Arthritis Care & Research

Aims and Scope

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

Volume 74 No. 5 May 2022

Arthritis Care & Research

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

Cover design: Sandra Pulmano

♾ This journal is printed on acid-free paper.

Arthritis Care & Research

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

VOLUME 74 • May 2022 • NO. 5

Special Articles

Clinicopathologic Conference: Starting off on the Right Foot: A 22-Year-Old Woman With Leg Swelling
Alfredo Aguirre, Anatoly Urisman, and Mary Margaretten 701

Clinicopathologic Conference: Recurrent Pleural Effusions, Lymphedema, and Abnormal Nails in a 61-Year-Old Man
Avni Amratia, Madiha Ahmad, and Arezou Khosroshahi 709

Editorial: A Moving Target: Lessons From Long-Term Studies in Juvenile Idiopathic Arthritis
Jaime Guzman and Ross E. Petty 716

Pediatrics

Changing Patterns in Treatment, Remission Status, and Categories in a Long-Term Nordic Cohort Study of Juvenile Idiopathic Arthritis
Mia Glerup, Ellen D. Arnstad, Veronika Rypdal, Suvi Peltoniemi, Kristiina Aalto, Marite Rygg, Susan Nielsen, Anders Fasth, Lillemor Berntson, Ellen Nordal, and Troels Herlin, for the Nordic Study Group of Pediatric Rheumatology (NoSPeR) 719

Brief Report: Epidemiology of Childhood-Onset Systemic Lupus Erythematosus: A Population-Based Study
Maria O. Valenzuela-Almada, Mehmet Hocaoglu, Jesse Y. Dabit, Shirley-Ann Osei-Onomah, Matthew L. Basiaga, Amir B. Orandi, Rachel E. Giblon, Kamil E. Barbour, Cynthia S. Crowson, and Alí Duarte-García 728

COVID-19

Medication Interruptions and Subsequent Disease Flares During the COVID-19 Pandemic: A Longitudinal Online Study of Patients With Rheumatic Disease
Tiffany Dharia, Shilpa Venkatachalam, Joshua F. Baker, Shubhasree Banerjee, David Curtis, Maria I. Danila, Kelly Gavigan, Jennifer Gordon, Peter A. Merkel, Dianne G. Shaw, Kalen Young, Jeffrey R. Curtis, William B. Nowell, and Michael D. George 733

Brief Report: Increased Risk of COVID-19 in Patients With Rheumatoid Arthritis: A General Population-Based Cohort Study
Yilun Wang, Kristin M. D'Silva, April M. Jorge, Xiaoxiao Li, Houchen Lyv, Jie Wei, Chao Zeng, Guanghua Lei, and Yuqing Zhang 741

Spondyloarthritis

One-Year Treatment Outcomes of Secukinumab Versus Tumor Necrosis Factor Inhibitors in Spondyloarthritis: Results From Five Nordic Biologic Registries Including More Than 10,000 Treatment Courses
Bente Glintborg, Ulf Lindström, Daniela Di Giuseppe, Sella Aarrestad Provan, Bjorn Gudbjornsson, Merete Lund Hetland, Brigitte Michelsen, Johan K. Wallman, Kalle Aaltonen, Anna-Mari Hokkanen, Dan Nordström, Tanja Schjødt Jørgensen, Rebekka Lund Hansen, Arni Jon Geirsson, Kathrine Lederballe Grøn, Niels Steen Krogh, Johan Askling, Lars Erik Kristensen, and Lennart T. H. Jacobsson, on behalf of the Danish Rheumatology Database (DANBIO), Anti-Rheumatic Therapy in Sweden/Swedish Rheumatology Quality (ARTIS/SRQ), Center for Rheumatology Research (ICEBIO), Finnish Register of Biological Treatment (ROB-FIN), and Norwegian Antirheumatic Drug Register (NOR-DMARD) registries 748

Secukinumab and Sustained Reduction in Fatigue in Patients With Ankylosing Spondylitis: Long-Term Results of Two Phase III Randomized Controlled Trials
Tore K. Kvien, Philip G. Conaghan, Laure Gossec, Vibeke Strand, Mikkel Østergaard, Denis Poddubnyy, Nicole Williams, Brian Porter, Abhijit Shete, Isabelle Gilloteau, and Atul Deodhar 759

Long-Term Association Between Disease Activity and Disability in Early Axial Spondyloarthritis: Results From a Prospective Observational Study of Inflammatory Back Pain
Pedro D. Carvalho, Adeline Ruyssen-Witrand, Ana Marreiros, and Pedro M. Machado 768

Osteoarthritis

Does Screening for Depressive Symptoms Help Optimize Duloxetine Use in Knee Osteoarthritis Patients With Moderate Pain? A Cost-Effectiveness Analysis
Nora K. Lenhard, James K. Sullivan, Eric L. Ross, Shuang Song, Robert R. Edwards, David J. Hunter, Tuhina Neogi, Jeffrey N. Katz, and Elena Losina 776

Screening to Identify Postoperative Pain and Cross-Sectional Associations Between Factors Identified in This Process With Pain and Function, Three Months After Total Knee Replacement
Vikki Wylde, Emily Sanderson, Tim J. Peters, Wendy Bertram, Nicholas Howells, Julie Bruce, Christopher Eccleston, and Rachael Gooberman-Hill 790

Novel Framework for Measuring Whole Knee Osteoarthritis Progression Using Magnetic Resonance Imaging
Jeffrey B. Driban, Lori Lyn Price, Michael P. LaValley, Grace H. Lo, Ming Zhang, Matthew S. Harkey, Amanda Canavatchel, and Timothy E. McAlindon 799

Systemic Sclerosis

Osteitis in Systemic Sclerosis: A Nationwide Case–Control Retrospective Study
Cyril Cosse, Solen Kernéis, Alain Lescoat, Gregory Pugnet, Marie-Elise Truchetet, Pascal Priollet, Elisabeth Diot, Mickael Martin, François Maurier, Jean François Viallard, Christian Agard, Brigitte Granel, Sabine Berthier, Dorothée Fagedet, Bénédicte Watelet, Ségolène Toquet, David Luque Paz, Chloé Giret, Olivier Cerles, Jérémie Dion, Christelle Nguyen, Loïc Raffray, Julien Bertolino, Wendy Jourde, Claire Le Jeunne, Luc Mouthon, and Benjamin Chaigne 809

Work Productivity and Economic Burden of Systemic Sclerosis in a Multiethnic Asian Population
Ling Xiang, Sandra M. Y. Kua, and Andrea H. L. Low 818

Brief Report: Finger Systolic Blood Pressure Index Measurement: A Useful Tool for the Evaluation of Arterial Disease in Patients With Systemic Sclerosis
Sophie Blaise, Carine Boulon, Marion Mangin, Patricia Senet, Isabelle Lazareth, Bernard Imbert, François-Xavier Lapebie, Philippe Lacroix, Joël Constans, and Patrick Carpentier 828

Ehlers-Danlos Syndrome

Electromyographic Muscle Activity and Three-Dimensional Scapular Kinematics in Patients With Multidirectional Shoulder Instability: A Study in the Hypermobile Type of the Ehlers-Danlos Syndrome and the Hypermobility Spectrum Disorders
Valentien Spanhove, Patrick Calders, Kelly Berckmans, Tanneke Palmans, Fransiska Malfait, Ann Cools, and Inge De Wandele 833

Altered Multisegment Ankle and Foot Kinematics During Gait in Patients With Hypermobile Ehlers-Danlos Syndrome/Hypermobility Spectrum Disorder: A Case–Control Study
Stefan Vermeulen, Sophie De Mits, Roel De Ridder, Patrick Calders, Joris De Schepper, Fransiska Malfait, and Lies Rombaut 841

Systemic Lupus Erythematosus

Developing and Validating Methods to Assemble Systemic Lupus Erythematosus Births in the Electronic Health Record
April Barnado, Amanda M. Eudy, Ashley Blaske, Lee Wheless, Katie Kirchoff, Jim C. Oates, and Megan E. B. Clowse 849

Gout

Allopurinol and Cardiovascular Events: Time-Related Biases in Observational Studies
Samy Suissa, Karine Suissa, and Marie Hudson 858

Cover image: The figure on the cover (from Driban et al, page 799) shows the measurement of medial tibiofemoral bone marrow lesion volume (BML; mm^3) and whole-knee effusion-synovitis volume (mm^3) on paired baseline (left; top and bottom) and 24-month follow-up (right; top and bottom) magnetic resonance images, as well as the segmented BML (top; yellow) and effusion-synovitis segmentation (bottom; yellow).

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 701–708
DOI 10.1002/acr.24738

American College of Rheumatology
Empowering Rheumatology Professionals

CLINICOPATHOLOGIC CONFERENCE

Starting off on the Right Foot: A 22-Year-Old Woman With Leg Swelling

Alfredo Aguirre, Anatoly Urisman, and Mary Margaretten

CASE PRESENTATION

Chief symptoms

The patient was a 22-year-old pregnant woman at 22 weeks and 6 days of gestation with a chief symptom of acute swelling of the bilateral lower extremities.

History of present illness

A previously healthy 22-year-old woman was in the second trimester of her first pregnancy when she presented to the emergency room with new edema of the lower extremities. Her pregnancy course had been unremarkable, and she was receiving routine obstetric care from her family physician. One week prior to admission, she developed progressive swelling of the legs. Examination by her family physician revealed a systolic blood pressure of 160 mm Hg, and dipstick urinalysis was abnormal. She was directed to the local safety-net hospital and admitted to the obstetric service there for further evaluation.

Past medical, social, and family history

The patient had no remarkable past medical history and took over-the-counter vitamins. She reported not taking nonsteroidal antiinflammatory drugs (NSAIDs). She was born in the Guangdong province of China and had moved to the US 6 years earlier. The patient lived with her male partner with whom she was monogamous. She reported no tobacco or drug use and had no family history of autoimmune, cardiac, or renal disease.

Review of systems

She endorsed cough, nausea, diarrhea, and weight gain in the week leading up to her presentation to the emergency room. She denied fevers, chest pain, dyspnea, abdominal pain, Raynaud's phenomenon, joint pain or swelling, rashes, arthralgias, photosensitivity, oral ulcers, alopecia, sicca symptoms, and a prior history of clots, serositis, seizures, or cytopenia.

Physical examination

In the emergency room, the patient had a temperature of 36.5°C, blood pressure reading of 163/97 mm Hg, heart rate of 90 beats per minute, and oxygen saturation 100% on room air. She was a young, anxious-appearing woman in no acute distress. Mucous membranes were moist, and the oropharynx was clear. The neck was supple with no cervical lymphadenopathy. Cardiac examination was normal without elevated jugular venous distension, murmurs, or rubs. She had normal work of breathing, and lung fields were clear without crackles, rhonchi, or wheezing. Her abdomen was gravid without tenderness. She had 2+ bilateral lower extremity pitting edema up to her mid-shins. There were no rashes, synovitis, joint laxity, or neurologic deficits observed on examination.

Laboratory and radiographic evaluation

The patient's white blood cell count was 6,500/mm^3, hemoglobin was 11.4 gm/dl (reference range: 11.7–15.7 mg/dl), and platelet count was 242,000/mm^3. Absolute lymphocyte count was 860/mm^3. Serum creatinine level was 0.94 mg/dl; one month prior, it was 0.5 mg/dl. Blood urea nitrogen level was 32 mg/dl (reference range: 6–20 mg/dl). The remaining electrolytes were all normal. Albumin level was 1.6 gm/dl (reference range: 3.4–4.8 gm/dl), and results from liver function testing were normal. Erythrocyte sedimentation rate was 106 mm/hour, and high-sensitivity C-reactive protein level was 16 mg/liter. Urinalysis showed 3+ protein without pyuria or hematuria. A spot urine test for protein/

Alfredo Aguirre, MD, Anatoly Urisman, MD, PhD, Mary Margaretten, MD, MAS: University of California, San Francisco.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24738&file=acr24738-sup-0001-Disclosureform.pdf.

Address correspondence to Alfredo Aguirre, MD, Division of Rheumatology, University of California, San Francisco, 4150 Clement Street (111R), San Francisco, CA 94121. Email: alfredo.aguirre@ucsf.edu.

Submitted for publication March 24, 2021; accepted in revised form June 29, 2021.

creatinine ratio (UPCR) was 14.7 gm/gm. Total cholesterol level was 345 mg/dl (reference range: <200 mg/dl). Antinuclear antibody testing was positive, with a titer greater than 1:640 in a speckled pattern. Complement 3 (C3) was 78 mg/dl (reference range: 86–184 mg/dl), and C4 was 16 mg/dl (reference range: 12–40 mg/dl). A radiograph of the chest revealed small bilateral pleural effusions. A renal ultrasound demonstrated mild right-sided hydronephrosis with normal kidney size and echogenicity.

CASE SUMMARY

A previously healthy 22-year-old woman at 22 weeks and 6 days of gestation presented with acute kidney injury, lower extremity edema, hypoalbuminemia, hyperlipidemia, and severe proteinuria consistent with nephrotic syndrome.

DIFFERENTIAL DIAGNOSIS

Nephrotic syndrome. Nephrotic syndrome results from disruption of the glomerular filtration barrier, leading to urinary loss of proteins, such as albumin, from the circulation. Heavy proteinuria in nephrotic syndrome (defined as ≥3.5 gm/day) has important pathophysiologic consequences, including hypoalbuminemia, edema, hyperlipidemia, and hypercoagulability (1). The differential diagnosis of nephrotic-range proteinuria is broad and can be divided into primary renal disorders and disorders secondary to systemic diseases (Table 1).

