
Caregiving (days/week)						
None/infrequent	44,678	60.1	120	56.9	103	56.0
Up to 2 times/week	18,869	25.4	54	25.6	49	26.6
3 or more times/week	10,748	14.5	37	17.5	32	17.4
Missing	391	–	0	–	–	–
Depression (CES-D ≥0.06)						
No	65,827	89.8	167	82.7	148	84.6
Yes	7,471	10.2	35	17.3	27	15.4
Missing	1,388	–	9	–	–	–
Joint pain/stiffness						
No symptoms	22,137	29.7	25	11.8	23	12.5
Mild	34,776	46.7	87	41.2	76	41.3
Moderate	13,838	18.6	71	33.6	62	33.7
Severe	3,706	5.0	28	13.3	23	12.5
Missing	229	–	–	–	–	–
Doctor-diagnosed arthritis‡	31,978	42.8	107	50.7	92	50.0
Moderate-severe joint pain/stiffness						
+Diagnosed arthritis	12,430	16.7	64	30.3	55	29.9
+No diagnosed arthritis	3,808	5.1	17	8.1	14	7.6
Missing arthritis	1,306	–	18	–	16	–

* CES-D = Center for Epidemiologic Studies Depression Scale; RA = rheumatoid arthritis; SLE = systemic lupus erythematosus.
† Total sample includes 8 cases of RA with SLE and 27 cases of SLE-only.
‡ Doctor-diagnosed arthritis other than RA, other nonspecified or unknown type.

Table 3. Associations of stressful life events or stressors with incident RA or SLE adjusting for age and covariates*

Stressful life events†	No. of noncases	No. of RA/SLE cases	HR (95% CI); age-adjusted‡	HR (95% CI); fully adjusted‡
All life events/stressors				
0	16,797	39	1.0	1.0
1 to 2	41,294	106	1.11 (0.77, 1.60)	1.16 (0.78, 1.73)
≥3	16,595	66	1.70 (1.14, 2.53)	1.75 (1.14, 2.69)
P for trend			*0.0026*	*0.0039*
Interpersonal				
0	22,895	60	1.0	1.0
1	29,581	79	1.02 (0.73, 1.43)	1.17 (0.82, 1.67)
≥2	22,210	72	1.23 (0.87, 1.73)	1.25 (0.86, 1.81)
P for trend			*0.2403*	*0.2465*
Financial stress				
No	56,448	151	1.0	1.0
Yes	18,106	60	1.22 (0.90, 1.64)	1.15 (0.83, 1.59)
P for trend			*0.2078*	*0.4020*
Abuse				
None	66,731	180	1.0	1.0
Verbal only	7,215	26	1.34 (0.89, 2.02)	1.36 (0.89, 2.10)
Physical (± verbal)	740	5	2.48 (1.02, 6.03)	2.50 (1.02, 6.14)
P for trend			*0.0614*	*0.0524*
Caregiving				
None	44,678	120	1.0	1.0
≤2 times/week	18,869	54	1.06 (0.77, 1.46)	1.04 (0.74, 1.47)
>2 times/week	10,748	37	1.25 (0.87, 1.81)	1.31 (0.89, 1.92)
P for trend			*0.2571*	*0.2115*

* RA = rheumatoid arthritis; SLE = systemic lupus erythematosus.
† Categories collapsed to reduce the variance in multivariable models.
‡ Hazard ratios (HR) and 95% confidence intervals (95% CIs) estimated from Cox proportional hazards regression models, adjusted for age, or fully adjusted for age, race/ethnicity, education, occupation, pack-years of smoking and body mass index.

seen for increased interpersonal events (HR 1.23 [95% CI 0.87, 1.73]), financial stress (HR 1.22 [95% CI 0.90, 1.64]), and more frequent caregiving (>2 days/week versus none [HR 1.25 [95% CI 0.87, 1.81]), though confidence limits did not exclude the null, and trend tests were not statistically significant. Estimates did not appear to be confounded by demographic or behavioral factors (race/ethnicity, education, occupation, pack-years of smoking, and BMI), as results were similar in fully adjusted models. In sensitivity analyses (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25164), associations of RA/SLE with the number of life events and physical abuse persisted after excluding women with moderate or severe joint pain or stiffness not due to diagnosed arthritis or depressive symptoms. The trend for more frequent caregiving became statistically significant after excluding women with moderate-to-severe joint pain in the absence of diagnosed arthritis (*P* for trend = 0.0424) but was attenuated excluding those with depressive symptoms.

In secondary analyses exploring potential modifiers (Figure 2 and Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25164), several associations were more apparent in women with lower occupational status (i.e., those with nonprofessional/managerial jobs or homemakers) or social support, or ages 50–64 years. Abuse showed strong, statistically significant associations with RA/SLE in women with lower occupational status, lower social support, and ages 50–64 years (odds ratios [ORs] 1.60–1.79), though interactions were not statistically significant. The association of financial stress with RA/SLE was also more apparent in women with a lower occupational status (*P* for interaction = 0.0147) and in women ages 50–64 years (*P* for interaction = 0.0814). By contrast, the association of RA/SLE with interpersonal events was more apparent in women ages ≥65 years (*P* for interaction = 0.0742).

DISCUSSION

Results of this prospective study support the hypothesis that stressful life events may play a role in the development of RA and SLE in postmenopausal women. We found that women reporting a greater number of events in the past year, including death of a partner or close friend, conflict with children or grandchildren, divorce, abuse (physical or verbal), financial problems, and major accidents or disasters, had a 70% increased risk of being diagnosed with probable RA or SLE within 3 years of enrollment. These results extend prior literature showing associations with diverse, contemporary psychosocial stressors in a population of middle-aged and older women.

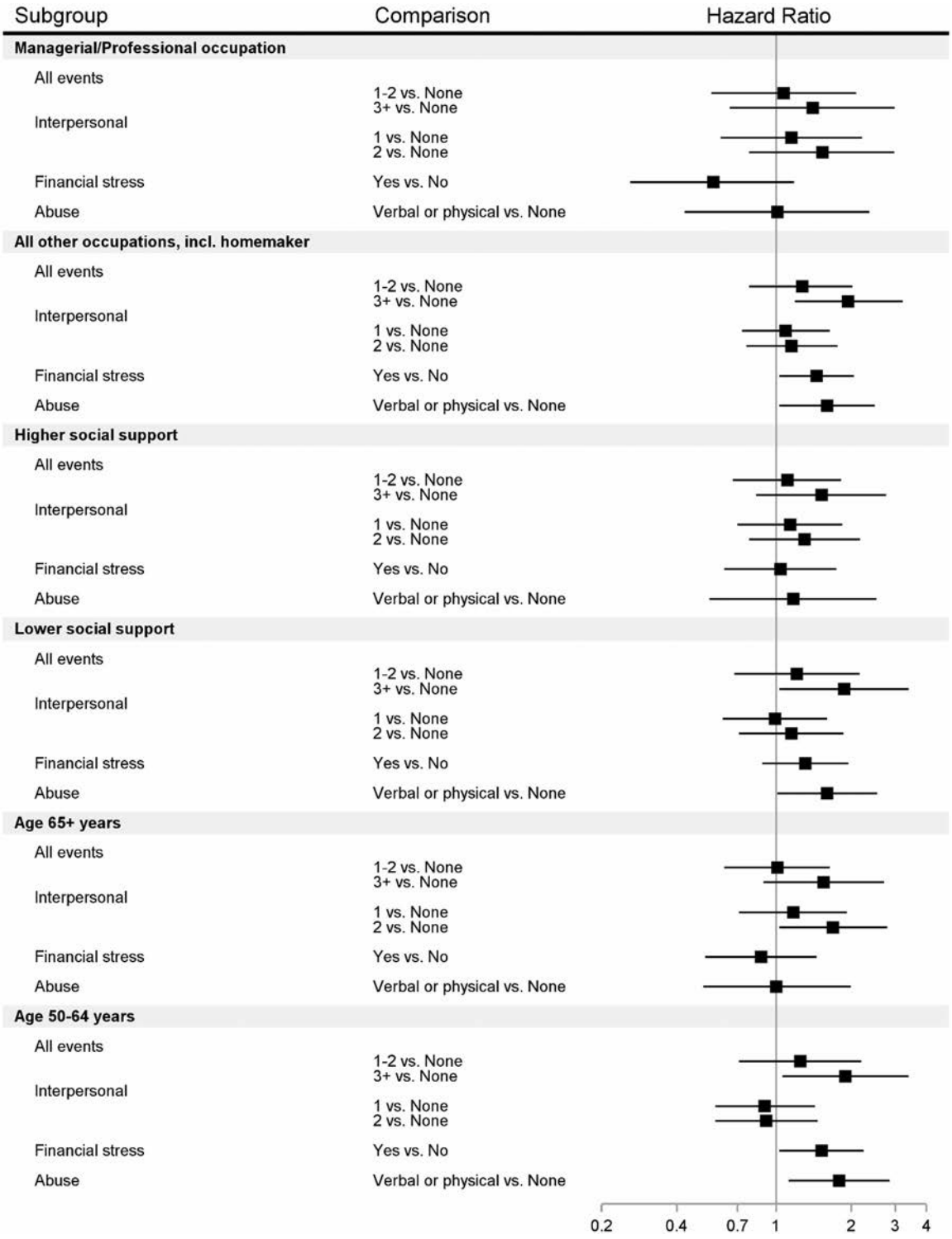


Figure 2. Life events associated with rheumatoid arthritis/systemic lupus erythematosus, stratified by social and demographic buffering factors; age-adjusted hazard ratios and 95% confidence intervals.

Our prospective findings for probable RA/SLE are consistent with a retrospective study of 2,774 RA cases in Sweden, which showed that having 3 or more life events in the past 5 years was associated with RA in women and with anti–citrullinated peptide antibody (ACPA)–negative cases (ORs 1.3–1.4), though their study sample was younger (median age 55 years [range 18–70 years]) and clinically validated (23). Looking at individual types of events, they saw stronger (OR ≥ 1.4), statistically significant associations for interpersonal conflict (with a spouse or children or at work), increased or decreased responsibility at work,

unemployment, change in residence, and divorce. In a retrospective analysis of 85 SLE cases, the strongest associations (though none of them statistically significant) included severe conflict or being deeply offended by someone, along with death of a child (24). In our overall sample, RA/SLE was not associated with increased interpersonal events (a group that included events such as divorce, death of a family member or friend, major conflict with children or grandchildren, loss of a job, or retirement), or with financial problems. Focused research is needed to address the impacts of stress from interpersonal conflict and loss.

While relatively uncommon, we found that physical abuse (with or without verbal abuse) in the past year was associated with risk of RA/SLE. Partner abuse is an important, understudied traumatic stressor in women, and these findings add to a growing literature focused on posttraumatic stress disorder (PTSD) and trauma as risk factors for RA and SLE (10–12,18,42). Previous research in the WHI has shown that abuse is related to depressive symptoms and overall mortality risk (43,44). Though focused on events in the past year, the question specifies abuse in a close relationship, which may also indicate an extended pattern of traumatic stress (45). We did not look at associations specifically in SLE cases, but note that 25.7% reported physical or verbal abuse compared to 12.5% of those with RA and 10.7% of noncases. Given the small number of exposed cases, these findings warrant cautious interpretation.

Our results suggest that overall caregiving was not associated with developing RA/SLE. We saw an elevated HR for more frequent caregiving (3 or more days per week), but the CI included the null. Notably, these results were sensitive to excluding women with baseline symptoms of joint pain (increased HR) or depression (attenuated HR). Caregiving is a common experience in mid-life and older women caring for parents or partners and is typically viewed as a source of chronic stress. But research on caregiving often fails to capture the context or heterogeneity of caregiving and any potential positive effects on health and wellbeing (46–49). Conversely, caregiving may be accompanied by life events such as illness or death of a spouse or relationship conflict. In the current study sample, women with more life events also reported more frequent caregiving. Caregiver stress may be better captured in future studies using methods such as a latent class analysis (50).

We explored the use of markers of Reserve Capacity (i.e., occupational class and social support) to contextualize the environment in which stressors occur, and age, as potential effect modifiers on the relationship of life events with RA/SLE. In stratified models, most differences were in the expected direction, with greater risk in women with fewer resources, i.e., lower occupational status and lower social support. Most did not reflect significant interactions (i.e., $P < 0.10$), except for financial stress, which was associated with increased RA/SLE in women with lower occupational status and in women ages <65 years. Having more interpersonal events was also associated with RA/SLE in those who were 65 and older, but not in women <65. These differences could be related to exposure heterogeneity across subgroups or variable response depending on life stages and co-occurring events. Cautious interpretation is warranted given the exploratory nature of these analyses.

Psychological stress or distress is experienced when demands outweigh psychological, social, and materials resources, with impacts on health through diverse pathways, including psychopathology. We did not adjust for depressive symptoms, which may be a marker for physiologic effects on the pathway to disease. Baseline depressive symptoms in the past 4 weeks were more common in women who later developed RA/SLE, however observed associations with life events persisted in women without symptoms. We did not examine diagnosed depression or antidepressant use, nor did we evaluate potential mediation or modification by depression, which warrants further consideration as it offers a potential opportunity in clinical settings to identify and support individuals at risk.

Stress can affect health through different biologic mechanisms, depending on the type or timing of the stressor, and other historical and concurrent exposures that modify neuroendocrine responses. Chronic and acute stress due to experiences viewed as threatening, unpredictable, or uncontrollable may lead to immune dysfunction and dysregulation, resulting in a variety of long- and short-term changes, including increased susceptibility to infection, reduced healing, and inflammation (51,52). Autoantibodies in RA (and other autoimmune diseases) can arise well in advance of disease onset as part of a larger constellation of factors, including systemic inflammation, preceding clinically apparent disease (32,53–55). Stress effects on the immune system could act at an earlier stage on the causal pathway leading to autoantibody production, for example, in the mucosal origins hypothesis for RA, or at a later stage contributing to clinical pathology leading to diagnosis (56). Time to diagnosis following initial symptoms varies and it is likely that some cases in our study had undiagnosed disease at baseline. Early symptoms of disease may include joint-specific inflammation; however, results were unchanged in models excluding those reporting moderate or severe joint pain at baseline not due to other or unknown forms of doctor-diagnosed arthritis.