Primary renal disorders. Membranous nephropathy (MN), minimal change disease (MCD), and focal segmental glomerulosclerosis (FSGS) were the leading causes of glomerular disease and nephrotic syndrome to be considered in this patient. Primary MN is characterized by subepithelial immune complex deposition and antibodies to M-type phospholipase A2 receptor (PLA2R) detected in the serum or on glomerular staining (2). In contrast, MCD is characterized by podocyte foot process effacement on ultrastructural analysis and has a generally favorable prognosis (3). Podocyte dysfunction is also a hallmark of FSGS but can be distinguished from MCD by the histologic features of glomerular scarring (4). Renal biopsy is critical to distinguish among these entities. Also, it is important to consider medications (e.g., NSAIDs), infections (e.g., hepatitis, HIV), malignancies, and rheumatic diseases as secondary causes of MN, MCD, and FSGS (Table 1).

Systemic lupus erythematosus (SLE). While lupus nephritis uncommonly presents de novo during pregnancy (5), lupus nephritis was a primary concern for this young woman with a positive antinuclear antibody test result and suspected glomerular disorder. Bland urine sediment observed during evaluation of the patient raised suspicion for isolated secondary MN caused by SLE (class V lupus nephritis), which occurs in 8–18% of biopsy-proven lupus nephritis and presents with nephrotic-range

Table 1. Differential diagnosis of nephrotic-range proteinuria in pregnancy*

Primary renal disorder	Renal disorder secondary to systemic disorder
Primary FSGS	Secondary FSGS Genetic (mutations in podocyte genes) Infection (HIV, human parvovirus B19, CMV, EBV) Autoimmune (SLE) Drugs/toxins (heroin, interferon therapy) Adaptive (secondary to reduced renal mass)
Primary MN	Secondary MN Infection (syphilis, HIV, HCV, HBV) Malignancy Autoimmune (SLE, MCTD, IgG4-RD) Drugs/toxins (NSAID, probenecid, lithium)
Primary MCD	Secondary MCD Infections (mycoplasma, viruses, parasites) Malignancy Autoimmune (SLE) Allergy (food and environmental allergy) Drugs (NSAID, interferon therapy) Diabetic nephropathy Amyloidosis Preeclampsia TMA (APS)

* APS = antiphospholipid syndrome; CMV = cytomegalovirus; EBV = Epstein-Barr virus; FSGS = focal segmental glomerulosclerosis; HBV = hepatitis B virus; HCV = hepatitis C virus; IgG4-RD = IgG4-related disease; MCD = minimal change disease; MCTD = mixed connective tissue disease; MN = membranous nephropathy; NSAID = nonsteroidal antiinflammatory drug; SLE = systemic lupus erythematosus; TMA = thrombotic microangiopathy. Adapted from refs. 1–4.

proteinuria in ~60% of cases (6–8). The clinical course of isolated membranous lupus nephritis is thought to be more benign than proliferative lupus nephritis, with a 10-year renal and overall survival rate of >70% (6,9,10). However, patients with membranous lupus nephritis tend to present with relatively few extrarenal manifestations and low serologic activity as compared to proliferative disease (6), which may complicate diagnosis in situations where renal biopsy is unavailable or high risk.

If SLE is confirmed, the presenting features of hypertension, proteinuria, and kidney injury should prompt consideration of antiphospholipid syndrome (APS) nephropathy, which presents with a variety of symptoms and signs depending on the renal vasculature involved (11). Examination should proceed with focused history and examination for APS, testing for lupus anticoagulant and antiphospholipid antibodies as well as renal imaging to exclude thrombus. In patients with SLE and APS nephropathy, nephrotic-range proteinuria has been observed, often in the context of concurrent lupus nephritis (12,13).

Preeclampsia. The presenting features of hypertension and proteinuria during pregnancy in this patient could also be consistent with preeclampsia, which is a relatively common disorder of pregnancy affecting ~5% of gestations (14,15). Preeclampsia is defined by new-onset hypertension (blood pressure of ≥140/90 mm Hg) and proteinuria (≥300 mg/day) or other maternal end-organ dysfunction occurring after 20 weeks of gestation (Table 2). History of prior renal disease and systemic autoimmunity increases the risk of preeclampsia, with up to 35% of pregnancies in individuals with SLE affected by this disorder (15). Preeclampsia is associated with fetal/neonatal complications (e.g., fetal growth restriction, preterm delivery) and adverse maternal outcomes (e.g., eclampsia, stroke, elevated long-term risks of cardiovascular disease), sometimes culminating in fetal or maternal death (14). In more recent classification, proteinuria is not a requirement for a diagnosis of preeclampsia, although it is present in ~75% of cases (16). Nephrotic-range proteinuria has been described in up to 20% of individuals with preeclampsia (17,18). A diagnosis of preeclampsia has significant implications for disease management, with a focus on intensified monitoring, hypertension control, management of complications, and timely delivery of the fetus (14).

Table 2. Diagnosis of preeclampsia*

1) New-onset hypertension after 20 weeks gestation (required)
Systolic blood pressure of ≥140 mm Hg or diastolic blood pressure of ≥90 mm Hg on 2 occasions at least 4 hours apart, or
Systolic blood pressure of ≥160 mm Hg or diastolic blood pressure of ≥110 mm Hg on 1 occasion
2) Organ dysfunction (≥1 of the following)
Proteinuria
≥300 mg per 24-hour urine test, UPCR of ≥0.3 gm/gm, or urine dipstick 2+
Hematologic involvement
Thrombocytopenia (platelet count of <100,000/μl)
Renal insufficiency
Serum creatinine level of >1.1 mg/dl or doubling of serum creatinine in absence of other renal disease
Impaired liver function
Transaminitis (ALT or AST >2 × ULN), severe right upper quadrant or epigastric pain
Pulmonary involvement
Pulmonary edema
Neurologic dysfunction
Visual symptoms, new-onset severe headache unresponsive to medication

* ALT = alanine aminotransferase; AST = aspartate aminotransferase; ULN = upper limit of normal; UPCR = urine protein/creatinine ratio. Adapted from the American College of Gynecology Practice Bulletin on Gestational Hypertension and Preeclampsia (ref. 27).

Preeclampsia is the consequence of abnormal placentation, which leads to maternal endothelial dysfunction and an imbalance of circulating angiogenic factors injurious to endothelial cells and podocytes (19–21). Notably, this mechanism is distinct from the largely immunologic mechanisms of the diseases described above. Patients with autoimmune disease are at higher risk for preeclampsia, and thus it may be difficult to distinguish between preeclampsia, flare of autoimmune disease, or flare of autoimmune disease with superimposed preeclampsia in clinical settings.

CLINICAL COURSE

Rheumatology was consulted to evaluate a young pregnant woman with nephrotic syndrome in the setting of a positive high-titer ANA. Additional laboratory evaluation revealed a uric acid level of 7.2 mg/dl (normal range: 2.5–7.5 mg/dl) and a 24-hour urine protein level of 7.81 gm. Serologic testing for HIV, syphilis, and hepatitis B and C was negative. The patient was positive for anti-RNP antibody at 80 AU/ml (reference range: <40 AU/ml) and negative for double-stranded DNA, SSA/SSB, and Smith antibodies. Testing for Russel's viper venom time, anticardiolipin IgG/IgM, IgG/IgM β_2-glycoprotein I, and serum anti-PLA2R was negative.

Given the high suspicion for glomerular disease, specifically membranous lupus nephritis with possible superimposed preeclampsia, we began treatment with high-dose oral glucocorticoids (1 mg/kg/day of prednisone) and hydroxychloroquine. An ultrasound-guided renal biopsy was performed without complication. The biopsy showed an adequate sample of cortical tissue with ~25 glomeruli, none of which were globally or segmentally sclerotic. On light microscopy, the glomeruli did not show endocapillary or mesangial hypercellularity (Figure 1) and were only notable for variable visceral epithelial cell prominence. Additionally, a single glomerulus demonstrated a tip-like lesion (Figure 2). There was no evidence of thrombotic microangiopathy. Focal areas of tubulointerstitial inflammation were observed without significant interstitial fibrosis or tubular atrophy. The arteries and arterioles were unremarkable. Immunofluorescence microscopy stains for IgG, IgM, IgA, C1q, C3, kappa and lambda light chains were negative. Electron microscopy demonstrated patent

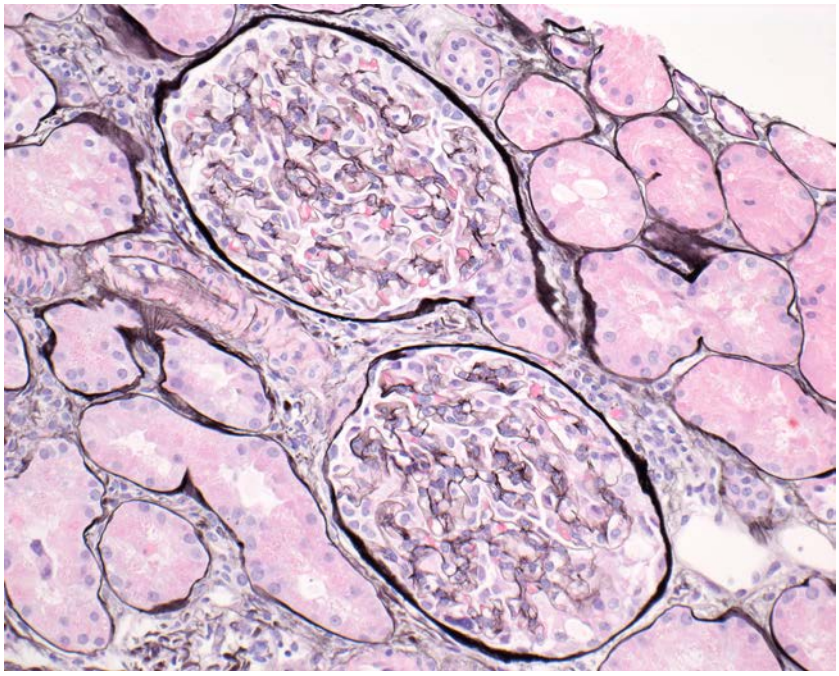

Figure 1. Methenamine silver–periodic acid–Schiff (Jones) staining of a renal biopsy specimen showing well-preserved cortical architecture. Most glomeruli did not reveal significant abnormalities aside from a prominence of variable visceral epithelial cells. Original magnification × 200.

capillary loops without endocapillary hypercellularity. The glomerular basement membranes showed normal thickness and contours. No electron-dense immune complex deposits were seen. The podocytes showed widespread foot process effacement and prominent microvillous transformation. There were focal areas of mesangial cell swelling, but no mesangial hypercellularity.

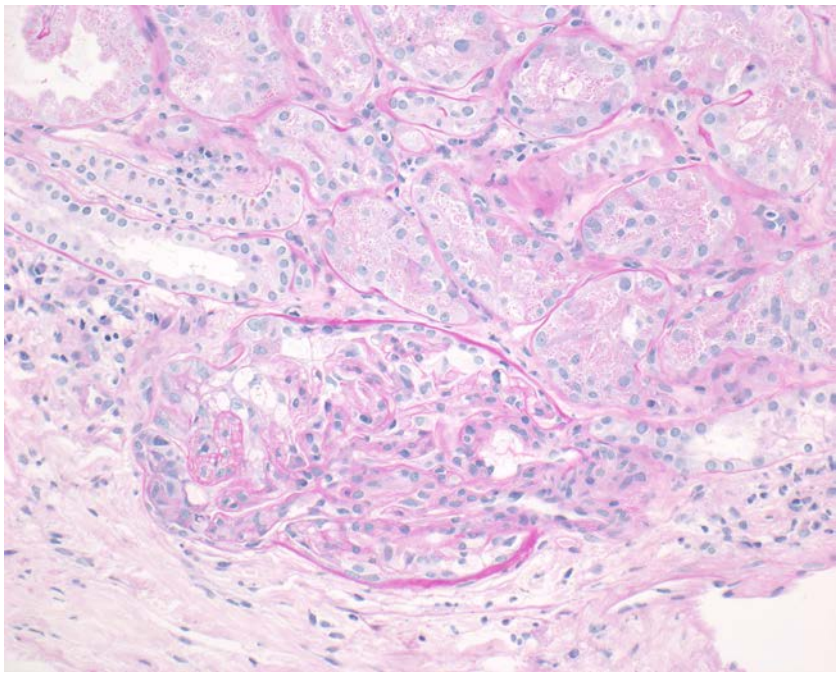

Figure 2. Periodic acid–Schiff staining of a renal biopsy specimen. A single glomerulus demonstrates a tip-like lesion at the tubular pole, with a near-collapsed distal segment surrounded by prominent visceral and parietal epithelial cell hyperplasia, including several injured epithelial cells. Tip-like lesions describe segmental acute podocyte injury involving the distal tip of the glomerulus, and these relatively nonspecific lesions can be observed in various podocytopathies, including both minimal change disease and focal segmental glomerulosclerosis. Original magnification × 200.

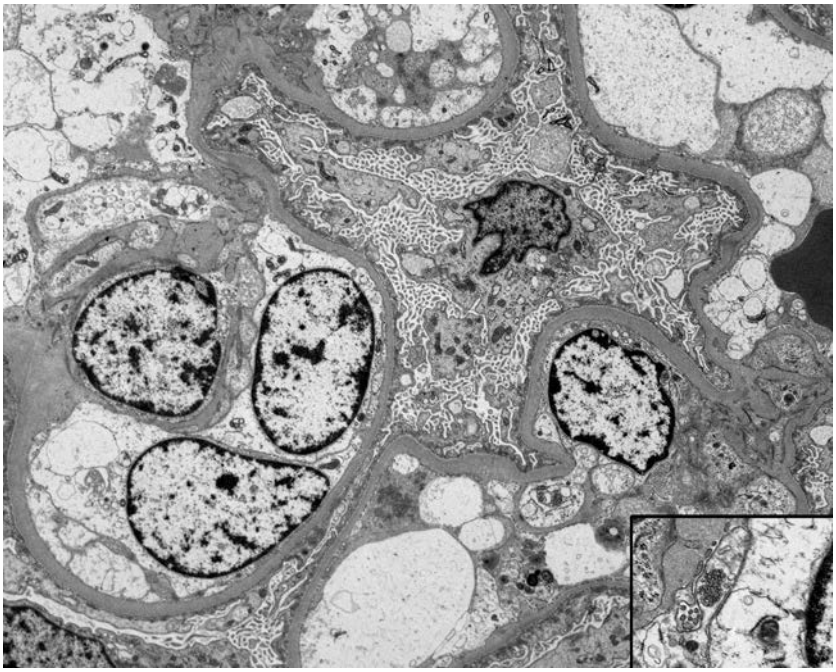

Figure 3. Electron microscopy showing diffuse podocyte foot process effacement and microvillous transformation of the podocyte cytoplasm. Variable endothelial cell swelling is also present, but no immune complex-type deposits were identified. Rare tubuloreticular inclusions are noted in the cytoplasm of glomerular capillary endothelial cells (**inset**).

Rare tubuloreticular inclusions in endothelial cell cytoplasm were noted (Figure 3). Overall, the findings were consistent with MCD secondary to SLE.

After discharge, the patient was treated with aspirin for prevention of preeclampsia. High-dose glucocorticoids for MCD were continued for one month followed by a slow taper. A decrease in proteinuria was prompt, with a UPCR of 0.21 gm/gm noted at the patient's 2-month follow-up visit. She was diagnosed with intrauterine growth restriction (estimated fetal weight <10th percentile) at 29 weeks. The rest of pregnancy course was unremarkable, and the patient had a spontaneous vaginal delivery at 37 weeks.

At a follow-up appointment 1-month post-partum, UPCR was found to be 3.0 gm/gm, and the patient reported having stopped all of her medications, including 30 mg of prednisone daily. Since she had been treated with high-dose steroids for a significant amount of time and no longer had evidence of serologic activity or other features of SLE, an angiotensin-converting enzyme inhibitor was initiated without glucocorticoids. Hydroxychloroquine was also restarted. The patient was closely monitored, and her proteinuria level decreased to 0.26 gm/gm within 2 months. Renal disease in this patient remained in remission at 1 year. The patient continues to see a rheumatologist at regular intervals.

Table 3. Distinguishing preeclampsia from lupus nephritis*

	Preeclampsia	Lupus nephritis
Hypertension	Present	May be present
Timing in pregnancy	>20 weeks of gestation	Any gestational age
SLE manifestations (arthritis, serositis, rash)	Not present	Present
ANA	Negative	Positive
Anti-dsDNA	Negative	May be positive
Complements	Normal	May be low
Serum uric acid	May be elevated	Normal (may be elevated with renal insufficiency)
Creatinine	Typically normal	Variable
Active urinary sediment	Absent	Present (in proliferative lupus nephritis)
Platelets	May be low	May be low
Liver function tests	May be elevated	Normal

* ANA = antinuclear antibody; Anti-dsDNA = anti–double-stranded DNA; SLE = systemic lupus erythematosus. Adapted from refs. 15 and 22.

DISCUSSION

This work highlights the challenges of distinguishing lupus nephritis from preeclampsia and the importance of constructing a broad differential diagnosis when evaluating a pregnant patient with suspected glomerular disease. Diagnostic clarity can be achieved by integrating history, physical examination, and laboratory data (Table 3) (15,22). The patient's presentation of hypertension and proteinuria in the second trimester of a first pregnancy could be consistent with preeclampsia. While testing for ANAs was positive, the patient did not have other overt symptoms of SLE and urine sediment was bland, pointing away from proliferative lupus nephritis as causing her symptoms. However, several features in this patient were atypical for preeclampsia. The onset of preeclampsia is usually later in pregnancy (after 34 weeks of gestation) (23). Low levels of C3 and lymphopenia were at odds with the normal physiologic changes of pregnancy, which include increased hepatic synthesis of complement proteins (24) as well as mild leukocytosis with preserved lymphocyte counts (25). The uric acid level in this patient was mildly elevated, which can be associated with preeclampsia (26,27); however, uric acid level is nonspecific and can be elevated in settings of renal insufficiency of any cause (26). These observations, in addition to positivity for ANAs and anti-RNP antibodies on testing, added to a growing diagnostic uncertainty in this patient.

Although typically not pursued in pregnancy, renal biopsy can be helpful in unclear clinical scenarios, severe renal disease, or first presentation of suspected glomerulopathy. A tissue diagnosis can direct immunosuppression in cases of biopsy-proven autoimmune glomerular disease as compared to supportive care for preeclampsia and its complications, such as hypertension (14). During pregnancy, a biopsy can also be critical in informing discussions with patients about pregnancy termination or early delivery of the fetus in cases of severe renal insufficiency or preeclampsia. While studies suggest a potential increased risk of bleeding complications for pregnant individuals undergoing this procedure, it can be clinically very useful, with up to 66% of renal biopsies in pregnant individuals leading to altered treatment decisions (28). Ultimately, this patient and her partner consented to a renal biopsy, which demonstrated MCD in the absence of immune complex–mediated tissue injury.

Case series of SLE patients with nephrotic syndrome and MCD or FSGS have been reported in the literature for decades (29–34). These syndromes are now understood as manifestations of "lupus podocytopathy," which is characterized by nephrotic-range proteinuria in a patient with SLE and diffuse podocyte foot process effacement without glomerular immune complex deposits or hypercellularity on biopsy (35). These patients typically have other evidence of active SLE, supporting a link between lupus and these renal lesions. Diffuse foot process effacement on biopsy can accompany three major patterns of injury, including mesangioproliferative lesions (the most common subtype), MCD, and FSGS (30). Histologic examination has prognostic and therapeutic implications, with best treatment response in the mesangioproliferative and MCD subtypes and higher rates of acute kidney injury, hypertension, and glucocorticoid-refractory disease in the FSGS subtype (30). In the context of low C3, lymphopenia, high-titer ANA, positive anti-RNP, and tubuloreticular inclusions on biopsy, the lesions of MCD in this patient were most likely secondary due to lupus podocytopathy.

The optimal treatment for lupus podocytopathy is still unclear, but glucocorticoid monotherapy is often used as a first-line treatment. Steroid regimens are based on recommendations for MCD—prednisone 1 mg/kg/day for a minimum of 4 weeks (or a maximum of 16 weeks if complete remission is not achieved) followed by a slow taper (35). An additional consideration in this patient was the safety of high-dose glucocorticoids during pregnancy. Moderate-to-high doses of glucocorticoids (e.g., >10–20 mg of prednisone per day) have been associated with preterm delivery and low birth weight (36–38), but uncontrolled autoimmune renal disease also carries significant risks to maternal-fetal health (15), which justified treatment with recommended doses. Hydroxychloroquine was also added to her medication regimen given its multiple benefits, including a strong safety record in pregnancy (39) and associations with improved pregnancy outcomes in SLE patients (40). Most patients with lupus podocytopathy will achieve partial or complete remission with these medications alone (29,31), although subsequent relapses occur in up to 60% of patients in long-term follow-up (30). Addition of a steroid-sparing immunosuppressive agent has been associated with reduced rates of relapse (41), although this was deferred in this patient given the favorable outcomes of the MCD subtype. She continues to be monitored closely for recurrence of proteinuria or development of other SLE features, as a relapse of MCD, transformation to FSGS, or development of frank proliferative lupus nephritis have all been described in lupus podocytopathy (30–32).

Although diagnosed with intrauterine growth restriction in her third trimester, the pregnancy course in this patient had a favorable outcome with an uncomplicated delivery at term. This work underscores the importance of considering a broad differential of renal disorders with a mechanistic approach to proteinuria during pregnancy. The decision to pursue renal biopsy during pregnancy clarified the diagnosis of a rare presentation of SLE and guided treatment with surveillance recommendations tailored to optimize renal recovery and maternal-fetal health, ensuring that this mother and her newborn both started off on the right foot.

FINAL DIAGNOSIS

Minimal change disease as a manifestation of lupus podocytopathy.

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

Study conception and design. Aguirre, Margaretten.
Acquisition of data. Aguirre, Urisman.
Analysis and interpretation of data. Aguirre, Urisman, Margaretten.

REFERENCES

1. Floege J, Johnson RJ, Feehally J. Introduction to glomerular disease. In: Comprehensive clinical nephrology. St. Louis: Elsevier; 2010. p. 193–207.
2. Couser WG. Primary membranous nephropathy. Clin J Am Soc Nephrol 2017;12:983–97.
3. Vivarelli M, Massella L, Ruggiero B, Emma F. Minimal change disease. Clin J Am Soc Nephrol 2017;12:332–45.
4. D'Agati VD. Focal segmental glomerulosclerosis. N Engl J Med 2011:14.
5. Buyon JP, Kim MY, Guerra MM, Lu S, Reeves E, Petri M, et al. Kidney outcomes and risk factors for nephritis (flare/de novo) in a multiethnic cohort of pregnant patients with lupus. Clin J Am Soc Nephrol 2017; 12:940–6.
6. Mercadal L. Factors affecting outcome and prognosis in membranous lupus nephropathy. Nephrol Dial Transplant 2002;17:1771–8.
7. Yap DY, Tang CS, Ma MK, Lam MF, Chan TM. Survival analysis and causes of mortality in patients with lupus nephritis. Nephrol Dial Transplant 2012;27:3248–54.
8. Hanly JG, Su L, Urowitz MB, Romero-Diaz J, Gordon C, Bae SC, et al. A longitudinal analysis of outcomes of lupus nephritis in an international inception cohort using a multistate model approach. Arthritis Rheumatol 2016;68:1932–44.
9. Sloan RP, Schwartz MM, Korbet SM, Borok RZ. Long-term outcome in systemic lupus erythematosus membranous glomerulonephritis. Lupus Nephritis Collaborative Study Group. J Am Soc Nephrol 1996;7:299–305.
10. Sun HO, Hu WX, Xie HL, Zhang HT, Chen HP, Zeng CH, et al. Long-term outcome of Chinese patients with membranous lupus nephropathy. Lupus 2008;17:56–61.
11. Tektonidou MG. Antiphospholipid syndrome nephropathy: from pathogenesis to treatment. Front Immunol 2018;9:1181.
12. Moss KE, Isenberg DA. Comparison of renal disease severity and outcome in patients with primary antiphospholipid syndrome, antiphospholipid syndrome secondary to systemic lupus erythematosus (SLE) and SLE alone. Rheumatol Oxf Engl 2001;40:863–7.
13. Daugas E, Nochy D, Piette JC, Hill G. Antiphospholipid syndrome nephropathy in systemic lupus erythematosus. J Am Soc Nephrol 2002:11.
14. Mol BW, Roberts CT, Thangaratinam S, Magee LA, de Groot CJ, Hofmeyr GJ. Pre-eclampsia. Lancet 2016;387:999–1011.
15. Clowse ME. Lupus activity in pregnancy. Rheum Dis Clin N Am 2007; 33:237–52.
16. Brown MA, Magee LA, Kenny LC, Karumanchi SA, McCarthy FP, Saito S, et al. The hypertensive disorders of pregnancy: ISSHP classification, diagnosis & management recommendations for international practice. Pregnancy Hypertens 2018;13:291–310.
17. Chan P, Brown M, Simpson JM, Davis G. Proteinuria in pre-eclampsia: how much matters? BJOG 2005;112:280–5.
18. Payne B, Magee LA, Côté AM, Hutcheon JA, Li J, Kyle PM, et al. PIERS proteinuria: relationship with adverse maternal and perinatal outcome. J Obstet Gynaecol Can 2011;33:588–97.
19. Phipps EA, Thadhani R, Benzing T, Karumanchi SA. Pre-eclampsia: pathogenesis, novel diagnostics and therapies. Nat Rev Nephrol 2019;15:275–89.
20. Stillman IE, Karumanchi SA. The glomerular injury of preeclampsia. J Am Soc Nephrol 2007;18:2281–4.
21. Craici IM, Wagner SJ, Weissgerber TL, Grande JP, Garovic VD. Advances in the pathophysiology of pre-eclampsia and related podocyte injury. Kidney Int 2014;86:275–85.
22. Stanhope TJ, White WM, Moder KG, Smyth A, Garovic VD. Obstetric nephrology: lupus and lupus nephritis in pregnancy. Clin J Am Soc Nephrol 2012;7:2089–99.
23. Lisonkova S, Joseph KS. Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. Am J Obstet Gynecol 2013;209:544. e1–12.
24. Buyon JP, Cronstein BN, Morris M, Tanner M, Weissman G. Serum complement values (C3 and C4) to differentiate between systemic lupus activity and pre-eclampsia. Am J Med 1986;81:194–200.
25. Kühnert M, Strohmeier R, Stegmüller M, Halberstadt E. Changes in lymphocyte subsets during normal pregnancy. Eur J Obstet Gynecol Reprod Biol 1998;76:147–51.
26. Bellomo G. Serum uric acid and pre-eclampsia: an update. Expert Rev Cardiovasc Ther 2012;10:701–5.
27. Gestational hypertension and preeclampsia: ACOG practice bulletin. Number 222. Obstet Gynecol 2020;135:e237–60.
28. Piccoli GB, Daidola G, Attini R, Parisi S, Fassio F, Naretto C, et al. Kidney biopsy in pregnancy: evidence for counselling? A systematic narrative review. BJOG 2013;120:412–27.
29. Kraft SW, Schwartz MM, Korbet SM, Lewis EJ. Glomerular podocytopathy in patients with systemic lupus erythematosus. J Am Soc Nephrol 2005;16:175–9.
30. Hu W, Chen Y, Wang S, Chen H, Liu Z, Zeng C, et al. Clinical-morphological features and outcomes of lupus podocytopathy. Clin J Am Soc Nephrol 2016;11:585–92.
31. Dube GK, Markowitz GS, Radhakrishnan J, Appel GB, D'Agati VD. Minimal change disease in systemic lupus erythematosus. Clin Nephrol 2002;57:120–6.
32. Hertig A, Droz D, Lesavre P, Grünfeld JP, Rieu P. SLE and idiopathic nephrotic syndrome: coincidence or not? Am J Kidney 2002;40:1179–84.
33. Wang SF, Chen YH, Chen DQ, Liu ZZ, Xu F, Zeng CH, et al. Mesangial proliferative lupus nephritis with podocytopathy: a special entity of lupus nephritis. Lupus 2018;27:303–11.
34. Wang Y, Yu F, Song D, Wang SX, Zhao MH. Podocyte involvement in lupus nephritis based on the 2003 ISN/RPS system: a large cohort study from a single centre. Rheumatology (Oxford) 2014;53:1235–44.
35. Oliva-Damaso N, Payan J, Oliva-Damaso E, Pereda T, Bomback AS. Lupus podocytopathy: an overview. Adv Chronic Kidney Dis 2019; 26:369–75.
36. Palmsten K, Rolland M, Hebert MF, Clowse ME, Schatz M, Xu R, et al. Patterns of prednisone use during pregnancy in women with rheumatoid arthritis: daily and cumulative dose. Pharmacoepidemiol Drug Saf 2018;27:430–8.
37. Laskin CA, Bombardier C, Hannah ME, Mandel FP, Ritchie JW, Farewell V, et al. Prednisone and aspirin in women with autoantibodies and unexplained recurrent fetal loss. N Engl J Med 1997;337: 148–53.
38. Sammaritano LR, Bermas BL, Chakravarty EE, Chambers C, Clowse ME, Lockshin MD, et al. 2020 American College of Rheumatology guideline for the management of reproductive health in rheumatic and musculoskeletal diseases. Arthritis Rheumatol 2020;72: 529–56.