This study has limitations. Although the sample was of sufficient size to detect modest associations with frequently reported life events, the low incidence of probable RA/SLE and short follow-up time limited our ability to examine less common individual exposures (e.g., physical abuse or experiencing an accident or disaster or another major traumatic event) or conduct analyses limited to SLE. The WHI-OS cohort is a volunteer sample of women who were ineligible for the clinical trials for various reasons; most were White, while Black and Hispanic women were more likely to be excluded from our study sample due to missing data, which further limited generalizability. The small number of non-White women in the study sample precluded analyses of racial/ethnic disparities in RA/SLE in relation to life events (57).

Self-report of RA/SLE is known to be nonspecific, so we limited our analyses to probable clinical cases based on their use of disease-specific medication. During our study period, the paradigm of early initiation of DMARDs following diagnosis was emerging, but not widespread, especially in older patients. Of those enrolled in Medicare Part D, only 24% of RA patients used DMARDs in 1996, rising to only 41% by 2003 (58, 59). Thus, our cases may represent those with more aggressive disease or otherwise better access to and uptake of DMARDs. We lacked data on disease phenotype, including anti–cyclic citrullinated peptide (anti-CCP) antibodies at RA diagnosis. Recent findings in male RA patients showed that PTSD symptoms were related to serum cytokine levels in those with anti-CCP antibodies (60), while a large retrospective study showed no associations of life events with anti-CCP negative RA in men, but few differences in associations with RA by anti-CCP status among women (23). Further research is warranted on the role of stress and RA/SLE in larger preclinical samples, including those with anti-CCP antibodies and individuals with a family history of RA/SLE.

Stress exposure assessment can be challenging in studies of autoimmune diseases, which themselves can be stressful or cause physical and psychological changes that may result in recall bias or reverse causality (39). Events were measured only at one time point, and no data were available on past adult life events, traumas, and childhood adverse experiences, which may proliferate and form trajectories impacting health in older women (61,62). Results for physical abuse, in particular, are likely to reflect a longer trajectory of abuse (physical and verbal) with cumulative effects on health. Repeated stress measures, also including perceived stress and psychological sequelae of traumatic stress, over a longer follow-up period, are needed to identify whether specific events versus sustained patterns of trauma or chronic stress confer increased risk of RA/SLE, especially given the potential latency of effects in the development and progression of disease.

We saw limited evidence of confounding by BMI and smoking, race/ethnicity, or socioeconomic factors (occupational status and education). We cannot rule out the possibility of unmeasured confounders.

In sum, the results of this prospective analysis highlight the possible role of stress due to recent life events as proximal risk factors for RA or SLE in postmenopausal women and support the need to consider a diverse range of stressors and contextual factors in future studies. If replicated, our findings also suggest opportunities to identify individuals who may be at higher risk for developing RA or SLE.

REFERENCES

1. Helmick CG, Felson DT, Lawrence RC, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part I. Arthritis Rheum 2008;58:15–25.
2. Tsao BP. The genetics of human systemic lupus erythematosus [review]. Trends Immunol 2003;24:595–602.
3. Lau CS, Yin G, Mok MY. Ethnic and geographical differences in systemic lupus erythematosus: an overview [review]. Lupus 2006;15:715–9.
4. Lockshin MD. Sex ratio and rheumatic disease: excerpts from an Institute of Medicine report [review]. Lupus 2002;11:662–6.
5. Alamanos Y, Drosos AA. Epidemiology of adult rheumatoid arthritis [review]. Autoimmun Rev 2005;4:130–6.
6. Woo JM, Parks CG, Jacobsen S, et al. The role of environmental exposures and gene-environment interactions in the etiology of systemic lupus erythematous [review]. J Intern Med 2022;291:755–78.
7. Klareskog L, Rönnelid J, Saevarsdottir S, et al. The importance of differences; on environment and its interactions with genes and immunity in the causation of rheumatoid arthritis [review]. J Intern Med 2020;287:514–33.
8. Gourley M, Miller FW. Mechanisms of disease: environmental factors in the pathogenesis of rheumatic disease [review]. Nat Clin Pract Rheumatol 2007;3:172–80.
9. Sharif K, Watad A, Coplan L, et al. The role of stress in the mosaic of autoimmunity: an overlooked association [review]. Autoimmun Rev 2018;17:967–83.
10. Dube SR, Fairweather D, Pearson WS, et al. Cumulative childhood stress and autoimmune diseases in adults. Psychosom Med 2009;71:243–50.
11. Boscarino JA. Posttraumatic stress disorder and physical illness: results from clinical and epidemiologic studies. Ann N Y Acad Sci 2004;1032:141–53.
12. Boscarino JA, Forsberg CW, Goldberg J. A twin study of the association between PTSD symptoms and rheumatoid arthritis. Psychosom Med 2010;72:481–6.
13. Lee YC, Agnew-Blais J, Malspeis S, et al. Post-traumatic stress disorder and risk for incident rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:292–8.
14. Roberts AL, Malspeis S, Kubzansky LD, et al. Association of trauma and posttraumatic stress disorder with incident systemic lupus erythematosus in a longitudinal cohort of women. Arthritis Rheumatol 2017;69:2162–9.
15. Song H, Fang F, Tomasson G, et al. Association of stress-related disorders with subsequent autoimmune disease. JAMA 2018;319:2388–400.
16. Bookwalter DB, Roenfeldt KA, LeardMann CA, et al. Posttraumatic stress disorder and risk of selected autoimmune diseases among US military personnel. BMC Psychiatry 2020;20:23.
17. Spitzer C, Wegert S, Wollenhaupt J, et al. Gender-specific association between childhood trauma and rheumatoid arthritis: a case-control study. J Psychosomatic Res 2013;74:296–300.
18. Case SM, Feldman CH, Guan H, et al. Post-traumatic stress disorder and risk of systemic lupus erythematosus among Medicaid recipients. Arthritis Care Res (Hoboken) 2023;75:174–9.
19. Stojanovich L. Stress and autoimmunity [review]. Autoimmun Rev 2010;9:A271–6.
20. Segerstrom SC, Miller GE. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. Psychol Bull 2004;130:601–30.
21. Pedersen AF, Zachariae R, Bovbjerg DH. Psychological stress and antibody response to influenza vaccination: a meta-analysis. Brain Behav Immun 2009;23:427–33.
22. Pearlin LI. The life course and the stress process: some conceptual comparisons. J Gerontol B Psychol Sci Soc Sci 2010;65B:207–15.
23. Wesley A, Bengtsson C, Skillgate E, et al. Association between life events and rheumatoid arthritis: results from a population-based case-control study. Arthritis Care Res (Hoboken) 2014;66:844–51.

24. Bengtsson AA, Rylander L, Hagmar L, et al. Risk factors for developing systemic lupus erythematosus: a case-control study in southern Sweden. Rheumatology (Oxford) 2002;41:563–71.
25. Matthews KA, Gallo LC. Psychological perspectives on pathways linking socioeconomic status and physical health [review]. Ann Rev Psychol 2011;62:501–30.
26. Gallo LC, Matthews KA. Understanding the association between socioeconomic status and physical health: do negative emotions play a role? [review] Psychol Bull 2003;129:10–51.
27. Mouton CP, Rodabough RJ, Rovi SL, et al. Prevalence and 3-year incidence of abuse among postmenopausal women. Am J Public Health 2004;94:605–12.
28. Michael YL, Carlson NE, Chlebowski RT, et al. Influence of stressors on breast cancer incidence in the Women's Health Initiative. Health Psychol 2009;28:137–46.
29. Springfield S, Qin F, Hedlin H, et al. Modifiable resources and resilience in racially and ethnically diverse older women: implications for health outcomes and interventions. Int J Environ Res Public Health 2022;19:7089.
30. Parks CG, Walitt BT, Pettinger M, et al. Insecticide use and risk of rheumatoid arthritis and systemic lupus erythematosus in the Women's Health Initiative Observational Study. Arthritis Care Res (Hoboken) 2011;63:184–94.
31. Uebelacker LA, Eaton CB, Weisberg R, et al. Social support and physical activity as moderators of life stress in predicting baseline depression and change in depression over time in the Women's Health Initiative. Soc Psychiatry Psychiatr Epidemiol 2013;48:1971–82.
32. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. Controlled Clinical Trials 1998;19:61–109.
33. Walitt BT, Constantinescu F, Katz JD, et al. Validation of self-report of rheumatoid arthritis and systemic lupus erythematosus: the Women's Health Initiative. J Rheumatol 2008;35:811–8.
34. Ruberman W, Weinblatt E, Goldberg JD, et al. Psychosocial influences on mortality after myocardial infarction. N Engl J Med 1984;311:552–9.
35. Brown LJ, Potter JF, Foster BG. Caregiver burden should be evaluated during geriatric assessment. J Am Geriatr Soc 1990;38:455–60.
36. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32:705–14.
37. Burnam MA, Wells KB, Leake B, et al. Development of a brief screening instrument for detecting depressive disorders. Med Care 1988;26:775–89.
38. Harth M, Nielson WR. Pain and affective distress in arthritis: relationship to immunity and inflammation [review]. Expert Rev Clin Immuno 2019;15:541–52.
39. Huerta PT, Kowal C, DeGiorgio LA, et al. Immunity and behavior: antibodies alter emotion. Proc Natl Acad Sci U S A 2006;103:678–83.
40. Wilson-Genderson M, Heid AR, Cartwright F, et al. Adverse childhood experiences, adult trauma, and depressive symptom trajectories. Aging Ment Health 2022;26:2170–8.
41. Yazawa A, Shiba K, Inoue Y, et al. Early childhood adversity and late-life depressive symptoms: unpacking mediation and interaction by adult socioeconomic status. Soc Psychiatry Psychiatr Epidemiol 2022;57:1147–56.
42. Cozier YC, Barbhaiya M, Castro-Webb N, et al. Association of child abuse and systemic lupus erythematosus in Black women during adulthood. Arthritis Care Res (Hoboken) 2021;73:833–40.
43. Baker MW, LaCroix AZ, Wu C, et al. Mortality risk associated with physical and verbal abuse in women aged 50 to 79. J Am Geriatr Soc 2009;57:1799–809.
44. Mouton CP, Rodabough RJ, Rovi SL, et al. Psychosocial effects of physical and verbal abuse in postmenopausal women. Ann Fam Med 2010;8:206–13.
45. Byers AL, Covinsky KE, Neylan TC, et al. Chronicity of posttraumatic stress disorder and risk of disability in older persons. JAMA Psychiatry 2014;71:540–6.
46. Young HM, Bell JF, Whitney RL, et al. Social determinants of health: underreported heterogeneity in systematic reviews of caregiver interventions [review]. Gerontologist 2020;60 Suppl 1:S14–28.
47. Lynch SH, Shuster G, Lobo ML. The family caregiver experience - examining the positive and negative aspects of compassion satisfaction and compassion fatigue as caregiving outcomes. Aging Ment Health 2018;22:1424–31.
48. O'Reilly D, Rosato M, Maguire A. Caregiving reduces mortality risk for most caregivers: a census-based record linkage study. Int J Epidemiol 2015;44:1959–69.
49. Bacon KL, Heeren T, Keysor JJ, et al. Longitudinal and reciprocal relationships between depression and disability in older women caregivers and noncaregivers. Gerontologist 2016;56:723–32.
50. McCool-Myers M, Grasso D, Kozlowski D, et al. The COVID-19 pandemic's intersectional impact on work life, home life and wellbeing: an exploratory mixed-methods analysis of Georgia women's experiences during the pandemic. BMC Public Health 2022;22:1988.
51. Bagby SP, Martin D, Chung ST, et al. From the outside in: biological mechanisms linking social and environmental exposures to chronic disease and to health disparities. Am J Public Health 2019;109:S56–63.
52. Marsland AL, Walsh C, Lockwood K, et al. The effects of acute psychological stress on circulating and stimulated inflammatory markers: a systematic review and meta-analysis [review]. Brain Behav Immun 2017;64:208–19.
53. Rantapää-Dahlqvist S, de Jong BA, Berglin E, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;48:2741–9.
54. Bemis EA, Demoruelle MK, Seifert JA, et al. Factors associated with progression to inflammatory arthritis in first-degree relatives of individuals with RA following autoantibody positive screening in a non-clinical setting. Ann Rheum Dis 2021;80:154–61.
55. Greenblatt HK, Kim HA, Bettner LF, et al. Preclinical rheumatoid arthritis and rheumatoid arthritis prevention [review]. Curr Opin Rheumatol 2020;32:289–96.
56. Holers VM, Demoruelle MK, Kuhn KA, et al. Rheumatoid arthritis and the mucosal origins hypothesis: protection turns to destruction [review]. Nat Rev Rheumatol 2018;14:542–57.
57. Koffer RE, Thurston RC, Bromberger JT, et al. Racial/ethnic differences in women's life event exposure across midlife. J Gerontol B Psychol Sci Soc Sci 2022;77:272–83.
58. Fries JF. Current treatment paradigms in rheumatoid arthritis [review]. Rheumatology (Oxford) 2000;39 Suppl 1:30–5.
59. Schmajuk G, Schneeweiss S, Katz JN, et al. Treatment of older adult patients diagnosed with rheumatoid arthritis: improved but not optimal. Arthritis Rheum 2007;57:928–34.
60. Maloley PM, England BR, Sayles H, et al. Post-traumatic stress disorder and serum cytokine and chemokine concentrations in patients with rheumatoid arthritis. Semin Arthritis Rheum 2019;49:229–35.
61. Bürgin D, Boonmann C, Schmeck K, et al. Compounding stress: childhood adversity as a risk factor for adulthood trauma exposure in the health and retirement study. J Trauma Stress 2021;34:124–36.
62. Arpawong TE, Mekli K, Lee J, et al. A longitudinal study shows stress proliferation effects from early childhood adversity and recent stress on risk for depressive symptoms among older adults. Aging Ment Health 2022;26:870–80.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2529–2536
DOI 10.1002/acr.25174
© 2023 American College of Rheumatology.