39. Diav-Citrin O, Blyakhman S, Shechtman S, Ornoy A. Pregnancy outcome following in utero exposure to hydroxychloroquine: a prospective comparative observational study. Reprod Toxicol 2013;39: 58–62.
40. Duan J, Ma D, Wen X, Guo Q, Gao J, Zhang G, et al. Hydroxychloroquine prophylaxis for preeclampsia, hypertension and prematurity in pregnant patients with systemic lupus erythematosus: a meta-analysis. Lupus 2021;30:1163–74.
41. Hu WX, Chen YH, Bao H, Liu ZZ, Wang SF, Zhang HT, et al. Glucocorticoid with or without additional immunosuppressant therapy for patients with lupus podocytopathy: a retrospective single-center study. Lupus 2015;24:1067–75.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 709–715
DOI 10.1002/acr.24761

American College of Rheumatology
Empowering Rheumatology Professionals

CLINICOPATHOLOGIC CONFERENCE

Recurrent Pleural Effusions, Lymphedema, and Abnormal Nails in a 61-Year-Old Man

Avni Amratia, Madiha Ahmad, and Arezou Khosroshahi

CASE PRESENTATION

Chief symptoms

A 61-year-old man with a history of hypertension, coronary artery disease (CAD) with prior stent placement, and chronic obstructive pulmonary disease (COPD) presented with a 4-month history of recurrent pleural effusions and progressive diffuse edema.

History of present illness

Four months prior to hospital presentation, the patient began to notice new onset dyspnea on exertion, with accompanying fatigue and chest pain. Findings from the initial evaluation with troponin and electrocardiogram were negative. A radiograph of the chest was performed and revealed a moderate-sized left pleural effusion. The patient did not have a history of pleural effusions or heart failure. Therefore, he underwent diagnostic thoracentesis, and fluid studies were consistent with an exudative pleural effusion without malignant cells or microbial growth.

One week following the procedure, the patient continued to note dyspnea. A radiograph of the chest was repeated and revealed recurrence of left-sided pleural fluid. Thus, he had a second thoracentesis, with pleural studies revealing similar negative findings. Given progressive symptoms, a computed tomography (CT) scan of the chest was performed, which then revealed bilateral pleural effusions. Further diagnostic evaluation was conducted with a bronchoscopy; however, this was also nondiagnostic for underlying causes. The patient was subsequently referred to cardiothoracic surgery for a pleural biopsy.

The patient was admitted to the hospital for an elective left thoracotomy with total pulmonary decortication and a left pleural biopsy. The procedure went well without any complications. However, shortly after the procedure, he developed new onset lower extremity and scrotal edema. Evaluation for sudden onset edema included a normal transthoracic echocardiogram and negative findings on Doppler ultrasound of the lower extremities. He was started on a regimen of diuretics with some improvement of symptoms, but the underlying cause for edema remained unclear. The patient was discharged home with a plan for close outpatient follow-up.

Due to progressively worsening diffuse edema and new abdominal distention, the patient returned to our hospital for treatment 5 days after discharge. A repeat CT of the chest revealed worsening right pleural effusion, new left lower lobe consolidation, and new right upper lobe ground glass opacities. There was concern for possible hospital-associated pneumonia, and the patient was treated empirically with antibiotics. Despite diuretic therapy during admission, he continued to have progressive edema, which now involved the hands, arms, and face. His albumin level was mildly low, but further evaluation with CT of the abdomen and pelvis did not reveal pathologic changes of the liver or ascites. The patient also underwent a third thoracentesis, which remained negative for infection and malignancy. Interestingly, results from pleural biopsy revealed subacute fibrosing pleuritis with increased number of IgG4-positive plasma cells. Rheumatology was consulted for possible immunoglobulin G4–related disease (IgG4-RD).

Past medical history

The patient reported a history of CAD with previous stent placement, squamous cell carcinoma of the skin with resection, hypertension, hyperlipidemia, COPD, and erectile dysfunction.

Family and social history

Family history was notable for diabetes mellitus and myocardial infarction in his mother. He did not have a family history of lung cancer. At the time of his presentation to the hospital, the patient

Avni Amratia, MD, Madiha Ahmad, MD, Arezou Khosroshahi, MD: Emory University, Atlanta, Georgia.

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

Address correspondence to Arezou Khosroshahi, MD, Department of Internal Medicine, Emory University, 100 Woodruff Circle, Atlanta, GA 30322. Email: akhosroshahi@emory.edu.

Submitted for publication January 26, 2021; accepted in revised form July 29, 2021.

lived with his wife at home, consumed alcohol socially, and was a current smoker with a 45 pack-year history. He previously worked as a plant manager for a cement factory.

Medications and allergies

At the time of admission, the patient's medication list included aspirin, atorvastatin, metoprolol, furosemide, sildenafil, tiotropium, oxycodone, and albuterol. He had no known drug allergies.

Review of systems

The patient reported dyspnea on exertion, fatigue, edema of the bilateral lower extremities, scrotal edema, and abdominal distension. He also described a wet, but non-productive chronic cough. He denied fevers, chills, vision changes, dry eyes, oral ulcers, Raynaud's phenomenon, headaches, dizziness, weakness, paresthesia, chest pain, hematuria, rashes, arthralgias, or myalgias. He denied parotid or lacrimal gland swelling and any history of pancreatitis. He endorsed recurrent episodes of sinusitis and nails with significantly slow growth.

Physical examination

The patient was a pleasant male adult in no acute distress. He was afebrile with a pulse of 70, blood pressure reading of 114/68 mm Hg, and oxygen saturation of 95% on a 1-liter nasal cannula. He had normal mucous membranes with absence of oral or nasal ulcers. He did not have lymphadenopathy. Lacrimal, parotid, and submandibular glands were not enlarged. He had normal work of breathing with end expiratory wheezes and crackles heard throughout the bilateral lung fields. He had symmetric 2+ pitting edema of the bilateral lower extremities extending to his thighs as well as edema of the scrotum, face, and upper extremities. Furthermore, thickened skin over the dorsum of the second toe was noted, and the examiner was unable to pinch the skin, indicative of a positive Stemmer's sign. The patient did not have any rashes, nodules, or telangiectasias. His fingers had absent cuticles and lunulae with increased curvature of several nails and mild yellow staining of the peripheral left first and second nails (Figure 1). The patient's toenails were also significantly thickened and yellow in color (Figure 2).

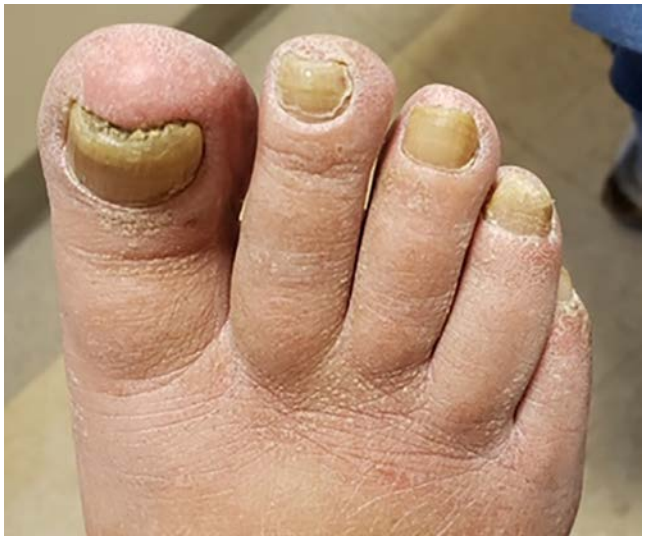

Figure 2. Thickened and yellow toenails observed in the patient.

Laboratory and radiographic evaluation

Pertinent negative laboratory findings during the patient's hospitalization are summarized in Table 1. Albumin level was 3.1

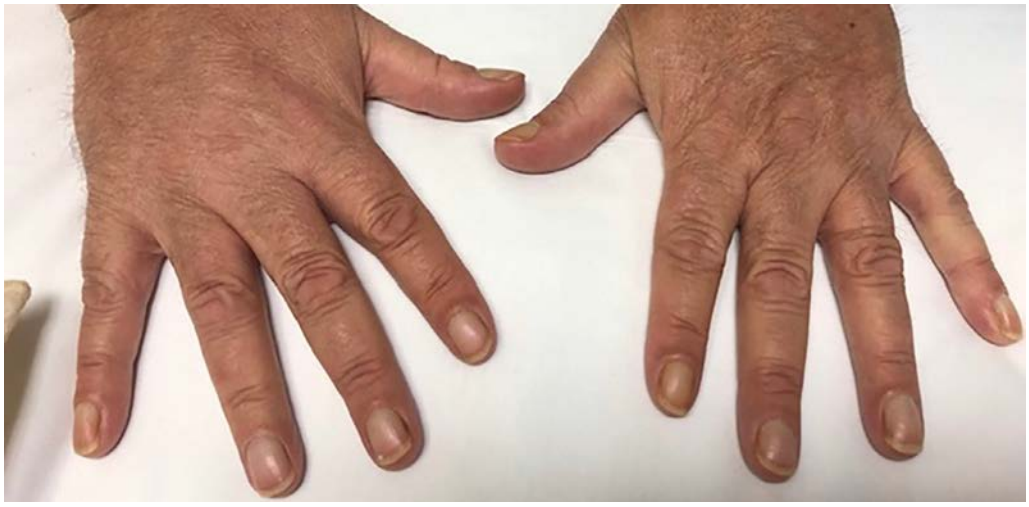

Figure 1. Absence of cuticles and increased nail curvature in several fingers of both hands. Mild yellow staining of peripheral first and second nails of the left hand was noted.

Table 1. Pertinent negative laboratory results on admission

	Laboratory value	Reference range
Sodium, mmoles/liter	140	136–145
Potassium, mmoles/liter	3.6	3.5–5.1
Chloride, mmoles/liter	99	98–107
Carbon dioxide, mmoles/liter	32	23–29
Anion gap, mmoles/liter	9	2–11
Blood urea nitrogen, mg/dl	13	7–25
Creatinine, mg/dl	0.94	0.70–1.30
Glucose, mg/dl	137	70–105
Calcium, mg/dl	9	8.6–10.3
Bilirubin (total), mg/dl	0.3	0.3–1.0
Aspartate aminotransferase, units/liter	14	13–39
Alanine aminotransferase, units/liter	14	7–52
Alkaline phosphatase, units/liter	98	34–104
Troponin, ng/ml	<0.03	≤0.04
Brain natriuretic peptide, pg/ml	42	≤99
International normalized ratio	1.04	<1.1

gm/dl (normal range: 3.5–5.7 gm/dl) with total protein of 6.5 gm/dl (normal range: 6.4–8.9 gm/dl). Complete blood cell count included a white blood cell count of 10.6/μl (normal range 4.2–9.1/μl), hemoglobin of 12.6 gm/dl (normal range: 12.9 to 16.1 gm/dl), and platelet count of 517/μl (normal range: 150–400/μl). Urine protein and urine creatinine were recorded as 8 mg/dl and 165 mg/dl, respectively, resulting in a notable protein/creatinine ratio of 0 gm/24-hour urine protein clearance.