American College of Rheumatology
Empowering Rheumatology Professionals

Social Determinants of Health Documentation Among Individuals With Rheumatic and Musculoskeletal Conditions in an Integrated Care Management Program

Sciaska N. Ulysse,[1] Mia T. Chandler,[2] Leah Santacroce,[1] Tianrun Cai,[1] Katherine P. Liao,[1] and Candace H. Feldman[1]

Objective. Social determinants of health (SDoH), such as poverty, are associated with increased burden and severity of rheumatic and musculoskeletal diseases. This study was undertaken to study the prevalence and documentation of SDoH-related needs in electronic health records (EHRs) of individuals with these conditions.

Methods. We randomly selected individuals with ≥1 International Classification of Diseases, Ninth/Tenth Revision (ICD-9/10) code for a rheumatic/musculoskeletal condition enrolled in a multihospital integrated care management program that coordinates care for medically and/or psychosocially complex individuals. We assessed SDoH documentation using terms for financial needs, food insecurity, housing instability, transportation, and medication access according to EHR note review and ICD-10 SDoH billing codes (Z codes). We used multivariable logistic regression to examine associations between demographic factors (age, gender, race, ethnicity, insurance) and ≥1 (versus 0) SDoH need as the odds ratio (OR) with 95% confidence interval (95% CI).

Results. Among 558 individuals with rheumatic/musculoskeletal conditions, 249 (45%) had ≥1 SDoH need documented in EHR notes by social workers, care coordinators, nurses, and physicians. A total of 171 individuals (31%) had financial insecurity, 105 (19%) had transportation needs, 94 (17%) had food insecurity; 5% had ≥1 related Z code. In the multivariable model, the odds of having ≥1 SDoH need was 2.45 times higher (95% CI 1.17–5.11) for Black versus White individuals and significantly higher for Medicaid or Medicare beneficiaries versus commercially insured individuals.

Conclusion. Nearly half of this sample of complex care management patients with rheumatic/musculoskeletal conditions had SDoH documented within EHR notes; financial insecurity was the most prevalent. Only 5% of patients had representative billing codes suggesting that systematic strategies to extract SDoH from notes are needed.

INTRODUCTION

Social determinants of health (SDoH), the conditions in which people work, live, and grow, contribute significantly to health behaviors and to inequities in health care access and outcomes (1,2). These nonmedical factors exist across medical specialties including rheumatology. In a national sample of patients with rheumatoid arthritis (RA), researchers demonstrated that faster declines in function over time and overall poorer functional status were observed in patients living in areas with more deprivation (measured using the Area Deprivation Index) (3). In a study that identified adults with RA from the National Health and Nutrition Examination Survey, >30% had food insecurity, which was associated with higher odds of depression (4). A study of patients with systemic lupus erythematous (SLE) showed that moving out of poverty led to lower mean scores of newly accumulated disease damage, similar to scores of participants who were never in poverty (5). Across rheumatic conditions, living in areas of high heat or social vulnerability has been associated with higher odds of recurrent hospitalizations (6).

The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grants P30-AR-072577 and K23-AR-071500).

[1]Sciaska N. Ulysse, BA, Leah Santacroce, MS, Tianrun Cai, MD, Katherine P. Liao, MD, MPH, Candace H. Feldman, MD, MPH, ScD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts;

[2]Mia T. Chandler, MD, MPH: Brigham and Women's Hospital, Harvard Medical School, and Boston Children's Hospital, Boston, Massachusetts.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25174.

Address correspondence via email to Candace H. Feldman, MD, MPH, ScD, at cfeldman@bwh.harvard.edu.

Submitted for publication January 24, 2023; accepted in revised form June 13, 2023.

SIGNIFICANCE & INNOVATIONS

- Through intensive chart reviews, we found a significant burden of social determinant of health (SDoH)–related needs, especially financial insecurity, and transportation challenges, in a subset of patients with rheumatic and musculoskeletal diseases enrolled in a multihospital complex care management program.
- While nearly 50% of individuals with rheumatic conditions enrolled in a complex care management program had at least 1 documented SDoH-related need, only 5% had an International Classification of Diseases, Tenth Revision SDoH-related billing (Z) code. Use of Z codes did not increase over time, suggesting that structured claims data do not capture the burden of need.
- There was nearly no documentation of SDoH-related needs by rheumatologists, suggesting that further education of rheumatologists, and rheumatology-based infrastructure to screen for and address these needs, are essential.

To date, certain SDoH have been examined among individuals with rheumatic diseases through quantitative research studies and qualitative interviews. However, we do not have a clear understanding of how SDoH-related needs are documented in the electronic health records (EHRs) of individuals with rheumatic diseases or the burden among individuals who may be at highest risk for adverse health outcomes related to these needs. We therefore aimed to systematically assess the documentation of SDoH-related needs in unstructured notes in the EHRs of patients with rheumatic conditions. We also aimed to determine whether structured billing International Classification of Diseases, Tenth Revision (ICD-10) Z codes, a set of standardized codes that are used to report the determinants that affect health-related outcomes (7), were being utilized and the degree to which they overlapped with unstructured documentation of SDoH.

Extraction of SDoH using structured billing codes would be a significantly easier way to understand population-level needs compared to manual chart review. However, we hypothesized that these Z codes would be underutilized and thus would underestimate the extent of SDoH-related needs. An understanding of both the burden of documented SDoH-related needs and the way in which they are documented will inform the development of algorithms that, if indicated, could combine natural language processing of unstructured notes with structured data (e.g., billing codes) to extract SDoH to guide clinical care and future research studies. In addition, by understanding the distribution and prevalence of SDoH-related needs according to rheumatic condition, resources can be better allocated to help rheumatology clinics develop infrastructure to better meet the needs of their patients, in turn reducing disparities in access and outcomes.

PATIENTS AND METHODS

Patient population. In our multi-institution academic hospital system, Mass General Brigham (MGB), a subset of medically and psychosocially complex individuals who receive their primary care through MGB-affiliated hospitals are enrolled in an integrated care management program (iCMP) (8). The iCMP includes a multidisciplinary team of nurses, social workers, community health workers, community resource specialists, and pharmacists and aims to coordinate and improve care and reduce costs. Individuals are identified for iCMP enrollment either through referral by their primary care physician or by a claims-based algorithm (9), which includes combinations of health care utilization patterns (e.g., recurrent emergency department visits), presence of complex and/or multiple medical issues, and/or a history of psychosocial needs (e.g., mental health diagnoses). During the timeframe of this study, the algorithm did not include granular SDoH-related needs. A qualitative study of 20 providers demonstrated that disease characteristics, including complexity, the diagnoses themselves and disease control, the environment of the patient (notably availability of social support), and ability to navigate the health care system, were considered when referring patients for the iCMP (10).

Each patient enrolled in the iCMP is assigned a specific care management lead (e.g., a nurse for patients where multiple medical issues drive complexity, or a social worker for psychologically complex patients). This care management lead conducts an initial assessment, creates a care plan, and manages the patient with the assistance of other members of the team depending on needs they uncover. The program, established in 2006, was initially supported by the Medicare Care Management for High-Cost Beneficiaries demonstration program (8). In 2012, the iCMP was extended and expanded without significant changes to the structure through MGB's participation in the Pioneer Accountable Care Organization (ACO) contract and is now seen as the main driver of the positive performance of the multihospital system on ACO risk contracts (8). The shared savings earned from those risk contracts provide the main funding mechanism for iCMP; enrollment is restricted to aligned beneficiaries (8).

SDoH are not systematically collected or documented as part of routine rheumatic disease care. However, screening for SDoH-related needs is part of the iCMP care manager's initial assessment for enrolled patients and was repeated as indicated, and more recently, is encouraged at least once yearly. Care managers' initial high-risk assessments include both free text documentation of reasons that render the patient "high risk" including living situation, functional status and financial concerns, and a series of multiple choice questions (see Supplementary Material 1 "High Risk Baseline Assessment," available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25174/abstract).

This information, however, is not consistently available or complete for all historically enrolled patients, is not in a location

that is readily accessible by other providers, does not include coded fields to facilitate structured data extraction, and at times lacks the granular detail included in more descriptive notes. SDoH-related details and needs uncovered during subsequent conversations are often documented in free text notes. To understand variations in documentation including but not limited to information collected at the baseline assessment, and the prevalence of SDoH among individuals with the highest likelihood of being screened in detail for SDoH needs (compared to the general population), we included adults ≥18 years old with ≥1 ICD-9 or ICD-10 code for a systemic rheumatic condition, crystalline arthritis, or osteoarthritis (OA) enrolled in the iCMP across MGB between January 1, 2012 (the year the iCMP was expanded) to October 18, 2021. To qualify for iCMP, individuals were required to have primary care providers within the MGB system, which ensured that they also had notes in the EHR (11).

Literature review. The study team (SNU, MTC) conducted literature reviews through PubMed and leveraged the Unified Medical Language System (UMLS), a collection of medical terms which includes some SDoH, to expand terms from the literature review to develop a dictionary of SDoH terms to guide in-depth EHR review. SDoH were defined using the categories of financial insecurity, food insecurity, housing instability, access to transportation, education, childcare, and access to medications. Together, the reviewers (SNU, MTC, CHF) developed a detailed standard operating procedure for the EHR review defining the date range of notes to review, types of notes, and search terms informed by the literature review (see Supplementary Material 2, "Search Terms from Standard Operating Procedure," available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25174/abstract), a prior pilot study of SDoH in a small subset of individuals with SLE and EHR data linked to billing claims (12), and UMLS.

SDoH data extraction: manual EHR review. Notes eligible for review included text written by physicians, nurses, social workers, psychiatrists, psychologists, dieticians/nutritionists, pharmacists, physical therapists, rheumatologists, and other health care professionals. Study team members (SNU, MTC, CHF) first reviewed the same 5 charts and adjudicated discrepancies. A second set of 5 charts was then reviewed, and the adjudication process was repeated. After discrepancies were adjudicated and definitions were agreed upon, the review team refined the systematic method for data extraction and divided the remaining charts between reviewers with weekly meetings to review charts together that raised questions for the primary reviewer. For each SDoH examined in this study, reviewers reported whether there were definite needs documented ("yes"), documentation of no need ("no"), the possibility of a need ("possible"), or no documentation ("not mentioned").

SDoH variables included financial insecurity, food insecurity, housing instability, transportation, education, childcare, medication access, and medication adherence. If an individual ever had a description of a need, they were categorized as having that need ("yes"), regardless of whether another note at a different time countered that narrative. Chart review was conducted during the dates of iCMP enrollment to focus on those notes most likely to have existing needs documented. An SDoH-related need was categorized as "possible" if there was a suggestion of a need but after comprehensive chart review by the team, a definite conclusion could not be reached. For each chart reviewed, demographic information was also extracted (age, gender, race, ethnicity, primary language, primary and secondary insurance), and the rheumatic condition.

ICD-10 Z Code identification. Z codes, introduced at the end of 2015, with studies of uptake beginning in 2016, are a subset of ICD-10, Clinical Modification codes used to report SDoH (7). The MGB Research Patient Data Registry (RPDR) was used to identify relevant ICD-10 Z codes for SDoH during the dates of iCMP enrollment. RPDR is a clinical data warehouse that integrates clinical information across the MGB health care system for research purposes (13). The Z codes provided by RPDR were from claims submitted by providers within our system. While a range of SDoH Z codes exist, we focused on those most relevant to the SDoH we were studying (e.g., problems related to education, employment, housing and economic circumstances, and problems related to medical facilities and other health care) (14), and stratified by the number of Z codes per year beginning in 2016.

Statistical analysis. Descriptive analyses were used to examine the overall prevalence of SDoH in this population from chart review and by Z code, and then identified SDoH were stratified by rheumatic condition. We used multivariable logistic regression including age, gender, race, ethnicity, insurance status, and rheumatic condition to examine associations between demographic factors and ≥1 (versus 0) SDoH-related needs (odds ratio [OR] with 95% confidence intervals [95% CIs]). We conducted an additional analysis removing transportation from the outcome, recognizing that this isolated need may be distinct from other SDoH measured. Additionally, we determined whether ICD-10 Z codes for SDoH were documented and the distribution of these codes over time. Analyses were conducted using SAS version 9.4 (SAS Institute) and R version 4.2.2. *P* values were 2-sided, and values less than 0.05 were considered statistically significant. This study was approved by the MGB Institutional Review Board.

RESULTS

Among 20,395 individuals (≥18 years old) with rheumatic conditions enrolled in MGB iCMP, we randomly selected 600 individuals for inclusion in this study. We excluded individuals without iCMP documentation (n = 35), or a clear rheumatic or

musculoskeletal disease diagnosis (n = 7). Among the 558 remaining individuals, the mean ± SD age was 73.7 ± 13.2 years, 62% were female, 80% were White, 9% were Black, and 82% were non-Hispanic (Table 1). The mean ± SD period of iCMP enrollment was 3.3 ± 2.4 years; there was no statistically significant difference in presence compared to the absence of SDoH need documentation according to mean enrollment time. There were 148 patients (27%) with a systemic rheumatic disease, 120 (22%) with crystalline arthritis, and 290 (52%) with OA without systemic or crystalline disease (categories are not mutually exclusive). Systemic rheumatic conditions included RA, palindromic rheumatism, SLE, systemic sclerosis, juvenile idiopathic arthritis, ankylosing spondylitis/sacroiliitis, Sjögren's/sicca syndrome, psoriatic arthritis, mixed/undifferentiated connective tissue disease, vasculitis, sarcoidosis, inflammatory myositis, polymyalgia rheumatica. Crystalline disease included gout and pseudogout.

Of the 558 charts reviewed, 249 (45%) had documentation of at least 1 definite ("yes") SDoH-related need. Overall, 171 individuals (31%) had definite evidence of financial needs, 105 (19%) had transportation needs, 94 (17%) had food insecurity, and 30 (5%) had housing instability. Inclusive of charts that were marked "possible," 88% of charts (490) contained documentation indicating at least 1 yes or 1 possible SDoH-related need. There were 126 individuals (23%) with possible evidence of financial needs, 176 (32%) with possible transportation needs, 92 (16%) with possible food insecurity, and 40 (7%) with possible housing instability. We also stratified documentation by age <65 years old compared to ≥65 years old and found that among those with definite SDoH needs, 88 individuals (35%) were <65 years old and 161 (65%) were ≥65 years old. In addition to these SDoH-related needs, we also assessed for documentation of education (n = 10) and childcare-related (n = 8) concerns. As this was an older population, we found that these needs were infrequently described and therefore they were not included in the final models.

Table 1. Baseline characteristics of 558 individuals with rheumatic/musculoskeletal conditions*

Variable	Value
Age, mean ± SD	73.7 ± 13.2
Gender	
Male	210 (38)
Female	348 (62)
Race	
Black	50 (9)
White	449 (80)
Other/not disclosed	60 (11)
Ethnicity	
Hispanic	49 (9)
Non-Hispanic	459 (82)
Other	50 (9)
Primary language	
English	516 (92)
Spanish	28 (5)
Other	14 (3)
Primary insurance	
Medicaid	32 (6)
Medicare	462 (83)
Commercial	39 (7)
Other	25 (4)
Rheumatic condition	
Systemic rheumatic condition†	148 (27)
Crystalline disease‡	120 (22)
Osteoarthritis	290 (52)

* Except where indicated otherwise, values are the number (%) of patients.
† Includes rheumatoid arthritis, palindromic rheumatism, systemic lupus erythematosus, systemic sclerosis, juvenile idiopathic arthritis, ankylosing spondylitis/sacroiliitis, Sjögren's/sicca syndrome, psoriatic arthritis, mixed/undifferentiated connective tissue disease, vasculitis, sarcoidosis, inflammatory myositis, and polymyalgia rheumatica.
‡ Includes gout and pseudogout.

Prevalence of documented SDoH needs varied by rheumatic condition (Figure 1). Among individuals with a systemic rheumatic disease, 39 individuals (26%) had evidence of financial insecurity, 32 (22%) had transportation needs, and 22 (15%) had food insecurity. For individuals with OA, 98 individuals (34%) had evidence of financial insecurity, 47 (16%) had transportation needs, and 51 (18%) had food insecurity. Furthermore, among individuals with a crystalline disease, 34 individuals (28%) had evidence of financial insecurity, 26 (22%) had transportation needs, and 21 (18%) had food insecurity.

We found significant heterogeneity in the descriptions and terms used for each SDoH need in the notes. Descriptions of financial insecurity included terms like "can't afford," "financial assistance," "limited income," and "struggling financially." Food insecurity was described with terms such as "food stamps," and housing instability was often implied with discussions of "subsidized housing." In our patient population, transportation needs were indicated by "PT-1," or "The Ride," a Massachusetts-based public transportation service for patients with temporary or permanent disabilities (15). We also noted terms and descriptions that required more context for interpretation. For example, "home delivered meals" was often used to describe Meals on Wheels, a service provided to seniors who experience physical declines or financial hardship (16); however, this is not a universal term for this service. While the terms "poor" and "poverty" are often used to describe individuals with insufficient funds, they were more frequently used in notes by physicians to describe "poor functional status" or "poverty of speech" rather than financial insecurity, emphasizing the importance of context alongside commonly used terms when delineating SDoH-related needs.

Descriptions and terms were found in various structured and unstructured note types such as telephone encounters, patient care coordination notes, templates, discharge summaries, progress notes, and consults. Notes were written by physicians from various fields such as rheumatology, primary care, psychiatry, and physical rehabilitation services and by iCMP nurses, physical therapists, occupational therapists, pharmacists, and medical

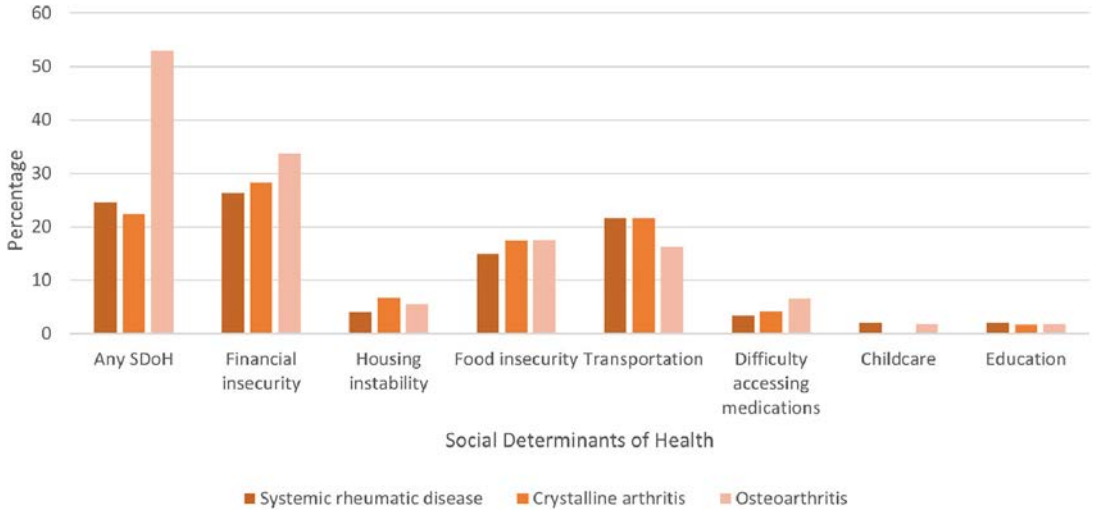


Figure 1. Percentage of individuals with social determinants of health–related needs according to rheumatic condition.

assistants. Among the notes reviewed by the study team, 245 notes had evidence of financial insecurity. Ninety-five of the notes (39%) indicating financial insecurity were recorded by an iCMP nurse, 53 of the notes (22%) were documented by a social worker, and only 1 note (0.4%) was written by a rheumatologist. Most mentions of housing instability and food insecurity were similarly included in notes written by the iCMP nurses or social workers. For housing instability, study team members extracted a total of 42 notes with a positive mention; 13 of the notes (31%) were written by a social worker and 11 (26%) were written by an iCMP nurse. For food insecurity, study team members extracted a total of 119 with a positive mention; 51 of the notes (43%) were written by an iCMP nurse, and 20 (17%) were written by a social worker. There were no notes written by rheumatologists that indicated housing or food needs among those patients with clear documentation of these needs elsewhere in their charts.

In the multivariable model, the odds of having ≥1 SDoH need was 2.45 times higher (95% CI 1.17–5.11) for Black individuals compared to White individuals, 6.72 times higher (95% CI 2.79–16.21) for Medicaid insurance beneficiaries compared to commercial insurance beneficiaries, 3.04 times higher (95% CI 1.32–6.97) for Medicare insurance beneficiaries compared to commercial insurance beneficiaries, and 4.12 times higher (95% CI 1.30–13.04) for individuals without insurance compared to commercial insurance beneficiaries (Table 2). We did not observe statistically significant differences by age, rheumatic condition, gender, or ethnicity. The multivariable model without transportation in the outcome resulted in similar findings (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25174/abstract).

We also examined the overall prevalence of SDoH by Z code (Table 3). Among our sample of 558 charts, we found that 26 individuals (5%) were assigned at least 1 SDoH Z code. The most frequently used SDoH Z code was Z59.9, defined as "problem related to housing and economic circumstances, unspecified." SDoH Z codes were also examined per year starting in 2016. After excluding individuals who died prior to 2016 or did not have a diagnosis or encounter code after 2016, uptake remained <5% each year and was highest in 2019 (3.4%) (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25174/abstract). One individual did not have an SDoH need in the reviewed categories but was assigned a billing Z code of Z75.8 ("other problems related to medical facilities and other

Table 2. Multivariable logistic regression model examining the odds of ≥1 social determinants of health–related needs versus no need in 558 individuals with rheumatic/musculoskeletal conditions*

Descriptive categories	Odds ratio	95% confidence interval
Age, years	0.97†	0.95–0.98†
Rheumatic Conditions		
Osteoarthritis	Ref.	Ref.
Systemic rheumatic disease	0.79	0.51–1.22
Crystalline arthritis	1.15	0.71–1.84
Gender		
Female	Ref.	Ref.
Male	0.70	0.47–1.05
Race		
White	Ref.	Ref.
Black	2.45†	1.17–5.11†
Other	0.99	0.48–2.04
Ethnicity		
Non-Hispanic	Ref.	Ref.
Hispanic	1.67	0.77–3.65
Other	1.41	0.75–2.66
Insurance		
Commercial	Ref.	Ref.
Medicaid	6.72†	2.79–16.21†
Medicare	3.04†	1.32–6.97†
No insurance	4.12†	1.30–13.04†

* Ref. = reference.
† Value was statistically significant.

Table 3. Number of individuals with an SDoH-related billing Z code (n = 26)*

ICD-10 codes to identify SDoH	Number of individuals with ICD-10 codes
Z55: Problems related to education and literacy	
Z55.0 Illiteracy and low-level literacy	1
Z55.9 Problems related to education and literacy, unspecified	5
Z56: Problems related to employment and unemployment	
Z56.0 Unemployment, unspecified	6
Z59: Problems related to housing and economic circumstances	
Z59.0 Homelessness	9
Z59.1 Inadequate housing	2
Z59.4 Lack of adequate food and safe drinking water	7
Z59.48 Other specified lack of adequate food	2
Z59.6 Low income	1
Z59.7 Insufficient social insurance and welfare support	4
Z59.8 Other problems related to housing and economic circumstances	4
Z59.9 Problem related to housing and economic circumstances, unspecified	10
Z75: Problems related to medical facilities and other health care	
Z75.8 Other problems related to medical facilities and other health care	1

* A total of 532 individuals were missing a relevant International Classification of Diseases, Tenth Revision (ICD-10) code. SDoH = social determinants of health.

health care"). A secondary analysis yielded similar results when this individual, who had no SDoH identified by manual chart review, was reclassified as having ≥1 SDoH need in the multivariable model (see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25174/abstract).

DISCUSSION

SDoH play a central role in disparities in care and outcomes in rheumatic and musculoskeletal conditions. Detailed chart review in this population of medically and/or psychosocially complex patients with rheumatic conditions receiving care at a multihospital academic medical center uncovered documentation of SDoH-related needs in nearly half of the charts, and when expanded to include those with possible needs, documentation increased to nearly 90% of the charts. The large discrepancy between the positive and possible mentions demonstrates the difficulty of accurately capturing information through unstructured data. Notably, we found that SDoH documentation was often recorded by iCMP nurses and social workers rather than by rheumatologists, with only 1 rheumatologist's note indicating a financial, food, or housing-related need, considerably less frequent compared to in primary care physicians' notes. Similarly, while 88% of the charts reviewed in this population had a definite or possible indication of an SDoH-related need, only 5% of charts had a related SDoH billing Z code without any significant increases in uptake over time. Prior studies in the general population have similarly demonstrated low uptake of Z codes to date (17–19), although utilization overall has increased since Z codes were implemented (7).

While the utilization of Z codes remains a promising option to understand population-level SDoH needs, rheumatologists at our multihospital institution are not trained or incentivized to use these codes or to screen for these needs. Patients seen in rheumatology who have primary care physicians within our system and who qualified for enrollment in iCMP were screened by iCMP care managers, and there were community resource specialists available to address the needs that were uncovered. However, both structured SDoH needs assessments and resources to meet these needs were not available as part of the rheumatology clinic infrastructure, which may in part explain the absence of described needs in rheumatology notes.

We also observed a higher prevalence of financial insecurity compared to other SDoH-related needs. Higher SDoH-related burden was noted among Black individuals and among individuals without insurance, as well as Medicaid and Medicare beneficiaries. The stark differences between racial and socioeconomic groups support findings from prior studies (20–23) and contribute to the underlying differences in health and health outcomes, even more so among a population with complex care needs (24). We found an overall higher prevalence of SDoH needs, and financial insecurity in particular, among individuals with OA compared to other rheumatic conditions; however, this difference was not statistically significant in adjusted analyses.

Several studies to date demonstrate the relevance of financial needs among individuals with rheumatic conditions. A 2016 study found that Medicaid patients were less likely to receive care from a rheumatologist and more likely to have delays in care in receiving medications, highlighting the role of socioeconomic status (SES) in health care access (25). In another study by Callahan et al, researchers found an association between lower individual- and community-level SES and poor physical health outcomes, highlighting the of role SES in disease outcomes for individuals with arthritis (26). Additionally, for individuals with RA, low SES has been associated with worse clinical outcomes and delays in treatment (27). Fewer studies to date examine the burden of financial needs, food insecurity, and transportation needs among individuals with rheumatic conditions, and more research is needed to demonstrate associations with medication and health care use and outcomes.

Strengths of our study include an understanding of SDoH documentation within EHR notes of medically and/or

psychosocially complex individuals with rheumatic and musculoskeletal conditions. Through intensive chart reviews, we have identified SDoH-related needs among individuals with rheumatic conditions, the way this information is being documented, and the providers who are recording this information. We have also examined uptake of ICD-10 Z codes among patients with rheumatic conditions and identified a gap for further educational efforts to promote greater utilization, and possibly another avenue to understand patient needs and complexity at a population level for individuals with rheumatic conditions. Prior studies that examine SDoH among patients with rheumatic conditions use methods including literature reviews (28), patient questionnaires (29,30), observational studies (31), and scoping reviews (32). However, we aimed to examine the prevalence of various SDoH-related needs using EHRs, which allows us to approximate real-world data. By understanding SDoH documentation in EHR notes, we can inform both the allocation of resources to meet these needs and future systematic data extraction strategies.