Pleural fluid studies showed an exudative effusion with a protein ratio of 0.6. Findings were notable for a total white blood cell count of 1,750 cells/μl with a cell differential of 3% neutrophils, 87% lymphocytes, and 10% macrophages. Cholesterol value of the pleural fluid studied was 76 mg/dl. Evaluation for infectious causes was negative. This included a negative *Legionella* urine antigen screen, *Histoplasma* antigen, *Blastomyces* antibody, *Coccidiodes* antibody, bacterial blood cultures, stool culture, *Clostridium difficile* by polymerase chain reaction test, and blood Cryptococcal antigen. Furthermore, assays for syphilis, hepatitis B, and hepatitis C, and HIV were nonreactive. Additionally, bacterial, fungal, and acid-fast bacillus cultures collected from pleural fluid were negative. Bronchoalveolar lavage cultures prior to this admission grew *Mycobacterium avium* complex and Exophiala dermatitidis. However, infectious disease consultants felt these results suggested colonization and not an active infection due to the patient's immunocompetent status and lack of risk factors. QuantiFeron-TB testing was also negative.

Serum protein electrophoresis testing was unremarkable. Serologic markers including antinuclear antibody, antineutrophilic cytoplasmic antibody, double-stranded DNA antibody, and rheumatoid factor were all negative. Erythrocyte sedimentation rate was mildly elevated at 34 mm/hour (normal range: 1–30 mm/hour). Immunoglobulin levels included total IgG of 611 mg/dl (normal range: 610–1,616 mg/dl), IgA of 140 mg/dl (normal range: 84–499 mg/dl), total IgM of 216 mg/dl (normal range: 35–242 mg/dl), and total IgE level of 315 kU/liter (normal high: <180 kU/liter). Levels of IgG subclasses included an IgG1 level of 329 mg/dl (normal range: 240–1,118 mg/dl), IgG2 of 228 mg/dl (normal range: 124–549 mg/dl), IgG3 of 73 mg/dl (normal range: 21–134 mg/dl), and IgG4 34 mg/dl (normal range: 1–34 mg/dl).

A CT of the chest performed during the patient's second admission was compared to the CT scan of the chest performed earlier. The second CT showed an interval increase in the size of the moderate right pleural effusion, trace left pleural effusion, surrounding right lower lobe atelectasis, and patchy ground glass opacities of the right upper lobe and left lower lobe; lymphadenopathy was absent on the repeated CT of the chest. CT of the abdomen and pelvis was normal without evidence of ascites or lymphadenopathy. There was also no evidence of liver, pancreas, or renal abnormalities. Of note, these imaging studies did not reveal any vascular abnormalities. Scrotal ultrasound with Doppler showed normal testicles and normal Doppler signals bilaterally.

Pleural biopsy revealed fibrosing pleuritis with intense mononuclear inflammatory infiltrate and patches of non-storiform fibrosis. The inflammatory infiltrate comprised a mixture of CD3-positive T cells, CD20-positive B cells, and plasma cells. There were 71 IgG4 positive plasma cells per high-power field with a ratio of IgG4/IgG-positive plasma cells of 32%. Plasma cells were polytypic by immunoglobulin κ and λ light chains in situ hybridizations. Images from the surgical pleural biopsy are shown in Figure 3.

CASE SUMMARY

The patient was at the time of presentation a 61-year-old man with history of cardiovascular disease and COPD who presented with chronic recurrent exudative and lymphocytic pleural effusions, distinct nail changes, and diffuse body edema not explained by cardiac, renal, hepatic, or gastrointestinal disease. Pleural biopsy revealed fibrosis and lymphocytic infiltration with frequent IgG4+ plasma cells but no evidence of infection or malignancy.

DIFFERENTIAL DIAGNOSIS

Here we discuss the differential diagnosis of the constellation of recurrent pleural effusions and diffuse edema in the patient with the following possible etiologic causes.

Common causes of recurrent pleural effusions. The differential diagnosis for the most common causes of recurrent pleural effusions includes congestive heart failure, cirrhosis of the liver, nephrotic syndrome, and protein losing enteropathies. In this

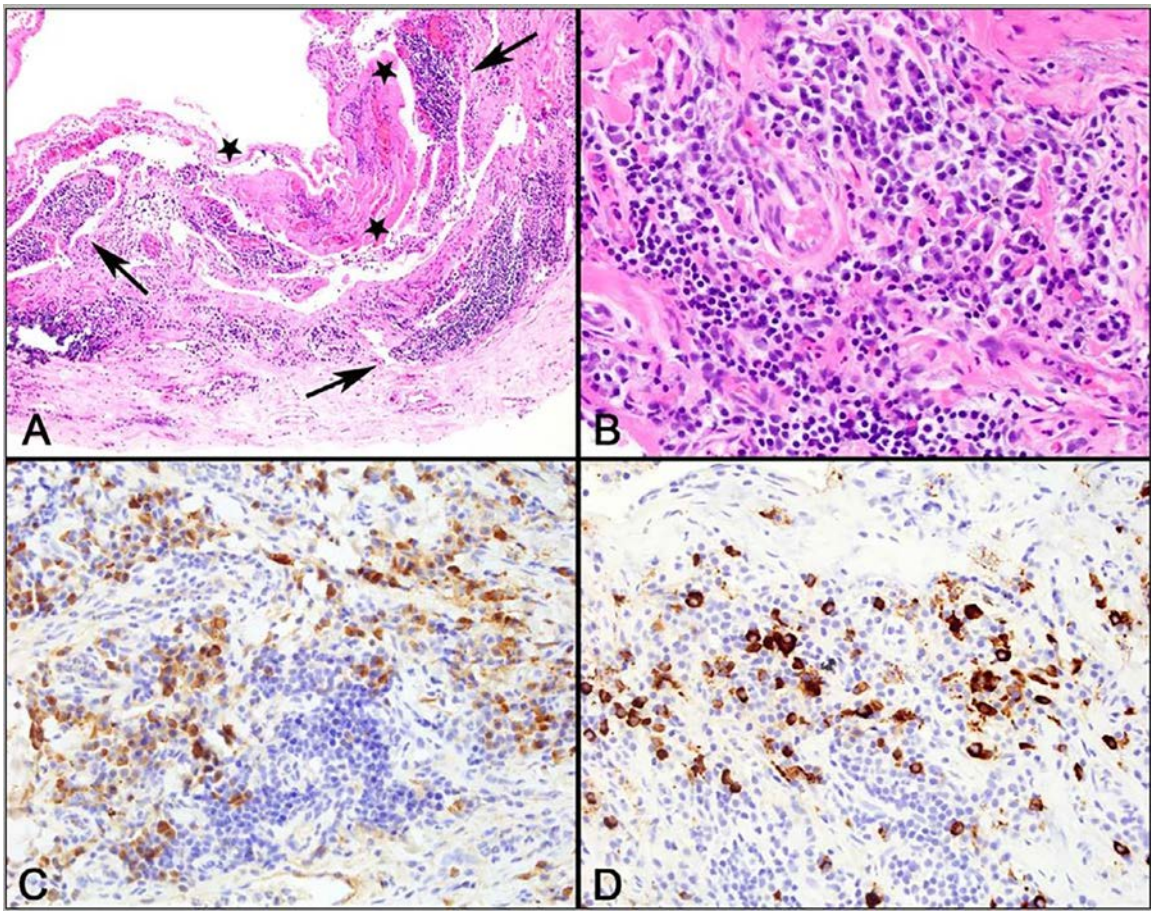

Figure 3. Surgical pleural biopsy (**A**) showing subacute pleuritis characterized by surface fibrin (**stars**) and a moderately intense mononuclear inflammatory infiltrate (**arrows**) with many plasma cells (**B**). A large subset of the plasma cells (**C**), highlighted by IgG expression, show expression of IgG4 (**D**). Sections were stained with hematoxylin and eosin in **A** and **B**, and immunohistochemical staining directed against IgG and IgG4 was performed in **C** and **D**, respectively. Original magnification × 40 in **A**; × 200 in **B**–**D**.

patient, all these diagnoses were excluded based on an extensive evaluation outlined above. Specifically, the patient had a normal echocardiogram, normal brain natriuretic peptide, and recurrent exudative pleural effusions, which argued against a diagnosis of congestive heart failure (1). We also considered liver cirrhosis and development of hepatic hydrothorax. Although the patient had recurrent effusions refractory to thoracentesis and diuretics, liver function tests, hepatitis panel, HIV testing, autoimmune markers, and international normalized ratio were all normal. More importantly, he did not have ascites or liver abnormalities on the abdomen and pelvis CT (2–4). Nephrotic syndrome was also excluded due to normal kidney function and urinalysis without proteinuria (5). The patient's presentation was most consistent with lymphedema rather than decreased plasma oncotic pressure as seen in liver failure and nephrotic syndrome. Lastly, protein-losing enteropathies must be considered in patients with recurrent pleural effusions and edema. Protein-losing enteropathy is a condition characterized by an abnormally rapid loss of serum proteins, commonly albumin, and can be associated with malignancy, infections, or autoimmune diseases (6). An absence of diarrhea and a relatively normal albumin level suggested that a diagnosis of protein-losing enteropathy was not applicable in this patient.

Malignancy. Malignancy was highly suspected given the systemic and progressive nature of his presentation. Lung cancer, lymphoma, mesothelioma, and breast cancer are the most common malignancies that are associated with malignant pleural effusions (7). Pleural effusions can be the first manifestation of metastatic cancer and are often associated with a poor prognosis. Additionally, cancer infiltration into the lymphatic system can lead to secondary lymphedema and peripheral edema. Notably, lung cancer with regional metastasis can cause thoracic outlet syndrome leading to edema of the facial and upper extremity regions (8). The patient had an extensive evaluation that included multiple thoracenteses, a bronchoscopy, a pleural biopsy, and imaging, which were all negative for malignancy.

Silicosis. Silicosis was also considered given the patient's occupation as a cement factory plant manager. Silicosis is often isolated to the lungs and produces progressive and irreversible

lung disease after prolonged dust inhalation. Typical findings of this condition include multiple discrete pulmonary nodules, interstitial lung disease, lung fibrosis, and pleural thickening. In rare cases, silicosis can produce right heart failure (9). Pleural effusions and extrapulmonary manifestations of silicosis are extremely rare presentations, but there are no significant data to suggest diffuse edema is present in this disease (10). Pleural biopsy findings were consistent with hyalinized collagen bundles with dust-filled histiocytes and interstitial fibrosis (11). Pleural biopsy findings, lack of pulmonary nodules on imaging, and a normal transthoracic echocardiogram made silicosis an unlikely cause of the patient's presentation.

Autoimmune conditions. Pleural involvement is common in patients with autoimmune diseases, including systemic lupus erythematosus, rheumatoid arthritis, and inflammatory myositis. However, the absence of a medical history typical of autoimmune disease, physical examination findings, and negative findings on serologic testing made these diagnoses unlikely for this patient.

IgG4-related disease (IgG4-RD). IgG4-RD is an inflammatory condition of unclear etiology that can manifest with single or multiple organ involvement. The most commonly affected organs include the pancreaticobiliary system, salivary glands, lung, and orbits (12). Clinical presentation of IgG4-RD can mimic many conditions including infections, malignancies, and other inflammatory diseases. Thus, thorough evaluation to exclude these mimickers is necessary prior to diagnosis.

Although the patient did have chest pain, cough, and dyspnea, which can be presenting symptoms in IgG4-RD, a CT chest did not show findings of prebronchovascular and septal thickening of airways that characterize this disease. Pleural involvement in IgG4-RD can vary from extensive adhesive fibrinous pleural thickening to incidental pleural effusions on imaging (13,14).

Lastly, while pleural biopsy revealed an increased amount of IgG4-positive plasma cells, the ratio of IgG4 to total IgG was only 32%. Diagnosis of this disease is based on an increased ratio of IgG4 to total IgG-positive plasma cells with a cutoff ratio of >40% as well as >10 IgG4-positive plasma cell/high-power field. More importantly, histopathology of IgG4-RD has 3 characteristic findings—a storiform pattern of fibrosis, obliterative phlebitis, and dense lymphoplasmacytic infiltrates with IgG4-positive plasma cells (15). Pleural biopsy findings in the patient revealed lymphoplasmacytic infiltrates. This can be seen in other conditions and is not specific for IgG4-RD. Also, he did not have any of the patterns of organ involvement typically observed in IgG4-RD.

Classification criteria were developed for IgG-RD in 2019. It should be noted that these criteria were not created for use in clinical practice or to establish a diagnosis of IgG4-RD. Instead, the intent was to provide a framework for clinicians considering IgG4-RD diagnosis and mostly for research purposes. This classification criteria were validated with a specificity of 97.8% and a sensitivity of 82% for diagnosis of IgG4-RD. Our patient met inclusion criteria because of pleural biopsy findings, but his score was only 4 (a score greater than 20 is needed to meet classification criteria for IgG4-RD) (16). When combined with the absence of other clinical or laboratory features suggestive of IgG4-RD, the pathology is not specific. Furthermore, we were not able to explain his lymphedema, recurrent pleural effusions, and nail discoloration with IgG4-RD.

Yellow nail syndrome (YNS). YNS is a rare syndrome characterized by a triad of nail changes, lymphedema, and recurring respiratory manifestations in patients 50 years or older. Its incidence has been relatively unknown with less than 400 cases documented in the literature. The pathophysiology of YNS is unclear, and in isolated cases, has been associated with autoimmune conditions and malignancy.

While YNS is characterized by a triad of features, only 2 are required for diagnosis. The complete triad is found in merely 27–60% of patients. Nail changes are the primary manifestations and include nail discoloration ranging from pale yellow to dark green, absent or insufficient cuticles, thickened nail plates, reduced longitudinal nail growth, and increased nail thickness. Respiratory manifestations are found in 56–71% of patients and include bronchiectasis, pleural effusions, sinusitis, recurrent pneumonias, and chronic cough. The most common respiratory manifestations are sinusitis and exudative pleural effusions (17–19). Pleural effusions tend to be persistent and often reoccur within days of thoracentesis. Pleural biopsies, though they do not provide diagnostic information, can be normal or show fibrosing pleuritis with lymphocytic cellular infiltrate (20,21). Lastly, lymphedema may be the first manifestation of YNS. Stemmer's sign, which is defined by inability to pinch the skin on the dorsal side or the base of the second toe, is pathognomonic on physical examination. It primarily involves the lower extremities though extension to the face, upper extremities, and peritoneal cavity has been documented (22).

CLINICAL COURSE

The patient was diagnosed with YNS following exclusion of infectious, autoimmune, and malignant causes. He started receiving high-dose vitamin E and torsemide. The patient was seen at an outpatient follow-up visit 1 month after discharge with significant improvement in edema and dyspnea on exertion. Radiographs of the chest performed 6 months after initial presentation demonstrated notable improvement with only trace bilateral pleural effusions observed.

DISCUSSION

The patient was diagnosed as having YNS. The diagnosis of YNS is one of exclusion, and we felt that the patient had an

extensive evaluation to determine other possible causes of his presentation. This included evaluation by multiple specialists, imaging, and laboratory testing. The patient was diagnosed as having YNS given the discoloration of his nails, history of slow growth, frequent sinus infections, recurrent exudative pleural effusions, chronic cough, and positive Stemmer's sign. Dermatology consultation confirmed this diagnosis. Furthermore, the results of the pleural biopsy can be explained by a diagnosis of YNS given that fibrosing pleuritis is seen in this condition. Literature review indicates histopathologic examination of the pleura in YNS usually shows features of fibrosing pleuritis with presence of plasma cells; however, IgG4 expression is limited in reported cases (23–26). Expression of IgG4 in tissue is not exclusive for YNS as IgG4 plasma cells can be present in many other inflammatory process like IgG4-RD, sarcoidosis, vasculitis, and rheumatoid arthritis.

Although YNS has been known for several decades, the pathophysiology behind it is largely unknown due to the scarcity of the disease. The leading hypothesis suggests abnormalities in lymphatic drainage causing increased microvascular permeability and subsequent protein leakage (27). Though it is difficult to confirm lymphatic impairment, it has been evaluated by lymphangiography. A study of 4 patients with YNS undergoing lower limb lymphangiography with 99mTc-colloidal antimony sulfide showed reduced uptake (impaired drainage) within the lower limb lymph nodes (28). Additionally, in comparison to primary lymphedema, there is a significantly higher rate of uptake observed in YNS patients, suggestive of decreased lymph transport rather than hypoplasia or aplasia.

Impairment in lymph drainage leading to subungual tissue sclerosis has also been thought to contribute to nail changes seen in YNS. Dilated and tortuous capillary loops can be seen on nailfold capillaroscopy. Histologic examination of nail-matrix tissues under light microscopy is notable for replacement of normal fibrovascular structure with dense collagen deposits as well as dilation and fibrosis of numerous lymph channels (29). The findings of fibrotic and dilated lymphatic vessels have also been seen in the parietal pleura of YNS patients. However, this theory does not explain many other pulmonary manifestations such as bronchiectasis and sinusitis, and full exploration has been limited (30).

Treatment for YNS is based on symptoms exhibited by the patient and can include vitamin E and antifungal drugs for nail discoloration, antibiotics as warranted for sinusitis, and compressive therapy for lymphedema. In patients with recurrent pleural effusions, somatostatin analogs such as octreotide have been used with success. It should also be noted that YNS can resolve spontaneously (31).

FINAL DIAGNOSIS

Yellow nail syndrome.

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

Study conception and design. Amratia, Ahmad, Khosroshahi.

Acquisition of data. Amratia.

Analysis and interpretation of data. Amratia, Ahmad, Khosroshahi

REFERENCES

1. Inamdar AA, Inamdar AC. Heart failure: diagnosis, management and utilization. J Clin Med 2016;5:62.
2. Alberts WM, Salem AJ, Solomon DA, Boyce G. hepatic hydrothorax: cause and management. Arch Intern Med 1991;151:2383–8.
3. Garbuzenko DV, Arefyev NO. Hepatic hydrothorax: an update and review of the literature. World J Hepatol 2017;9:1197–204.
4. Lv Y, Han G, Fan D. Hepatic hydrothorax. Ann Hepatol 2018;17: 33–46.
5. Hull RP, Goldsmith DJ. Nephrotic syndrome in adults. BMJ 2008;336: 1185–9.
6. Braamskamp MJ, Dolman KM, Tabbers MM. Clinical practice. Protein-losing enteropathy in children. Eur J Pediatr 2010;169: 1179–85.
7. Skok K, Hladnik G, Grm A, Crnjac A. Malignant pleural effusion and its current management: a review. Medicina (Kaunas) 2019;55:490.
8. Shinoda M, Akutsu H, Ohtani T, Tamura T, Satoh H. Lower limb lymphedema in lung adenocarcinoma: two case reports. Mol Clin Oncol 2016;5:478–9.
9. Yen CM, Lin CL, Lin MC, Chen HY, Lu NH, Kao CH. Pneumoconiosis increases the risk of congestive heart failure: a nationwide population-based cohort study. Medicine (Baltimore) 2016;95:e3972.
10. Salih M, Aljarod T, Ayan M, Jeffrey M, Shah BH. Pulmonary silicosis presents with pleural effusion. Case Rep Med 2015;2015:543070.
11. Mlika M, Adigun R, Bhutta BS. Silicosis. In: StatPearls. Treasure Island, Florida: StatPearls Publishing; 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK537341/.
12. Kubo K, Yamamoto K. IgG4-related disease. Int J Rheum Dis 2016; 19:747–2.
13. Campbell SN, Rubio E, Loschner AL. Clinical review of pulmonary manifestations of IgG4-related disease. Ann Am Thorac Soc 2014; 11:1466–75.
14. Inoue D, Zen Y, Abo H, Gabata T, Demachi H, Kobayashi T, et al. Immunoglobulin G4–related lung disease: CT findings with pathologic correlations. Radiology 2009;251:260–70.
15. Obiorah IE, Henao Velasquez AH, Özdemirli M. The clinicopathologic spectrum of IgG4-related disease. Balkan Med J 2018;35:292–300.
16. Wallace ZS, Naden RP, Chari S, Choi H, Della-Torre E, Dicaire JF. The 2019 American College of Rheumatology/European League Against Rheumatism classification criteria for IgG4-related disease. Arthritis Rheumatol 2020;72:7–19.
17. Vignes S, Baran R. Yellow nail syndrome: a review. Orphanet J Rare Dis 2017;12:42.
18. Dornia C, Johst U, Lange T, Kab S, Hamer OW. Yellow nail syndrome: dystrophic nails, peripheral lymphedema and chronic cough. Can Respir J 2011;18:e68–9.
19. Kurin M, Wiesen J, Mehta AC. Yellow nail syndrome: a case report and review of treatment options. Clinical Respir J 2015;11:405–10.
20. Cordasco EM Jr, Beder S, Coltro A, Bavbek S, Gurses H, Mehta AC. Clinical features of yellow nail syndrome. Cleve Clin J Med 1990;57: 472–6.

21. Valdés L, Huggins JT, Gude F, Ferreiro L, Alvarez-Dobaño JM, Golpe A, et al. Characteristics of patients with yellow nail syndrome and pleural effusion. Respirology 2014;19:985–92.
22. Piraccini BM, Urciuoli B, Starace M, Tosti A, Balestri R. Yellow nail syndrome: clinical experience in a series of 21 patients. J Dtsch Dermatol Ges 2014;12:131–7.
23. Dallmann A, Attanoos RL. (2018). Yellow nail syndrome with bilateral pleural plaques and diffuse pleural thickening: a mimic of asbestos-related disease. Case Rep Pulmonol 2018;2018: 7302898.
24. Solal-Céligny P, Cormier Y, Fournier M. The yellow nail syndrome. Light and electron microscopic aspects of the pleura. Arch Pathol Lab Med 1983;107:183–5.
25. David I, Crawford FA Jr, Hendrix GH, Harley RA, Tucker T. Thoracic surgical implications of the yellow nail syndrome. J Thorac Cardiovasc Surg 1986;91:788–90.
26. Hase I, Kurasawa K, Takizawa H, Yamaguchi B, Sakuma H, Ishii Y. (2018). Hyperplasia of lymphoid follicles and lymphangiectasia in the parietal pleura in bucillamine-induced yellow nail syndrome. Inter Med 2018;57:1887–92.
27. Preston A, Altman K, Walker G. Yellow nail syndrome [case report]. Proc (Bayl Univ Med Cent) 2018;31:526–7.
28. Marks R, Ellis JP. Yellow nails: a report of six cases. Arch Dermatol 1970;102:619–23.
29. DeCoste SD, Imber MJ, Baden HP. Yellow nail syndrome. J Am Acad Dermatol 1990;22:608–11.
30. Kato T, Akagawa S, Kusaka K, Kawashima M, Ohshima N, Kitani M, et al. An autopsy case report of yellow nail syndrome coincided with primary biliary cholangitis. Respir Med Case Rep 2021;32:101332.
31. Cheslock M, Harrington DW. Yellow Nail Syndrome In: StatPearls. Treasure Island, Florida: StatPearls Publishing; 2022. URL: https://www.ncbi.nlm.nih.gov/books/NBK557760/?report=classic.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 716–718
DOI 10.1002/acr.24858

EDITORIAL

A Moving Target: Lessons From Long-Term Studies in Juvenile Idiopathic Arthritis

Jaime Guzman and Ross E. Petty

In this issue of *Arthritis Care & Research*, Glerup et al compare data on juvenile idiopathic arthritis (JIA) categories, disease activity, and treatments at disease onset, at 8 years later, and at 18 years later in the Nordic JIA cohort (1). The findings in this study, together with detailed 18-year findings published in 2020 (2), exemplify 4 main lessons learned from recent long-term JIA cohorts.

Lesson 1: if you reapply classification criteria, JIA categories become a moving target. The current study confirms that strict ascertainment and application of International League of Associations for Rheumatology (ILAR) JIA classification criteria (3) years after diagnosis results in reclassification of roughly 10–30% of patients. Should we be surprised by this? After all, it is well known to rheumatologists that psoriasis may develop years after arthritis and that nonradiographic spondyloarthropathy evolves into ankylosing spondylitis in some patients. Yet, it is important that we classify patients in a uniform way early in the disease to accurately advise patients and families on what to expect and to define groups for inclusion in mechanistic and treatment studies. Glerup and colleagues have clearly shown that patients continue to shift across JIA categories up to 18 years after diagnosis (1,2), so retrospective reclassification would produce different groupings depending on when the reclassification is done. And, obviously, retrospective reclassification can not be used to select participants for prospective studies early in the disease.

Changes to family history were a main reason for reclassification in the study by Glerup et al, and we know that family history of psoriasis is also the most common reason for classification of undifferentiated arthritis (4). Some JIA categories (e.g., systemic arthritis and rheumatoid factor–positive polyarthritis) were more stable than others (1), a fact recognized in recent efforts to revise the ILAR classification (5). From their inception, it was recognized that JIA classification criteria would need to be revised periodically, but until we have agreement on something demonstrably better, using the current criteria is much preferable to each study using different categories. Further, we should apply the criteria at the point where it is recommended to do so, which is 6 months after disease onset.

Lesson 2: JIA long-term prognosis has improved but is not optimal. In the 1980s it was believed that most children with arthritis would outgrow the disease (6). Although this belief was an unsubstantiated claim, it was a frequent clinical observation that the disease would eventually burn out and no longer exhibit active inflammation. Burned-out disease, however, often meant life-long disability (Steinbrocker criteria [7] class 3 or 4 in 10–30% of patients) (6), stunted growth, shortened lifespan, and multiple surgeries. Those who did bear witness to the signs of burned-out disease would regard the results reported by Glerup et al as a glass half full, even though others may look at them as a glass half empty because half of the patients did not fulfill criteria for remission at 8 years and 18 years. Of note, 18 years after disease onset, >75% of participants had a Health Assessment Questionnaire Disability Index score of 0 or 0.125 and no active joints (2). Further, 70% of participants were not taking any treatment by 18 years after onset, and 45% of patients in fact had not received any disease-modifying antirheumatic drug (DMARD) for 10 years! In oncology, such a treatment course would be called a cure.

Lesson 3: prognosis for some JIA categories has improved more than for others. Systemic arthritis used to be considered the most severe JIA category, but long-term cohorts in the Nordic countries, Canada, and The Netherlands

Dr. Guzman's work was supported by the British Columbia Children's Hospital Research Institute (Clinical Investigator award).