Limitations of this study include lack of data outside of our population cohort, as we have only analyzed individuals enrolled in the iCMP program who may be more likely to both have higher prevalence and documentation of SDoH-related needs compared to a less complex population. Compared to a prior study in individuals receiving rheumatic disease care within our multihospital system (6), this population, on average, was about 10 years older, had a similar gender and race distribution, and had a higher percentage of Hispanic individuals and Medicare beneficiaries. As iCMP only requires the primary care physician to be within the multihospital system and not the subspecialists, there may be misclassification of rheumatic/musculoskeletal conditions. Furthermore, our population was older due to Medicare insurance-related eligibility for the original iCMP, and as such, our findings may not represent the prevalence or distribution of SDoH in younger populations. With the absence of systematic screening at our institution in rheumatology clinics during the timeframe of this study, understanding needs in this enriched population allowed for the identification of strategies and infrastructure for future efforts. Additionally, the Z codes examined were limited to the billing data submitted by providers within our multihospital health care system and do not capture information from providers outside of our system. We do not expect that we are missing a significant number however, as the primary care setting would be the most likely for use of these codes, and all patients in this cohort had their primary care team within our system. Efforts to increase awareness of Z codes across providers at our institution may increase their use and allow for these factors to be better accounted for when understanding medical complexity, resource allocation, and care utilization patterns.

Another limitation to note is that reviewing charts and labeling them is difficult, since SDoH are, by definition, dynamic, making it challenging to classify individuals as having or not having needs overall rather than at specific time points. Z codes are important for population-level data, but our chart reviews demonstrated the importance of clinical context from narrative notes to truly understand the extent of SDoH-related needs. Further, SDoH documentation was infrequently structured or standardized in notes, and differences may, in part, reflect ascertainment bias within this academic-based complex care population. It is plausible that providers, including rheumatologists, may ask about SDoH but not routinely document them in their clinical notes. Alternatively, some providers may not screen for these needs since their clinics may not have the necessary tools to address them. As such, needs may be even higher than what was uncovered in this study. In addition, we did not examine neighborhood-level factors as our focus was on note-based documentation of SDoH; however, future analyses are planned to link these SDoH needs to area deprivation indices and neighborhood environmental exposures in this population.

This study illustrates the high burden of SDoH-related needs among individuals with rheumatic and musculoskeletal conditions and the importance of infrastructure to document and address these needs. Despite the high prevalence of SDoH in this population, in our chart review, we found only 1 note by a rheumatologist documenting financial, food or housing-related needs, suggesting that heightened awareness is needed for rheumatologists, and infrastructure is required in rheumatology clinics to meet the uncovered needs. Future studies should develop processes that effectively incorporate SDoH screening and EHR documentation into routine rheumatology care and that efficiently extract these data, and the actions taken in response. A strategy implemented in primary care at Boston Medical Center screened patients for SDoH, and their responses were linked to their EHR and incorporated into the structured data in their charts using ICD-10 Z codes. Then, if patients requested assistance, referrals (in the patients' primary languages) were provided to guide them to necessary resources (33). This strategy has the potential to be replicated in other institutions including our own.

We encourage rheumatologists to understand their patients' needs both biomedically and psychosocially and to advocate for resources and referral systems within their institutions to address these needs. In a clinical and public health sphere, SDoH information will allow for better care access and quality, more equitable enrollment in clinical trials, and more comprehensive research studies that appropriately account for the key contributions of SDoH to care utilization, medication adherence, and outcomes. Understanding SDoH allow health care providers to provide integrated care, and by connecting patients to appropriate services to address these needs, disparities in health care access and outcomes can be reduced (34).

ACKNOWLEDGMENTS

We would like to thank Christine Vogeli, PhD, and the Population Health team at Mass General Brigham for their assistance with data from individuals enrolled in the integrated care management program.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Ulysse, Chandler, Cai, Liao, Feldman.
Acquisition of data. Ulysse, Chandler, Santacroce, Liao, Feldman.
Analysis and interpretation of data. Ulysse, Chandler, Santacroce, Cai, Liao, Feldman.

REFERENCES

1. Healthy People 2030. Social determinants of health. URL: https://health.gov/healthypeople/priority-areas/social-determinants-health.
2. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes [review]. Public Health Rep 2014;129 Suppl 2:19–31.
3. Izadi Z, Li J, Evans M, et al. Socioeconomic disparities in functional status in a national sample of patients with rheumatoid arthritis. JAMA Netw Open 2021;4:e2119400.
4. Cai Q, Pesa J, Wang R, et al. Depression and food insecurity among patients with rheumatoid arthritis in NHANES. BMC Rheumatol 2022;6:6.
5. Yelin E, Trupin L, Yazdany J. A Prospective study of the impact of current poverty, history of poverty, and exiting poverty on accumulation of disease damage in systemic lupus erythematosus. Arthritis Rheumatol 2017;69:1612–22.
6. Santacroce L, Dellaripa PF, Costenbader KH, et al. Association of area-level heat and social vulnerability with recurrent hospitalizations among individuals with rheumatic conditions. Arthritis Care Res (Hoboken) 2023;75:22–33.
7. Centers for Medicare & Medicaid Services: Office of Minority Health. Utilization of Z codes for social determinants of health among Medicare fee-for-service beneficiaries, 2019. 2021. URL: www.cms.gov/files/document/z-codes-data-highlight.pdf.
8. Hsu J, Price M, Vogeli C, et al. Bending the spending curve by altering care delivery patterns: the role of care management within a pioneer ACO. Health Aff (Millwood) 2017;36:876–84.
9. Vogeli C, Spirt J, Brand R, et al. Implementing a hybrid approach to select patients for care management: variations across practices. Am J Manag Care 2016;22:358–65.
10. Haime V, Hong C, Mandel L, et al. Clinician considerations when selecting high-risk patients for care management. Am J Manag Care 2015;21:e576–82.
11. Massachusetts General Hospital. A team approach to care. URL: https://www.massgeneral.org/primary-care/integrated-care-management/our-team.
12. Taber KA, Williams JN, Huang W, et al. Use of an integrated care management program to uncover and address social determinants of health for individuals with lupus. ACR Open Rheumatol 2021;3:305–11.
13. Research Information Science Computing. Identify Subjects / Request Data. URL: https://rc.partners.org/research-apps-services/identify-subjects-request-data.
14. Johns Hopkins Medicine. Provider Update: ICD-10 Codes to identify social determinants of health. URL: https://www.hopkinsmedicine.org/johns_hopkins_healthcare/providers_physicians/resources_guidelines/provider_communications/2021/PRUP135_ICD10-km.pdf.
15. Massachusetts Bay Transportation Authority. Accessibility on the MBTA: the RIDE. URL: https://www.mbta.com/accessibility/the-ride.
16. Meals on Wheels America. URL: https://www.mealsonwheelsamerica.org/.
17. Weeks WB, Cao SY, Lester CM, et al. Use of Z-codes to record social determinants of health among fee-for-service Medicare beneficiaries in 2017. J Gen Intern Med 2020;35:952–5.
18. Liss DT, Cherupally M, Kang RH, et al. Social needs identified by diagnostic codes in privately insured U.S. adults. Am J Prev Med 2022;63:1007–16.
19. Torres JM, Lawlor J, Colvin JD, et al. ICD social codes: an underutilized resource for tracking social needs. Med Care 2017;55:810–6.
20. Healthy People 2030. Reduce the proportion of people living in poverty: SDOH-01. URL: https://health.gov/healthypeople/objectives-and-data/browse-objectives/economic-stability/reduce-proportion-people-living-poverty-sdoh-01.
21. Probst JC, Ajmal F. Rural & Minority Health Research Center: social determinants of health among the rural African American population. 2019. URL: https://www.sc.edu/study/colleges_schools/public_health/research/research_centers/sc_rural_health_research_center/documents/socialdeterminantsofhealthamongtheruralafricanamericanpopulation.pdf.
22. Noonan AS, Velasco-Mondragon HE, Wagner FA. Improving the health of African Americans in the USA: an overdue opportunity for social justice [review]. Public Health Rev 2016;37:12.
23. Williams DR, Mohammed SA, Leavell J, Collins C. Race, socioeconomic status, and health: Complexities, ongoing challenges, and research opportunities. Ann N Y Acad Sci 2010;1186:69–101.
24. Cherry AW. Maximus. The social determinants of health: why they matter to improving health outcomes. 2021. URL: https://maximus.com/article/social-determinants-health-why-they-matter-improving-health-outcomes.
25. Cifaldi M, Renaud J, Ganguli A, et al. Disparities in care by insurance status for individuals with rheumatoid arthritis: analysis of the medical expenditure panel survey, 2006–2009. Curr Med Res Opin 2016;32:2029–37.
26. Callahan LF, Martin KR, Shreffler J, et al. Independent and combined influence of homeownership, occupation, education, income, and community poverty on physical health in persons with arthritis. Arthritis Care Res (Hoboken) 2011;63:643–53.
27. Molina E, Del Rincon I, Restrepo JF, et al. Association of socioeconomic status with treatment delays, disease activity, joint damage, and disability in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2015;67:940–6.
28. Jackson LE, Danila MI. Healthcare disparities in telemedicine for rheumatology care [review]. Curr Opin Rheumatol 2022;34:171–8.
29. Looper KJ, Mustafa SS, Zelkowitz P, et al. Work instability and financial loss in early inflammatory arthritis. Int J Rheum Dis 2012;15:546–53.
30. Callhoff J, Ramos AL, Zink A, et al. The association of low income with functional status and disease burden in German patients with rheumatoid arthritis: results of a cross-sectional questionnaire survey based on claims data. J Rheumatol 2017;44:766–72.
31. Barton J, Trupin L, Schillinger D, et al. Racial and ethnic disparities in disease activity and function among persons with rheumatoid arthritis from university-affiliated clinics. Arthritis Care Res (Hoboken) 2011;63:1238–46.
32. Thomas M, Harrison M, De Vera MA. Reporting of determinants of health inequities in rheumatoid arthritis randomized controlled trials in Canada: a scoping review [review]. Arthritis Care Res (Hoboken) 2023;75:101–14.
33. Buitron de la Vega P, Losi S, Sprague Martinez L, et al. Implementing an EHR-based screening and referral system to address social determinants of health in primary care. Med Care 2019;57 Suppl 2:S133–9.
34. Health IT Analytics. Costs fell by 11% when payer addressed social determinants of health. 2018. URL: https://healthitanalytics.com/news/costs-fell-by-11-when-payer-addressed-social-determinants-of-health.

Arthritis Care & Research
Vol. 75, No. 12, December 2023, pp 2537–2545
© 2023 American College of Rheumatology.

American College of Rheumatology
Empowering Rheumatology Professionals

LETTERS

DOI 10.1002/acr.25194

Understanding stakeholders' perspectives to increase COVID-19 vaccine and booster: comment on the article by Ezeh et al

To the Editor:

We would like to share ideas on the article, Understanding Stakeholders' Perspectives to Increase COVID-19 Vaccine and Booster Uptake Among Black Individuals with Rheumatic Conditions, by Ezeh et al in a recent issue of *Arthritis Care and Research* (1). Ezeh et al sought community and physician perspectives on COVID-19 vaccination and booster hesitation, as well as ideas to enhance vaccine uptake among Black people with rheumatic and musculoskeletal disorders. To enhance vaccination uptake among Black individuals with rheumatic illnesses, Ezeh et al stated that initiatives should recognize and respond to racial, ethnic, and socioeconomic inequities that cause vaccine hesitation. According to Ezeh et al, messaging should be empathetic, specifically personalized, and should acknowledge variability in experiences and viewpoints (1).

Through interviews with community leaders and clinicians in greater Boston and Chicago, the paper delivers an essential study that tries to address vaccine reluctance and suggest solutions to target high-risk populations. With a well-defined moderator guide, a diversified sample size, and a rigorous qualitative analysis utilizing Dedoose, the study methodology looks to be sound. However, the study has several limitations that should be noted. For starters, the study only included community leaders and physicians from two specific geographic areas, which may not be typical of the larger population. Second, the study focuses on COVID-19 vaccine reluctance rather than other vaccination hesitancies. As a result, the study does not give a full assessment of vaccine apprehension in general. Furthermore, the study makes no mention of the potential drawbacks of the methods chosen for distributing vaccine-related information locally. Although overcoming vaccination hesitation may be helped by personal storytelling combined with an iterative and sympathetic approach, it might not be enough to dispel deeply ingrained attitudes and mistrust of health care systems.

It is important to emphasize that programs promoting vaccine acceptance should be praised. Concerns are voiced each time a new COVID-19 vaccination is created and made available to the general population. It should emphasize the significance of addressing concerns and encouraging vaccine adoption among various demographics. The advice from Ezeh et al for personalized and empathic messaging that recognizes the experiences and perspectives of different groups match with the greater need for effective promotion activities. The findings by Ezeh et al further underscore the importance of addressing underlying social and structural issues that contribute to vaccine hesitancy among Black people with rheumatic disorders. The COVID-19 vaccination might induce adverse responses in the general population, and this might be the reason some people hesitate (2). The onset of the COVID-19 outbreak, as well as the surrounding environment, have an impact on the resistance pattern (3). Because the hesitancy pattern is erratic, promotions will be more or less effective. As a result, information about the COVID-19 outbreak must be included in the study's context. If the pandemic's circumstances altered, the vaccination's acceptance rate may or may not change. As a result, it is vital to keep the shifting circumstances of the epidemic in mind when developing and conducting vaccination promotion efforts. This could include customizing messages and outreach activities to current community concerns and needs, as well as monitoring and responding to changes in vaccine acceptance rates.

As a result, governments and health care practitioners might utilize this data to create targeted interventions addressing unique concerns and barriers to immunization in diverse populations. Furthermore, the findings emphasize the need for accurate and trustworthy information regarding vaccines, as well as addressing misconceptions, to promote vaccine uptake. More study is needed to determine the generalizability of these findings to different demographics and circumstances, as well as to uncover effective vaccination hesitancy remedies.