Jaime Guzman, MD, MSc, Ross E. Petty, MD, PhD: British Columbia Children's Hospital and the University of British Columbia, Vancouver, British Columbia, Canada.

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

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24858&file=acr24858-sup-0001-Disclosureform.pdf.

Address correspondence to Jaime Guzman. 4500 Oak Street, Suite K4-122, Vancouver, British Columbia, Canada, V6H 3N1. Email: jguzman@cw.bc.ca.

Submitted for publication December 29, 2021; accepted in revised form January 6, 2022.

have shown that this is no longer the case (1,8–10). Although patients are acutely ill at onset, it now appears that systemic JIA is the category with the best long-term prognosis, whether it is treated early with prednisone and methotrexate (9) or with biologics (10). Enthesitis-related arthritis now appears to be a category with a more protracted course and long-term need for medications than previously realized (1,2). The Nordic cohort included very few patients with rheumatoid factor–positive polyarthritis, but other studies suggest this is a persisting problematic disease (8,9).

Oligoarthritis used to be considered the most benign category. This opinion may still be valid if only children with persistent oligoarthritis are considered. Such a consideration, however, would be a self-fulfilling prophecy because the moment the disease proves to be aggressive (extends beyond 4 joints), patients are taken out of the category. Of 230 participants with oligoarthritis at disease onset, Glerup et al reclassified 80 participants as having extended oligoarthritis by the 8-year follow-up and a further 17 participants by the 18-year follow-up (1). The separation of persistent and extended oligoarthritis clearly does not help for counseling patients at disease onset. Patients may be better served when oligoarthritis is described as a single group, and extension as an undesirable outcome. By providing such a description, oligoarthritis does not seem very benign, with up to 30% of patients not in remission 30 years after disease onset (8).

Lesson 4: patients may accept minimally active disease rather than taking treatment. Glerup et al reported that many patients categorized as having active disease were not receiving any treatment at the 18-year follow-up, and that only 42% of the overall cohort were followed regularly by a physician (2). In a 30-year follow-up study by Selvaag et al, 44% of patients categorized as having active disease were not receiving any DMARDs (8). These findings could have 3 explanations. First, remission criteria may be too stringent, since in some patients morning stiffness lasting >15 minutes was the only sign of active disease (2,11). Second, patients may consider minimal disease activity as an acceptable state, preferable to the inconveniences and risks of treatment (12). Third, these patients had an unsuccessful transition to adult care, and they should be reconnected to medical care and treated. In our opinion, the second explanation is the most likely, since receiving treatment has a significant psychosocial impact and risks of side effects that decrease quality of life (13).

The study by Glerup et al should be interpreted in light of 2 important considerations. First, 17% of patients were lost to follow-up 18 years after disease onset, and a further 20% declined to attend an in-person assessment. Those who declined had twice the chance of being in remission relative to those who volunteered to attend, based on participants' subjective assessments. Overall, 291 of 510 (57%) eligible participants attended in-person assessments at all 3 study points (1). Significant attrition is a major problem in most JIA long-term studies published to date and may mean that the true long-term prognosis is better than what is reported. Patients who are doing well may be less likely to volunteer for a clinic visit years after they have left the care of the pediatric rheumatologist.

Second, researchers in the Nordic cohort made a conscientious effort to recruit all children with JIA within defined geographic regions in Scandinavian countries renowned for their universal health care systems and health care registries. In deciding to what extent the JIA prognosis reported by Glerup et al (1,2) applies to their patients, clinicians should be mindful of referral biases and ethnic and health care differences that may influence their own clinic populations.

In conclusion, recent reports from long-term JIA cohorts paint a very different picture than what JIA looked like in the 1980s but, clearly, it is still a life-long concern for a sizable subset of patients. Further, long-term prognosis will continue to change as treatments evolve from a methotrexate era to a biologic era (14). Call us optimistic, but while JIA long-term prognosis has become a moving target, we see recent results as a very encouraging half-full glass.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Glerup M, Arnstad ED, Rypdal V, Peltoniemi S, Aalto K, Rygg M, et al. 2022. Changing patterns in treatment, remission status and categories in a long-term Nordic cohort study of juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 74;5:719–27.
2. Glerup M, Rypdal V, Arnstad ED, Ekelund M, Peltoniemi S, Aalto K, et al for the Nordic Study Group of Pediatric Rheumatology. Long-term outcomes in juvenile idiopathic arthritis: eighteen years of follow-up in the population-based Nordic juvenile idiopathic arthritis cohort. Arthritis Care Res (Hoboken) 2020;72:507–16.
3. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al, and the International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
4. Chan MO, Petty RE, Guzman J, ReACCh-Out Investigators. A family history of psoriasis in a first-degree relative in children with JIA: to Include or Exclude? J Rheumatol 2016;43:944–7.
5. Martini A, Ravelli A, Avcin T, Beresford MW, Burgos-Vargas R, Cuttica R, et al for the Pediatric Rheumatology International Trials Organization (PRINTO). Toward New Classification Criteria for Juvenile Idiopathic Arthritis: First Steps, Pediatric Rheumatology International Trials Organization international consensus. J Rheumatol 2019;46:190–7.
6. Ravelli A. Toward an understanding of the long-term outcome of juvenile idiopathic arthritis. Clin Exp Rheumatol 2004;22:271–5.
7. Steinbrocker O, Traeger CH, Batterman RC. Therapeutic criteria in rheumatoid arthritis. JAMA 1949;140:659–62.

8. Selvaag AM, Aulie HA, Lilleby V, Flatø B. Disease progression into adulthood and predictors of long-term active disease in juvenile idiopathic arthritis. Ann Rheum Dis 2016;75:190–5.
9. Chhabra A, Robinson C, Houghton K, Cabral DA, Morishita K, Tucker LB, et al. Long-term outcomes and disease course of children with juvenile idiopathic arthritis in the ReACCh-Out cohort: a two-centre experience. Rheumatology (Oxford) 2020;59: 3727–30.
10. Ter Haar NM, van Dijkhuizen EHP, Swart JF, van Royen-Kerkhof A, El Idrissi A, Leek AP, et al. Treatment to target using recombinant interleukin-1 receptor antagonist as first-line monotherapy in new-onset systemic juvenile idiopathic arthritis: results from a five-year follow-up study. Arthritis Rheumatol 2019;71:1163–73.
11. Allegra M, Gicchino MF, Giancane G, Alongi A, Tibaldi J, Mazzoni M, et al. The effect of morning stiffness duration on the definition of clinically inactive disease in juvenile idiopathic arthritis. J Rheumatol 2020;47:1238–41.
12. Consolaro A, Bracciolini G, Ruperto N, Pistorio A, Magni-Manzoni S, Malattia C, et al for the Paediatric Rheumatology International Trials Organization. Remission, minimal disease activity, and acceptable symptom state in juvenile idiopathic arthritis: defining criteria based on the juvenile arthritis disease activity score. Arthritis Rheum 2012;64:2366–74.
13. Oen K, Tian J, Loughin TM, Shiff NJ, Tucker LB, Huber AM, et al, for the ReACCh-Out Investigators. Causal pathways to health-related quality of life in children with juvenile idiopathic arthritis: results from the ReACCh-Out cohort. Rheumatology (Oxford) 2021;60:4691–702.
14. Giancane G, Muratore V, Marzetti V, Quilis N, Benavente BS, Bagnasco F, et al. Disease activity and damage in juvenile idiopathic arthritis: methotrexate era versus biologic era. Arthritis Res Ther 2019;21:168.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 719–727
DOI 10.1002/acr.24857

American College of Rheumatology
Empowering Rheumatology Professionals

Changing Patterns in Treatment, Remission Status, and Categories in a Long-Term Nordic Cohort Study of Juvenile Idiopathic Arthritis

Mia Glerup,[1] Ellen D. Arnstad,[2] Veronika Rypdal,[3] Suvi Peltoniemi,[4] Kristiina Aalto,[5] Marite Rygg,[6] Susan Nielsen,[7] Anders Fasth,[8] Lillemor Berntson,[9] Ellen Nordal,[3] and Troels Herlin,[1] for the Nordic Study Group of Pediatric Rheumatology (NoSPeR)

Objective. To explore sustainability of achieved remission off medication and defined International League of Associations for Rheumatology (ILAR) categories in juvenile idiopathic arthritis (JIA) and describe the trajectory of disease course over time by comparing treatment, disease activity, and ILAR categories from baseline, 8 years, and 18 years after disease onset.

Methods. A total of 373 of the 510 included patients were initially recruited consecutive cases of JIA from the prospective, longitudinal, population-based Nordic JIA cohort with disease onset during 1997–2000 from Denmark, Norway, Sweden, and Finland in an 18-year follow-up study. Clinical data were collected consecutively at baseline, 8 years, and 18 years after disease onset and were evaluated regarding treatment, disease activity, and ILAR category.

Results. Significantly more patients (70%) were off medication after 18 years of follow-up compared to after 8 years (59.7%); nevertheless, the number of patients in remission had not increased (52% off medication versus 51% on medication). Twelve percent of patients changed ILAR category between 8 years and 18 years after disease onset. Almost half of the changes were due to updated information about heredity in a first-degree relative. In the same period, the psoriatic arthritis group increased significantly in number ($P < 0.001$), in contrast to the oligoarticular category, which decreased ($P = 0.02$). The undifferentiated group increased 24% from 8 to 18 years of follow-up; however, this increase was not significant ($P = 0.06$).

Conclusion. In this Nordic JIA cohort study, the remission rate did not increase even though significantly more patients were off medication at the 18-year follow-up compared to at the 8-year follow-up after disease onset. The distribution of patients in the ILAR categories continued to change significantly throughout the 18-year study period.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is a chronic immune-mediated inflammatory disease in childhood with a miscellaneous disease spectrum from monophasic to chronic and an often fluctuating and unpredictable disease course. The variability of outcome and complications warrant grouping into homogeneous categories according to distinct phenotypes, pathophysiology, biochemical findings, disease course, and prognosis. Distinguishing the different classification criteria is essential for clinical trials and epidemiologic studies, such as long-term outcome investigations. According to the International League of Associations for Rheumatology (ILAR) consensus-based classification criteria, there are 7 exclusive JIA categories, including systemic JIA, oligoarthritis (persistent and extended), polyarticular rheumatoid factor (RF) negative, polyarticular RF positive, juvenile psoriatic, enthesitis-related (ERA), and undifferentiated (1). Undifferentiated JIA is applicable in cases where the criteria for other categories

[1]Mia Glerup MD, PhD, Troels Herlin, MD, DMSc: Aarhus University Hospital, Aarhus, Denmark; [2]Ellen D. Arnstad, MD, PhD: Norwegian University of Science and Technology, Trondheim, Norway, and Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway; [3]Veronika Rypdal, MD, Ellen Nordal, MD, PhD: University Hospital of North Norway and UiT The Artic University of Norway, Tromsø, Norway; [4]Suvi Peltoniemi, MD: Helsinki University Hospital, Helsinki, Finland; [5]Kristiina Aalto, MD, PhD: University of Helsinki, Helsinki, Finland; [6]Marite Rygg, MD, PhD: Norwegian University of Science and Technology and St. Olavs Hospital Trondheim University, Trondheim, Norway; [7]Susan Nielsen, MD: Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; [8]Anders Fasth, MD, PhD: Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden; [9]Lillemor Berntson, MD, PhD: Uppsala University, Uppsala, Sweden, for the Nordic Study Group of Pediatric Rheumatology (NoSPeR)

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24857&file=acr24857-sup-0001-Disclosureform.pdf.

Address correspondence to Mia Glerup, MD, PhD, Department of Pediatrics, Aarhus University Hospital, Palle Juul-Jensens Blvd. 99, 8200 Aarhus N. Email: miagleru@rm.dk.

Submitted for publication March 25, 2021; accepted in revised form January 6, 2022.

SIGNIFICANCE & INNOVATIONS

- We found that significantly more patients were off medication 18 years after disease onset compared to patients at the 8-year follow-up visit, but the number of patients in remission off medication had not increased correspondingly.
- The distribution of patients within the International League of Associations for Rheumatology (ILAR) categories defined from baseline were not sustained but changed significantly even beyond 8 years after disease onset.
- Almost half of the changes in the distribution between the ILAR categories were caused by updated information on heredity in a first-degree relative obtained during follow-up visits.

are neither met nor allow unambiguous classification. The ILAR criteria proposed in 1995 (2) were revised in 1998 (3) and in 2004 (1) to correct misconceptions. However, ongoing criticism about the current criteria has been raised and, contrary to the intentions of the ILAR criteria, the distribution of patients for most categories tends to change over time in the first decade of the disease course (4).

Current treatment recommendations propose the use of medication tailored according to clinical manifestations, as previously described (5–10). The scale of research in this field is still advancing, and modern therapies have evolved the outcomes of JIA, due to the increasing variety of targeted therapies that are available (11,12). Etanercept was the first biologic drug studied in polyarticular JIA, and the randomized controlled trial (RCT) was published in 2000 (13). In the subsequent 2 decades, several RCT studies on anti–tumor necrosis factor (anti-TNF) agents, CD28 receptor antagonist, interleukin-1 (IL-1) inhibitors, and anti–IL-6 receptor antagonist have been published in JIA (14–16). Despite the revolutionary leap in treatment options, we found that JIA continues to be an ongoing chronic disease, with only 33% of individuals in clinical remission off medication even 18 years after disease onset (17). Furthermore, we have shown that the ILAR categories defined at disease onset change considerably during the first 8 years of disease course (4).