It is essential to investigate how these underlying contexts affect the situation. Future use of the identified resistance pattern will be beneficial. If additional research is to be done, it should concentrate on identifying and addressing the core reasons for vaccination reluctance, such as misinformation, mistrust, and access hurdles, as well as designing and assessing efficient solutions to address vaccine hesitancy in various populations and circumstances.

In conclusion, Ezeh et al investigated COVID-19 vaccination hesitancy among Black people with rheumatic illnesses and proposed strategies to improve vaccine uptake, such as addressing disparities, employing individualized message, and accepting various experiences. While the methodology of the study is competent, drawbacks include a limited geographic emphasis and an exclusive focus on COVID-19 vaccine hesitancy. When advocating vaccination, the shifting circumstances of the pandemic should be considered.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25194.

Amnuay Kleebayoon, PhD
amnuaykleebai@gmail.com

Private Academic Consultant
Samraong, Cambodia
Viroj Wiwanitkit, MD
Chandigarh University
Punjab, India
and Joesph Ayobabalola University
Ikeji-Arakeji, Nigeria

1. Ezeh N, Sirek G, Ulysse SN, et al. Understanding stakeholders' perspectives to increase COVID-19 vaccine and booster uptake among Black individuals with rheumatic conditions. Arthritis Care Res (Hoboken) 2023;75(12):2508–18.
2. Mungmunpuntipantip R, Wiwanitkit V. COVID-19 vaccination hesitancy. Recenti Prog Med 2021;112:596.
3. Xiao J, Cheung JK, Wu P, et al. Temporal changes in factors associated with COVID-19 vaccine hesitancy and uptake among adults in Hong Kong: serial cross-sectional surveys. Lancet Reg Health West Pac 2022;23:100441.

DOI 10.1002/acr.25206

Reply

To the Editor:

We thank Drs. Kleebayoon and Wiwanitkit for your thoughtful remarks, your perspective regarding our article Understanding Stakeholders' Perspectives to Increase COVID-19 Vaccine and Booster Uptake among Black Individuals with Rheumatic Conditions, and your positive comments related to our study design and objectives.

In addressing your points regarding the limitations of the study, we would like to reinforce that this article reports findings from initial stakeholder engagement in two specific US cities that will ultimately inform a randomized controlled trial that leverages the Popular Opinion Leader (POL) model in those cities (1–3). We are currently developing a curriculum to address mistrust and misinformation guided by monthly physician/researcher and community stakeholder meetings to incorporate ever-changing sentiments and recommendations regarding COVID-19 vaccinations and boosters.

With regard to the scope of our study, we focused specifically on COVID-19 vaccine and booster hesitancy. Certainly, while some of the themes elucidated could be more broadly applicable to hesitancy related to other vaccines, this extrapolation would be out of the scope of this study and requires further investigation. While our study is limited geographically to two US cities, this was intentional, as our goal was to learn about the specific sentiments within the communities we aim to reach with our future intervention. Further studies are needed in other communities to appreciate the factors that contribute to COVID-19 vaccine and booster hesitancy that may be shared, and those that may differ. In addition, while we included two cities that have heterogenous representations of some of the diversity among individuals of African ancestry in the US, we explicitly acknowledge in our manuscript that there are other viewpoints within this varied community that are not represented in this work.

We agree that a multifaceted communication structure to improve vaccine and booster uptake is crucial. There is growing evidence that community engagement is critical to addressing the specific needs of historically marginalized and stigmatized communities (4–9). This should be in concert with global messaging regarding vaccine and booster uptake, not as a substitution. The POL model is an evidence-based social network model that is supported by the Centers for Disease Control and Prevention (10). This model, building on diffusion of innovation principles, has been successfully utilized to engage community leaders to disseminate information, change norms, and promote positive health behaviors associated with, but not limited to, lupus, HIV, violence prevention, concussion prevention, and organ donation within a community (11–19). The tailored messaging endorsed by this intervention will be community-driven to best reflect practices that key opinion leaders believe will be effective in their local communities.

Lastly, we have structurally addressed the impact of pandemic evolution in our study design that began with the interviews that informed this manuscript. We have routine meetings with our community partners which serve as a sustainable community engagement model to identify and address changing perceptions regarding COVID-19 infection, vaccination, and boosters. We have a consulting infectious disease physician on our study team to help revise our intervention as vaccine booster recommendations are updated. Additionally, we address both vaccine and booster hesitancy to acknowledge the shifting focus of communication efforts after the development of the COVID-19 vaccine boosters. The importance of this was notable as booster hesitancy was specifically identified during conversations with community leaders and physicians.

We thank you again for your commentary and engagement regarding the role of community-based interventions to improve COVID-19 vaccine and booster uptake for all. We agree that multilevel strategies are needed at the individual, community, national, and global level to address hesitancy and achieve equity in preventive care and we welcome future studies that expand upon this work.

Nnenna Ezeh, MD
Greta Sirek, BA
Sciaska N. Ulysse, BA
Candace H. Feldman, MD, MPH, ScD
cfeldman@bwh.harvard.edu
Brigham and Women's Hospital
Boston, Massachusetts
Rosalind Ramsey-Goldman, MD, DrPH
Northwestern University
Chicago, Illinois

1. Valente TW, Pitts SR. An appraisal of social network theory and analysis as applied to public health: challenges and opportunities [review]. Annu Rev Public Health 2017;38:103–18.

2. Leatherwood C, Canessa P, Cuevas K, et al. Community-engaged research: leveraging community-academic partnerships to reduce disparities and inequities in lupus care [review]. Rheum Dis Clin North Am 2021;47:109–18.
3. Arneson LC, Taber KA, Williams JN, et al. Use of popular opinion leader models to disseminate information about clinical trials to Black individuals with lupus in two US cities. Arthritis Care Res (Hoboken) 2023;75:44–52.
4. Kobau R, Carry M, Rubenstein BL, et al. Implementing the COVID-19 rapid community assessment on vaccine confidence: lessons learned from Alabama and Georgia. Public Health Rep 2022;137:832–40.
5. Marquez C, Kerkhoff AD, Naso J, et al. A multi-component, community-based strategy to facilitate COVID-19 vaccine uptake among Latinx populations: from theory to practice. PLoS One 2021; 16:e0257111.
6. Assoumou SA, Peterson A, Ginman E, et al. Addressing inequities in SARS-CoV-2 vaccine uptake: the Boston Medical Center health system experience [review]. Ann Intern Med 2022;175:879–84.
7. Mansfield LN, Carson SL, Castellon-Lopez Y, et al. Exploring perspectives on establishing COVID-19 vaccine confidence in Black communities. Ethn Dis 2022;32:341–50.
8. Ignacio M, Oesterle S, Mercado M, et al. Narratives from African American/Black, American Indian/Alaska Native, and Hispanic/Latinx community members in Arizona to enhance COVID-19 vaccine and vaccination uptake. J Behav Med 2023;46:140–52.
9. Kipke MD, Karimipour N, Wolfe N, et al. Community-based public health vaccination campaign (VaccinateLA) in Los Angeles' Black and Latino communities: protocol for a participatory study. JMIR Res Protoc 2023;12:e40161.
10. Kelly JA, Murphy DA, Sikkema KJ, et al. Randomised, controlled, community-level HIV-prevention intervention for sexual-risk behaviour among homosexual men in US cities. Lancet 1997;350:1500–5.
11. Mancera-Cuevas K, Canessa P, Chmiel JS, et al. Addressing lupus health disparities: the MONARCAS community and academic collaborative program. Health Equity 2018;2:145–51.
12. Phillip CR, Mancera-Cuevas K, Leatherwood C, et al. Implementation and dissemination of an African American popular opinion model to improve lupus awareness: an academic-community partnership. Lupus 2019;28:1441–51.
13. Li L, Guan J, Liang LJ, et al. Popular opinion leader intervention for HIV stigma reduction in health care settings. AIDS Educ Prev 2013;25: 327–35.
14. Theall KP, Fleckman J, Jacobs M. Impact of a community popular opinion leader intervention among African American adults in a southeastern United States community. AIDS Educ Prev 2015;27:275–87.
15. Young SD, Konda K, Caceres C, et al. Effect of a community popular opinion leader HIV/STI intervention on stigma in urban, coastal Peru. AIDS Behav 2011;15:930–7.
16. NIMH Collaborative HIV/STD Prevention Trial Group. The community popular opinion leader HIV prevention programme: conceptual basis and intervention procedures. AIDS 2007;21 Suppl 2:S59–68.
17. Edwards KM, Banyard VL, Waterman EA, et al. Use of social network analysis to identify popular opinion leaders for a youth-led sexual violence prevention initiative. Violence Against Women 2022;28:664–85.
18. Kerr ZY, Register-Mihalik JK, Haarbauer-Krupa J, et al. Using opinion leaders to address intervention gaps in concussion prevention in youth sports: key concepts and foundational theory. Inj Epidemiol 2018;5:28.
19. Shi J, Salmon CT. Identifying opinion leaders to promote organ donation on social media: network study. J Med Internet Res 2018;20:e7.

DOI 10.1002/acr.25177

Factors affecting serum urate monitoring among older adults with gout initiating urate-lowering therapy: comment on the article by Kwok et al

To the Editor:

With great interest, we read the recent article by Kwok et al (1), recently published in *Arthritis Care & Research*, investigating factors associated with serum urate (SUA) monitoring in older persons with gout who initiated urate-lowering therapy. The authors revealed that clinicians in other specialty fields were less likely to perform SUA tests with resulting odds ratios (ORs) ranging from 0.25 to 0.37 compared to rheumatologists. In the patient-level factors, older people with rural residence or lower socioeconomic status had a lower probability of SUA monitoring with ORs ranging from 0.81 to 0.91 than those without it, while patients with chronic kidney disease, hypertension, diabetes mellitus, or coprescription of colchicine/oral corticosteroids posed a higher frequency of SUA testing with an OR of 1.31 compared to those without (1). However, we believe there are some issues about the potential confounding effect of unrated covariates on this study.

Although considerable data on study participants (e.g., age, sex, comorbidities, income, colchicine/oral corticosteroids prescription, laboratory tests, etc.) based on multidimensional databases had been provided, the use of certain drugs related to alleviating or deteriorating hyperuricemia was not observed. Some drugs (such as sodium-dependent glucose cotransporter 2 inhibitors, statins, bile acid resins, niacin, fibrates, calcium-channel blockers, etc.) appear to have a urate-lowering effect or a lower risk of gout flares (2–5); in contrast, certain medications (e.g., diuretics, beta-blockers, levodopa, salicylates, especially low-dose aspirin <300 mg daily, angiotensin-converting enzyme inhibitors, and non-losartan angiotensin II receptor blockers) may have an increased risk of hyperuricemia or gout attack (6–8). Additionally, other covariates associated with the risk of hyperuricemia, including ethnicity, diet habits, level of smoking or drinking, body mass index, and worsening or new onset of comorbidities (8–11), were also not evaluated. These issues could affect the research results and need to be clarified.

Finally, we appreciate the impressive work of Kwok et al. However, we would like to draw the reader's attention to these potential limitations by carefully interpreting the main findings of this article and sincerely look forward to their response.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25177.

Kai-Lun Sheu, MD, MSc
Chun-Chieh Chen, MD, PhD
Shiuan-Chih Chen, MD, PhD
sccy399@yahoo.com.tw

Chung Shan Medical University School of Medicine and Chung Shan Medical University Hospital Health Management Center
Taichung, Taiwan

1. Kwok TS, Kuriya B, Hawker G, et al. Serum urate monitoring amongst older adults with gout initiating urate lowering therapy in Ontario, Canada. Arthritis Care Res (Hoboken) 2023;75(12):2463–71.
2. Waldman B, Ansquer JC, Sullivan DR, et al. Effect of fenofibrate on uric acid and gout in type 2 diabetes: a post-hoc analysis of the randomised, controlled FIELD study. Lancet Diabetes Endocrinol 2018; 6:310–8.
3. Choi HK, Soriano LC, Zhang Y, et al. Antihypertensive drugs and risk of incident gout among patients with hypertension: population based case-control study. BMJ 2012;344:d8190.
4. Juraschek SP, Simpson LM, Davis BR, et al. The effects of antihypertensive class on gout in older adults: secondary analysis of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial. J Hypertens 2020;38:954–60.
5. Zhou J, Liu X, Chou OH, et al. Lower risk of gout in sodium glucose cotransporter 2 (SGLT2) inhibitors versus dipeptidyl peptidase-4 (DPP4) inhibitors in type-2 diabetes. Rheumatology (Oxford) 2023; 62:1501–10.
6. Ben Salem C, Slim R, Fathallah N, et al. Drug-induced hyperuricaemia and gout [review]. Rheumatology (Oxford) 2017;56:679–88.
7. Li L, Zhang Y, Zeng C. Update on the epidemiology, genetics, and therapeutic options of hyperuricemia [review]. Am J Transl Res 2020; 12:3167–81.
8. Dalbeth N, Gosling AL, Gaffo A, et al. Gout [review]. Lancet 2021;397: 1843–55.
9. Roman YM. Moving the needle in gout management: the role of culture, diet, genetics, and personalized patient care practices [review]. Nutrients 2022;14:3590.
10. Thompson MD, Wu YY, Cooney RV, et al. Modifiable factors and incident gout across ethnicity within a large multiethnic cohort of older adults. J Rheumatol 2022;49:504–12.
11. Yang Y, Xian W, Wu D, et al. The role of obesity, type 2 diabetes, and metabolic factors in gout: a Mendelian randomization study. Front Endocrinol (Lausanne) 2022;13:917056.

DOI 10.1002/acr.25188

Reply

To the Editor:

We thank Drs. K. Sheu, C. Chen, and S. Chen for their interest in our study on the frequency of and associated patient- and physician-level factors correlated with SUA monitoring following urate-lowering therapy (ULT) initiation in older adults with gout (1). They raise concerns regarding the potential confounding of unmeasured covariates that may contribute to hyperuricemia.

We agree with Sheu et al that certain classes of medications and other clinical parameters may influence SUA levels. However, we wish to emphasize that the outcome of our study was focused on quality of gout care as measured by whether patients had an SUA test performed at an appropriate interval, rather than the specific SUA level achieved. SUA monitoring is necessary to optimize ULT management and this quality measure we assessed (adapted from the American College of Rheumatology electronic clinical quality measures for gout) is applicable to all gout patients initiating ULT, irrespective of other concomitant therapies (2). Moreover, if these additional drugs were indeed present in an individual, one might anticipate an even higher uptake of SUA monitoring, which makes our study results even more concerning from a gout quality of care perspective.

We agree that the incorporation of drug and clinical data, and in particular pharmacological classes of medications shown to affect SUA levels, would be critical in subsequent research on the achievement of target SUA levels and is an area of subsequent inquiry for our research team. In performing such a future study, we acknowledge that the incorporation of certain clinical variables Sheu et al raised (that may affect SUA levels such as body mass index) may not be possible with the inherent limitations of using health administrative data.

Timothy S. H. Kwok, MD, MSc, FRCPC
timothysh.kwok@mail.utoronto.ca
Bindee Kuriya, MD, SM, FRCPC
University of Toronto
Toronto, Ontario, Canada
Gillian Hawker, MD, MSc, FRCPC
University of Toronto,
Women's College Hospital,
and University of Toronto
Toronto, Ontario, Canada
Gregory Choy, MD, FRCPC
University of Toronto
and Sunnybrook Research Institute
Toronto, Ontario, Canada
Jessica Widdifield, PhD
Women's College Hospital,
University of Toronto,
Sunnybrook Research Institute,
and Institute for Clinical Evaluative Sciences
Toronto, Ontario, Canada

1. Kwok TS, Kuriya B, Hawker G, et al. Serum urate monitoring amongst older adults with gout: initiating urate-lowering therapy in Ontario, Canada. Arthritis Care Res (Hoboken) 2023;75(12):2463–71.
2. FitzGerald JD, Mikuls TR, Neogi T, et al. Development of the American College of Rheumatology electronic clinical quality measures for gout. Arthritis Care Res (Hoboken) 2018;70:659–71.

DOI 10.1002/acr.25182

Factors affecting the risk of falling in people with knee osteoarthritis: comment on the article by Wilfong et al

To the Editor:

We have read with interest the Canadian community-dwelling cohort study by Wilfong et al recently published in *Arthritis Care & Research* (1) investigating the risk of falls among adults with knee osteoarthritis (OA). According to analysis of large-scale, population-based data, Wilfong et al disclosed that people with knee OA had a higher risk of falling and were more likely to have a

history of falling while standing or walking indoors. However, despite showing promising results, we believe there is a potential effect of residual confounders in this study.

Although the authors have endeavored to adjust several covariates, including demographic characteristics, chronic disorders, and performance-related data, certain potential confounders were not evaluated. These unmeasured covariates, including the severity of chronic disorders (e.g., stage of chronic obstructive pulmonary disease or asthma, severity of cerebrovascular diseases, stage of hypertension, glycemic control status of diabetes mellitus, severity of mood disorder, and degree of cognitive impairment), the level of burden of composite comorbidity, nutritional status, and the use of medications related to an increased risk of falling (such as anticonvulsants, antidepressants, antipsychotics, sedatives, diuretics, etc.), chronic kidney injury, congestive heart failure, and severity of knee OA, would be associated with fall risk (2–7), which could affect the research results. This issue needs further clarification.

Finally, we appreciate the impressive work of Wilfong et al, however, we would like to alert the reader to the potential limitations of this study when reading the research results. We look forward to their response.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25182.

Chun-Chieh Chen, MD, PhD
Yuan-Ti Lee, MD, PhD
Shiuan-Chih Chen, MD, PhD
sccy399@yahoo.com.tw
Chung Shan Medical University School of Medicine and Chung Shan Medical University Hospital Health Management Center
Taichung, Taiwan

1. Wilfong JM, Perruccio AV, Badley EM. Examination of the increased risk for falls among individuals with knee osteoarthritis: a Canadian Longitudinal Study on Aging population-based study. Arthritis Care Res (Hoboken) 2023;75:2336–44.
2. Bowling CB, Bromfield SG, Colantonio LD, et al. Association of reduced eGFR and albuminuria with serious fall injuries among older adults. Clin J Am Soc Nephrol 2016;11:1236–43.
3. Carazo M, Sadarangani T, Natarajan S, et al. Prognostic utility of the Braden Scale and the Morse Fall Scale in hospitalized patients with heart failure. West J Nurs Res 2017;39:507–23.
4. Green AR, Reifler LM, Bayliss EA, et al. Drugs contributing to anticholinergic burden and risk of fall or fall-related injury among older adults with mild cognitive impairment, dementia and multiple chronic conditions: a retrospective cohort study. Drugs Aging 2019;36:289–97.
5. Schiek S, Hildebrandt K, Zube O, et al. Fall-risk-increasing adverse reactions–is there value in easily accessible drug information? A case-control study. Eur J Clin Pharmacol 2019;75:849–57.
6. Schoberer D, Breimaier HE, Zuschnegg J, et al. Fall prevention in hospitals and nursing homes: clinical practice guideline [review]. Worldviews Evid Based Nurs 2022;19:86–93.
7. Harris R, Strotmeyer ES, Sharma L, et al. The association between severity of radiographic knee OA and recurrent falls in middle and older aged adults: the Osteoarthritis Initiative. J Gerontol A Biol Sci Med Sci 2023;78:97–103.

DOI 10.1002/acr.25191

Reply

To the Editor:

We thank Chen et al for their letter and interest in our work. As is evident from the list of variables included in our regression analyses, we did control for the presence of a list of individual co-occurring conditions based on a review of the literature of risk factors for falls (1). In contrast to clinical studies that usually focus on specific aspects of a single disease, such as severity, our work was based on a larger population-based study concerned with overall health, capturing data across physical, mental, and social dimensions of health and across individual, neighborhood, and regional strata, with a broad focus on chronic conditions (2). The severity of individual conditions was not determined as part of this larger study and thus could not be considered in our work.

We agree with Chen et al that disease severity certainly can be an important factor in falls, and the interpretation of findings should consider the potential effects of varying disease severities across individuals. Having considered the presence of a list of individual conditions in our regression analyses, the issue of whether the absence of data around severity of co-occurring conditions influences the results may rest on whether, being present, the severity of other conditions differs between those with knee osteoarthritis (OA) and those without in our sample. Unfortunately, we do not know whether this is the case, and this may warrant further research. However, even if it is the case that co-occurring disease severities may be greater among those with knee OA, and this contributes to increased falls in this group, the fact remains that for those living with knee OA, there is an increased risk of falls. Therefore, we believe that fall prevention is an important clinical target in those with knee OA, particularly for those with additional risk factors for falls, as falling can cause additional joint damage and injury.

As we note in our Methods, our cardiovascular disease variable included heart disease. The heart disease question put to survey respondents was, "Has a doctor ever told you that you have heart disease (including congestive heart failure, or CHF)?" Therefore, congestive heart failure was included in our study. With regard to medication use, we note in our Methods that in identifying the presence of diabetes mellitus, hypertension, and depression, our definition included either a report of the disease or a report of medication use for these conditions. Medications used for these conditions have been shown to be associated with fall risk and were readily available in the dataset (3–5). Chen et al also identified chronic kidney disease as a possible confounder. While some evidence does show an association with fall risk (6), we had not adjusted for its presence. In the sample, 2.7% of individuals reported having kidney disease or kidney failure. Therefore, we reanalyzed the data with the addition of kidney disease. The results are presented below in the Table (Model 1 replicates

Table 1. Examination of the predictors of reporting an injurious fall in the overall sample (outcome: reported fall(s) vs. no reported fall)*

Model Variables	Variable Category	Model 1 OR (95% CI)	Model 2 OR (95% CI)
Knee OA (ref. No)	Yes	1.33 (1.14–1.56)†	1.32 (1.12–1.54)†
Age (ref. 75–85)	45–54	0.93 (0.76–1.13)	0.93 (0.76–1.15)
	55–64	0.87 (0.72–1.06)	0.88 (0.73–1.07)
	65–74	0.85 (0.70–1.03)	0.85 (0.70–1.03)
Sex (ref: male)	Female	1.44 (1.27–1.63)†	1.44 (1.28–1.63)†
BMI (ref. underweight/normal)	Overweight	0.96 (0.83–1.11)	0.97 (0.84–1.11)
	Obese	0.88 (0.74–1.03)	0.87 (0.74–1.03)
Alcohol use (0–1 times/week)	6+ times/week	1.18 (1.00–1.39)†	1.19 (1.01–1.40)†
	2–5 times/week	1.10 (0.96–1.25)	1.10 (0.97–1.26)
Baseline fall (ref. No)	Yes	1.85 (1.50–2.26)†	1.86 (1.52–2.28)†
Knee symptoms (ref. No)	Yes	1.13 (0.98–1.30)	1.13 (0.98–1.29)
Lower fracture (ref. No)	Yes	1.39 (1.23–1.57)†	1.38 (1.22–1.57)†
Vision problems (ref. No)	Yes	1.10 (0.89–1.36)	1.11 (0.89–1.37)
Respiratory (ref. No)	Yes	1.17 (1.01–1.35)†	1.16 (1.00–1.34)†
CVD (ref. No)	Yes	1.15 (0.97–1.36)	1.13 (0.95–1.34)
Urinary incontinence (ref. No)	Yes	1.29 (1.06–1.56)†	1.28 (1.05–1.55)†
Neurological (ref. No)	Yes	1.60 (1.15–2.22)†	1.61 (1.16–2.24)†
Diabetes mellitus (ref. No)	Yes	1.05 (0.89–1.24)	1.03 (0.87–1.22)
High blood pressure (ref. No)	Yes	0.97 (0.85–1.11)	0.97 (0.85–1.10)
Depression (ref. No)	Yes	1.40 (1.21–1.62)†	1.40 (1.21–1.62)†
One leg balance (ref. >4.5 seconds)	≤4.5 seconds	1.34 (1.12–1.60)†	1.33 (1.11–1.59)†
TUG time (ref. <14.2 seconds)	≥14.2 seconds	1.34 (0.96–1.86)	1.33 (0.96–1.85)
Chair rise test (ref. <15.9 seconds)	≥15.9 seconds	1.09 (0.94–1.26)	1.08 (0.93–1.25)
Chronic kidney disease (ref. No)	Yes	–	1.55 (1.15–2.09)†

* 95% CI = 95% confidence interval; CVD = cardiovascular disease; OA = osteoarthritis; OR = odds ratio; ref = reference; TUG = Timed-Up-and-Go.
† $P < 0.05$.

what appears in the original manuscript, Model 2 additionally includes kidney disease). Kidney disease was indeed a significant predictor of falls (odds ratio 1.55 [95% confidence interval 1.15–2.09]). However, note that the association between knee OA and falls was unchanged. Our conclusion, that individuals with knee OA are at an increased risk of falls, independent of other risk factors, remains the same.

As indicated, data on disease severity across the conditions were not available, including for knee OA. Chen et al cite the paper by Harris et al (7) in their letter, as did we in our manuscript (1). Harris et al did have data on knee OA severity for their study (as measured by Kellgren-Lawrence grading). While they did report some degree of elevated fall risk with increasing knee OA severity, fall risk was significant even among those with mild disease severity. As our population-based knee OA group likely included individuals with a full range of OA severities, we concur with the Harris et al (7) conclusion that fall prevention efforts should focus on "all stages of KOA [knee OA] from possible to moderate-severe."

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Dr. Perruccio and Ms. Wilfong's work was supported by the Arthritis Society Canada (grant STAR-20-0000000012 and a collaborative service agreement, respectively).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25191.

Anthony V. Perruccio, PhD
anthony.perruccio@uhnresearch.ca
Jessica M. Wilfong, MPH
Elizabeth M. Badley, DPhil
Schroeder Arthritis Institute
University Health Network
Dalla Lana School of Public Health
University of Toronto
and Arthritis Community Research and Epidemiology Unit
Toronto, Ontario, Canada

1. Wilfong JM, Perruccio AV, Badley EM. Examination of the increased risk for falls among individuals with knee osteoarthritis: a Canadian Longitudinal Study on Aging population-based study. Arthritis Care Res (Hoboken) 2023;75:2336–44.
2. Raina PS, Wolfson C, Kirkland SA, et al. The Canadian Longitudinal Study on Aging (CLSA). Can J Aging 2009; 28:221–9.
3. Alenazi A, Alhowimel A, Alshehri M, et al. Generalized and localized osteoarthritis and risk of fall among older adults: the role of chronic diseases and medications using real world data from a single center. Eur Rev Med Pharmacol Sci 2023;27:3957–66.
4. Tinetti ME, Han L, Lee DS, et al. Antihypertensive medications and serious fall injuries in a nationally representative sample of older adults. JAMA Intern Med 2014;174:588–95.
5. Darowski A, Chambers SA, Chambers DJ. Antidepressants and falls in the elderly [review]. Drugs Aging 2009;26:381–94.
6. Goto N, Weststrate A, Oosterlaan F, et al. The association between chronic kidney disease, falls, and fractures: a systematic review and meta-analysis. Osteoporos Int 2020;31:13–29.

7. Harris R, Strotmeyer ES, Sharma L, et al. The association between severity of radiographic knee OA and recurrent falls in middle and older aged adults: the Osteoarthritis Initiative. J Gerontol A Biol Med Sci 2023;78:97–103.

DOI 10.1002/acr.25183

Sleep-related disorders in adults with rheumatoid arthritis: comment on article by Katz et al

To the Editor:

We read with great interest the article by Katz et al recently published in *Arthritis Care & Research* (1) investigating sleep disorders in people with rheumatoid arthritis (RA). The study revealed that of patients with RA, 21% had a diagnosis or risk of obstructive sleep apnea (OSA), 30% had a diagnosis or symptoms of restless legs syndrome (RLS), and 43% reported short sleep (SS). An increased level of RA-related pain or Rheumatoid Arthritis Disease Activity Index score was associated with all sleep disorders. However, despite promising results, we would like to address concern about the potential confounding effect of unmeasured covariates on this study.

Several craniofacial abnormalities, including narrowing of the lateral peritonsillar, retrognathia, tonsillar hypertrophy, micrognathia, macroglossia, high-arched or narrow palate, enlarged or elongated uvula, nasal septal deviation, and nasal polyps, can narrow the upper airways and appear to pose an increased risk of having OSA (2–4). About 40–60% of patients with RLS have a family history indicating that RLS presents a high pedigree trait (5–7). Furthermore, low iron stores in the brain that can accompany reduced serum ferritin levels have shown to be associated with the risk of RLS, especially in older people or those without a family history of RLS (6–8). Some people with depletion of iron stores in the brain may not cause a decrease in hemoglobin or hematocrit levels, and those with RLS and iron deficiency would not have been anemic (9,10). Parkinson's disease, pregnancy, uremia, multiple sclerosis, and spinal cord disorders have also been reported to carry an increased risk of RLS (7,11). These comorbidities or conditions correlated with OSA or RLS were not evaluated in the present study, which could affect the study results. Thus, this issue needs to be clarified.

Finally, we appreciate the impressive work of Katz et al. Because there was limited information based on self-report data in this study, we fear the potential confounding effect could not be fully unraveled. Therefore, we would like to draw the reader's attention to the potential limitations in interpreting the important findings of this study.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25183.

Chun-Chieh Chen, MD, PhD
Po-Hui Wang, MD, PhD
Shiuan-Chih Chen, MD, PhD
sccy399@yahoo.com.tw

Chung Shan Medical University School of Medicine and Chung Shan Medical University Hospital Health Management Center
Taichung, Taiwan

1. Katz P, Pedro S, Michaud K. Sleep disorders among individuals with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2023;75:1250–60.
2. Young T, Skatrud J, Peppard PE. Risk factors for obstructive sleep apnea in adults [review]. JAMA 2004;291:2013–6.
3. Epstein LJ, Kristo D, Strollo PJ Jr, et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. J Clin Sleep Med 2009;5:263–76.
4. Patel SR. Obstructive sleep apnea [review]. Ann Intern Med 2019;171:ITC81–96.
5. Ondo W, Jankovic J. Restless legs syndrome: clinicoetiologic correlates. Neurology 1996;47:1435–41.
6. Zhu XY, Wu TT, Wang HM, et al. Clinical features and subtypes of restless legs syndrome in Chinese population: a study of 359 patients. Sleep Med 2019;59:15–23.
7. Trenkwalder C, Allen R, Högl B, et al. Comorbidities, treatment, and pathophysiology in restless legs syndrome [review]. Lancet Neurol 2018;17:994–1005.
8. Earley CJ, Connor JR, Beard JL, et al. Abnormalities in CSF concentrations of ferritin and transferrin in restless legs syndrome. Neurology 2000;54:1698–700.
9. Ulfberg J, Nyström B. Restless legs syndrome in blood donors. Sleep Med 2004;5:115–8.
10. Spencer BR, Kleinman S, Wright DJ, et al. Restless legs syndrome, pica, and iron status in blood donors. Transfusion 2013;53:1645–52.
11. Manconi M, Garcia-Borreguero D, Schormair B, et al. Restless legs syndrome [review]. Nat Rev Dis Primers 2021;7:80.

DOI 10.1002/acr.25208

Reply

To the Editor:

We thank Chen et al for their careful attention to our article, Sleep Disorders Among Individuals With Rheumatoid Arthritis (1,2). We agree there are covariates that may affect the presence of sleep disorders for which we were unable to account.

While these conditions may contribute to the underlying cause of sleep disorders, two questions regarding the high prevalence of sleep disorders among this cohort of individuals with RA are relevant: Are these risk factors more common among individuals with RA? Do these additional risks factors mitigate the importance of our findings?

Some risk factors described are indeed more common in individuals with RA or inflammatory arthritis. For example, as Chen et al mention, craniofacial abnormalities may create a structural risk for OSA, a risk that may be particularly relevant to adults who had juvenile-onset inflammatory arthritis (3). Higher rates of iron deficiency have been shown in adults with RA (4,5). Use of glucocorticoids may predispose to the onset of diabetes mellitus (6), which may lead to neuropathy, another risk factor for RLS (7). It appears, then, that at least some of the additional factors associated with sleep disorders

are more common among people with RA. It is possible, perhaps even likely, that we would have noted the associations of these additional factors with the presence of OSA and RLS if we had been able to include them in our analyses.

However, the greater presence of these risk factors makes it even more critical to consider sleep disorders in RA. In addition to the risks of sleep disorders that may be conferred by RA disease activity, inflammation, and pain, a greater prevalence of additional risk factors for sleep disorders should increase attention to these conditions that can have serious impacts on health and well-being.

Again, we thank Chen and colleagues for raising these important points and hope our findings will stimulate further research into the prevalence, risk, and impact of sleep disorders in RA and other rheumatic and autoimmune conditions.

Patricia Katz, PhD
patti.katz@ucsf.edu
University of California, San Francisco
Sofia Pedro, MSc
FORWARD, The National Databank for Rheumatic Diseases
Wichita, Kansas
Kaleb Michaud, PhD
FORWARD, The National Databank for Rheumatic Diseases
Wichita, Kansas
and University of Nebraska Medical Center
Omaha

1. Chen CC, Wang PH, Chen SC. Sleep-related disorders in adults with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2023;75(12):2543.
2. Katz P, Pedro S, Michaud K. Sleep disorders among individuals with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2023;75:1250–60.
3. Ma KS, Thota E, Huang JY, et al. Increased risk of temporomandibular joint disorders and craniofacial deformities in patients with juvenile idiopathic arthritis: a population-based cohort study. Int J Oral Maxillofac Surg 2022;51:1482–7.
4. Tanski W, Chabowski M, Jankowska-Polanska B, et al. Iron metabolism in patients with rheumatoid arthritis. Eur Rev Med Pharmacol Sci 2021;25:4324–35.
5. Demir S, Kucuk A, Altas M, et al. Restless leg syndrome and sleep disorders in patients with rheumatoid arthritis and its relation with anemia parameters. Acta Medica (Hradec Kralove) 2021;64:137–44.
6. Xie W, Yang X, Ji L, et al. Incident diabetes associated with hydroxychloroquine, methotrexate, biologics and glucocorticoids in rheumatoid arthritis: a systematic review and meta-analysis. Semin Arthritis Rheum 2020;50:598–607.
7. Garg A, Chilakamarri P, Koo BB. Diagnostic and treatment considerations in restless legs syndrome complicated by diabetic neuropathy [review]. Curr Diabetes Reports 2021;21:66.

DOI 10.1002/acr.25185

Consensus and agreements on the sonographic definitions of pediatric tenosynovitis: comment on the article by Collado et al

To the Editor:

We read with great interest the article by Collado et al recently published in *Arthritis Care & Research* (1), in which the authors developed a consensus among experts on the sonographic definitions of pediatric tenosynovitis using Delphi techniques. The article revealed the experts in Step 1 achieved strong group agreement (>86%) for definitions of tenosynovitis in children following the fashion of adult definitions. The final definitions were reached and validated for all tendons, except biceps tenosynovitis in children ages <4 years, after a 4-round process in Step 2. Despite encouraging results, as far as we are concerned, some methodologic issues remain.

First, since information about affiliated institutions (or countries) and academic years (or specialty seniority) of the 28 panel participants in Step 1 and 16 panel participants in Step 2 was not available in this study, the representatives of specializations cannot be fully understood. This could affect the construct validity of the expert panel in the Delphi consensus process.

Second, there was an apparent decrease in the number of participants on the expert panel from Step 1 (28) to Step 2 (16), implying that the coverage of the specializations in the panel participants may decrease in the consensus process. Therefore, the increased attrition over the subsequent step could affect the generalizability of the study results.

Third, the consensus process for this study appeared to have been implemented in a non-anonymous manner, and this may be an issue. In the Delphi approach, participants can be apprised of the responses of the other anonymous participants, having the voices of broadly gathering judgments and statements to be expressed rather than those of the few. The anonymity of the participants prevents the unnecessary influence of others who may be regarded as an expert who may be more knowledgeable than them. However, the process of consensus without anonymity can make them prone to the illusion of consensus, thus resulting in response bias (2–7). This issue requires further clarification.

Finally, we acclaim the hard work of Collado et al, meanwhile, we would like to inform readers of the potential limitations in expounding these research results. We look forward to a response.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25185.

Chun-Chieh Chen, MD, PhD
Chin-Feng Tsai, MD, PhD
Shiuan-Chih Chen, MD, PhD
sccy399@yahoo.com.tw
Chung Shan Medical University
School of Medicine and Chung Shan Medical University Hospital Health Management Center
Taichung, Taiwan

1. Collado P, Martire MV, Lanni S, et al. OMERACT International consensus for ultrasound definitions of tenosynovitis in juvenile idiopathic arthritis: systematic literature review and Delphi process. Arthritis Care Res (Hoboken) 2023;75:2277–84.
2. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies [review]. J Clin Epidemiol 2014;67:401–9.

3. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques [review]. Int J Clin Pharm 2016;38:655–62.
4. McPherson S, Reese C, Wendler MC. Methodology update: Delphi studies. Nurs Res 2018; 67:404–10.
5. Romero-Collado A. Essential elements to elaborate a study with the Delphi method [review]. Enferm Intensiva (Engl Ed) 2021;32:100–4.
6. Drumm S, Bradley C, Moriarty F. 'More of an art than a science'? The development, design and mechanics of the Delphi technique [review]. Res Social Adm Pharm 2022;18:2230–6.
7. Spranger J, Homberg A, Sonnberger M, et al. Reporting guidelines for Delphi techniques in health sciences: a methodological review [review]. Z Evid Fortbild Qual Gesundhwes 2022;172:1–11.

DOI 10.1002/acr.25198

Reply

To the Editor:

We thank Chen et al for their interest in our article, and the possibility to clarify some important methodologic aspects mentioned (1). The authors raise a question about the methodology described in the Delphi consensus process, particularly about the number of participants and the blinded response to the questions.

The study was conducted following the Outcome Measures in Rheumatology (OMERACT) stepwise approach for selecting and developing an outcome measurement instrument based on imaging (2). The authors comment on the scarcity of information about the panel participants. They are right. We did not include detailed information because we assume the reader is aware of the OMERACT Ultrasound Working Group (US WG).

The OMERACT US WG was established in 2004, with the aim to validate US-based outcome measurement instruments for rheumatic diseases (2). The participants in both steps are members of OMERACT US WG who work with US and have skill and knowledge about the topic of concern: tenosynovitis detected by US. They are health care professionals who perform pediatric US scans. Additionally, most have worked on several OMERACT studies, including the Delphi process applied to imaging instruments and patient reporting outcomes (3–4).

In regard to the decreasing number of participants, we would like to draw the reader's attention to the section "Consensus process," which included two steps. The aim of step 1 was to develop definitions. Based on the information obtained from the systematic review of the literature, the preliminary proposal was to assess whether the consensus definitions developed in adults with rheumatoid arthritis would be appropriate for children with juvenile idiopathic arthritis, rather than to develop new definitions.

The aim of step 2 was to validate the applicability of US tenosynovitis images obtained by US examination of children. Therefore, the typical Delphi method format is applied from step 2 (5). For step 2, expert recruitment was crucial because the process is strengthened by the commitment of the participants (6). Therefore, the number of participants decreased in step 2, because only the participants who declared interest in actively participating into the acquisition of standardized US images were included. There is no clear agreement on the optimal panel size and, so far, published studies show variable panel sizes (5).

We preferred to use the Delphi technique instead of the Nominal group technique as a consensus method to retain participant anonymity. Participant responses to the Delphi questionnaire were anonymous. After each iteration, participants received feedback in the form of a numeric level of agreement and a list of free-text comments.

We thank the authors for allowing us to reply and hope we have resolved their doubts about the methodology of our study.

Paz Collado, MD, PhD
paxko10@gmail.com
Hospital Universitario Severo Ochoa and Universidad Internacional Alfonso X El Sabio Madrid, Spain
Esperanza Naredo, MD, PhD
Hospital General Universitario Gregorio Marañón and Universidad Autónoma de Madrid Madrid, Spain
Maria Antonietta D'Agostino, MD, PhD
Fondazione Policlinico Universitario Agostino Gemelli IRCCS and Università Cattolica del Sacro Cuore Rome, Italy

1. Chen CC, Tsai CF, Chen SC. Consensus and agreements on the sonographic definitions of pediatric tenosynovitis. Arthritis Care Res (Hoboken) 2023;75(12):2544–5.
2. Terslev L, Naredo E, Keen HI, et al. The OMERACT stepwise approach to select and develop imaging outcome measurement instruments: the musculoskeletal ultrasound example. J Rheumatol 2019;46: 1394–1400.
3. Bruyn GA, Iagnocco A, Naredo E et al. OMERACT definitions for US pathologies and elementary lesions of rheumatic disorders 15 years on. J Rheumatol 2019;46:1388–93.
4. Roth J, Ravagnani V, Backhaus M, et al. Preliminary definitions for the sonographic features of synovitis in children. Arthritis Care Res (Hoboken) 2017;69:1217–23.
5. Drumm S, Bradley C, Moriarty F. 'More of an art than a science'? The development, design and mechanics of the Delphi technique [review]. Res Social Adm Pharm 2022;18:2230–6.
6. McMillan S, King M, Tully MP. How to use the nominal group and Delphi techniques [review]. Int J Clin Pharm 2016;38:655–62.