In the past 2 decades, 2 other studies have reported longitudinal data on long-term outcomes with a follow-up of >10 years (18,19) but both the 15- and 17-year follow-up in these studies are from the prebiologic era before year 2000. No previous study has reported on the longitudinal changes in ILAR criteria beyond 8 years after disease onset. There is a shortage of conclusive data about the sustainability of the defined ILAR categories, medical treatment, and achieved remission beyond the first decade of disease; these concerns have been addressed in the present study. We aimed to investigate the longitudinal trajectory of JIA disease course over time by comparing ILAR categories, treatment, and disease activity from baseline and 8 years and 18 years of disease.

PATIENTS AND METHODS

Study design. We performed a multicenter, prospective, population-based study of the Nordic JIA cohort, including data collected consecutively at baseline and at 8-year and 18-year follow-up visits. Study participants included were all consecutively, newly diagnosed patients from defined geographic areas of Denmark, Norway, Sweden, and Finland with onset of JIA between January 1, 1997 and June 30, 2000. JIA was classified according to the ILAR criteria (1). A total of 510 participants were included in the cohort.

For the 3 study visits, all previously included participants were invited regardless of disease status. Data from baseline and the 8-year and 18-year follow-up visits have previously been published individually (4,17,20).

Inclusion criteria. All eligible participants fulfilled the ILAR criteria (1) and had at least 3 study visits at baseline and 8 years and 18 years after disease onset. There were no exclusion criteria.

Data collection. Demographic characteristics, treatment, disease characteristics, and blood samples were collected at the study visits. Additionally, a clinical examination was performed. We offered a standardized telephone interview to those who could not attend a study visit and a cross-check of the validity of the information was performed in the medical records.

Treatment. Medications were categorized as nonsteroidal antiinflammatory drugs (NSAIDs), systemic steroids taken at the time of the follow-up visit, conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), and biologic DMARDs (bDMARDs). Methotrexate, azathioprine, hydroxychloroquine, leflunomide, sulfasalazine and mycophenolate mofetil were the included csDMARDs. The bDMARDs used included etanercept, infliximab, adalimumab, certolizumab, golimumab, rituximab, abatacept, anakinra, and tocilizumab. DMARDs refers to the use of csDMARDs and/or bDMARDs.

Inactive disease and remission. We applied the preliminary Wallace criteria (21) for clinical inactive disease, which includes: 1) no joints with active disease; 2) no fever, rash, serositis, splenomegaly, or generalized lymphadenopathy attributable to JIA; 3) no active uveitis; 4) normal erythrocyte sedimentation rate (ESR) and/or C-reactive protein (CRP) level; and 5) a physician global assessment of the disease activity (PhGA), with the best score attainable indicating no disease activity. For clinical remission on medication, the criteria for inactive disease on systemic antiinflammatory medication had to be fulfilled for a minimum of 6 continuous months. For clinical remission off

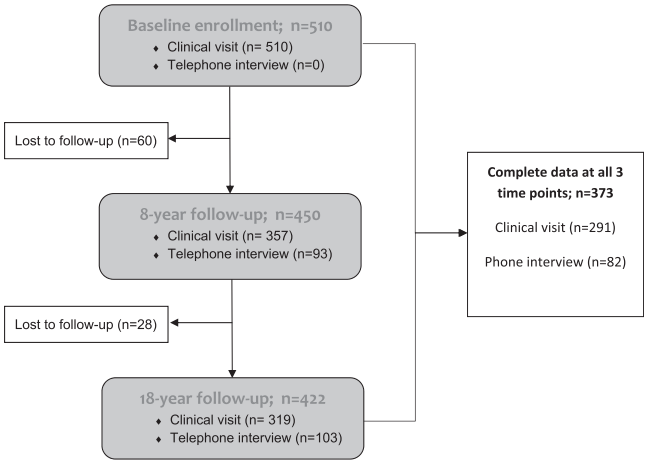

Figure 1. Flow chart of the study population throughout the 18-year observation period.

medication, the patients must have had inactive disease for a continuous period of at least 12 months off all anti-arthritis and anti-uveitis medication (21).

An joint with active disease was defined as a joint with swelling and/or a joint with limitation on motion accompanied by pain and/or tenderness. We defined a normal ESR as a value <20 mm/first hour, and a normal CRP level as <10 mg/liter. PhGA was assessed on a visual analog scale.

Statistical analysis. Descriptive statistics were used to describe demographic and clinical characteristics. Pearson's chi-square test was used to analyze odds ratios (ORs) in categorical, unmatched data. Fisher's exact test was used to analyze risk ratios (RRs) and 95% confidence intervals (95% CIs) between groups for categorical variables, and Mann-Whitney U test was used to compare medians for continuous variables. For the within-person analysis shown, we used the case–control OR calculator in STATA.

Ethical approval. Approval from medical ethics committees and data protection authorities were obtained. Written informed consent from all participants was achieved according to the regulations of each participating country.

RESULTS

Of 510 eligible participants, 422 (82.7%) attended the 18-year follow-up visit. In total, 373 completed at least 3 follow-up visits (Figure 1), at a median of 7 months, 8.1 years, and 17.6 years after disease onset (Table 1). At the time of inclusion, all participants attended a clinical visit. At the 8- and 18-year follow-up visits, 357 of 450 patients (79.3%) and 319 of 422 patients (75.6%) attended a clinical visit, respectively, while the remaining patients participated through a telephone interview

Table 1. Characteristics of the Nordic JIA cohort of participants with a follow-up at baseline, 8-years, and 18-years and those with a follow-up at all 3 time points (baseline, 8 years, and 18 years)*

	Baseline (n = 510)	8-year follow-up (n = 450)	18-year follow-up (n = 422)	Follow-up at 3 time points (n = 373)
Female sex, no./total no. (%)	340/510 (66.7)	299/450 (66.4)	288/422 (68.2)	249/373 (66.8)
Age at onset, median (IQR) years	5.9 (2.8–10.0)	5.5 (2.5–9.7)	5.7 (2.6–9.7)	5.5 (2.3–9.4)
Age at follow–up, median (IQR) years	6.6 (3.3–10.8)	14.2 (10.6–17.6)	23.4 (20.3–27.3)	23.3 (20.2–27.3)
Follow-up time, median (IQR) years	0.6 (0.5–0.7)	8.2 (7.9–8.5)	17.6 (16.8–18.4)	17.6 (16.8–18.4)
ANA positivity, no./total no. (%)	154/442 (34.8)	148/434 (34.1)	141/383 (36.8)	133/365 (36.4)
HLA–B27 positivity, no./total no. (%)	104/481 (21.6)	94/433 (21.7)	87/406 (21.4)	79/363 (21.8)
JIA categories				
Systemic JIA	18 (3.5)	17 (3.8)	14 (3.3)	13 (3.5)
Persistent oligoarthritis	275 (53.9)†	139 (30.9)	113 (26.7)	98 (26.2)
Extended oligoarthritis	–	80 (17.8)	84 (19.9)	78 (20.9)
Polyarticular RF negative	108 (21.1)	81 (18.0)	73 (17.3)	68 (18.2)
Polyarticular RF positive	10 (2.0)	4 (0.9)	6 (1.4)	3 (0.8)
Psoriatic arthritis	9 (2.0)	14 (3.1)	28 (6.6)	23 (6.1)
ERA	38 (7.5)	47 (10.4)	41 (9.7)	37 (9.9)
Undifferentiated arthritis	52 (10.2)	68 (15.1)	63 (14.9)	56 (15.0)
Treatment				
No systemic treatment	402 (78.8)	289 (64.2)	268 (63.5)	–
NSAID	262 (51.3)	51 (11.3)	79 (18.7)	–
Monotherapy NSAID	187 (36.7)	3 (0.7)	2 (0.5)	–
Systemic corticosteroids	27 (5.3)	5 (1.1)	4 (0.9)	–
Monotherapy csDMARDs	71 (13.9)	66 (14.7)	35 (8.3)	–
Monotherapy bDMARDs	0 (0.0)	15 (3.3)	37 (8.8)	–
csDMARD and bDMARDs	6 (1.2)	37 (8.2)	38 (9.0)	–

* Values are the number (%) unless indicated otherwise. ANA = antinuclear antibody; bDMARDs = biologic disease-modifying antirheumatic drugs; csDMARDs = conventional synthetic DMARDs; ERA = enthesitis-related arthritis; IQR = interquartile range; JIA = juvenile idiopathic arthritis; NSAID = nonsteroidal antiinflammatory drug; RF = rheumatoid factor.
† Oligoarticular, not yet differentiated as persistent and extended.

Table 2. Baseline characteristics of the 373 participants with a visit at baseline, 8 years, and 18 years of follow-up compared to the 137 participants lost to follow-up*

	Lost to follow-up (n = 137)	Follow-up at 3 time points (n = 373)	OR (95% CI)	*P*
Female sex	91 (66.4)	249 (66.8)	1.0 (0.6–1.5)	0.94
Age at onset, median (IQR) years	7.1 (3.0–11.2)	5.5 (2.3–9.4)	–	0.10†
ANA positivity, no./total no. (%)	21/77 (27.2)	133/365 (36.4)	0.7 (0.4–1.2)	0.13
HLA–B27 positivity, no./total no. (%)	25/118 (21.2)	79/363 (21.8)	1.0 (0.6–1.6)	0.89
Active joint count, median (IQR)	1 (0–3)	1 (0–3)	–	0.57†
JADAS-71, median (IQR)	5.9 (3.4–11.0)	4.5 (1.8–11.0)	–	0.27†
JIA category				
Systemic JIA	5 (3.6)	13 (3.5)	–	–
Oligoarticular	73 (53.3)	202 (54.2)	1.0 (0.6–1.5)	0.86
Polyarticular RF negative	26 (19.0)	82 (22.0)	0.8 (0.5–1.4)	0.46
Polyarticular RF positive	7 (5.1)	3 (0.0)	–	–
Psoriatic arthritis	3 (2.2)	6 (0.0)	–	–
ERA	10 (7.3)	28 (7.5)	1.0 (0.4–2.1)	0.94
Undifferentiated JIA	13 (9.5)	39 (10.5)	0.9 (0.4–1.8)	0.75

* Values are the number (%) unless indicated otherwise. 95% CI = 95% confidence interval; ANA = antinuclear antibody; ERA = enthesitis-related arthritis; IQR = interquartile range; JADAS-71 = Juvenile Arthritis Disease Activity Score in 71 joints; JIA = juvenile idiopathic arthritis; OR = odds ratio; RF = rheumatoid factor.
† By Mann-Whitney U test.

(Figure 1). In total, 291 of 373 participants (78.0%) attended a clinical visit at all 3 time points.

The demographic data of the cohort are shown in Table 1. Comparison of the participants to those lost to follow-up revealed no significant difference in baseline characteristics (Table 2).

Changes in therapeutic drug use. Of the 373 participants who had a follow-up visit at all 3 time points, 273 (73.2%) were not receiving DMARDs at baseline. At the 8-year follow-up, 59.7% of the participants were not receiving DMARDs, and at the 18-year follow-up this proportion increased significantly to 70.0% (RR 1.3, $P = 0.003$) (data not shown).

Of the 103 participants treated with csDMARDs (either as monotherapy or in combination with bDMARDs) at the 8-year follow-up, 44 (42.7%) were still taking csDMARDs at the 18-year follow-up (RR 0.4, $P < 0.001$) (data not shown). Additionally, of the 52 participants who were treated with bDMARDs (either as monotherapy or combined with csDMARDs) at the 8-year follow-up, 32 (61.5%) were still receiving bDMARDs at the 18-year follow-up (RR 0.6, $P = 0.02$). Only a few participants were taking NSAIDs as monotherapy (3 patients at the 8-year follow-up and 2 patients at the 18-year follow-up).

Among the 373 participants, 55 (14.7%) did not receive any treatment for JIA during the entire period from the 8-year follow-up to the 18-year follow-up. Conversely, 85 of the 373 participants (22.8%) were receiving continuous DMARD treatment due to either uveitis or arthritis during the same period. In total, 33 of the 85 participants (38.8%) were diagnosed with uveitis at some point of the disease course.

At some point during the disease course, 76.9% of the participants with systemic JIA were treated with DMARDs. For the other arthritis categories, 27.6% of patients with oligoarticular persistent, 77.3% with oligoarticular extended, 89.7% with

Table 3. Changes in disease status from 8-year to 18-year follow-up in participants with juvenile idiopathic arthritis (n = 363)*

		18-year follow-up†		
	8-year follow-up	Remission while off medication	Remission while on medication	Not in remission
Remission off medication	151 (42)	103 (68)	4 (3)	44 (29)
Remission on medication	34 (9)	12 (35)	6 (18)	16 (47)
Not in remission‡	178 (49)	40 (22)	23 (13)	115 (65)
Total	363 (100)	155 (43)	33 (9)	175 (48)

* Values are the number (%). Changes in disease status determined according to the preliminary criteria described by Wallace et al (ref. 21).
† Missing data (n = 10).
‡ Active disease or inactive disease not yet fulfilling the remission criteria either on or off medication.

polyarticular RF negative, 66.7% with polyarticular RF positive, 65.2% with juvenile PsA, 75.7% with ERA, 58.9% with undifferentiated, and 62.7% of the total cohort were treated with DMARDs.

In total, 287 of 373 participants (76.9%) had joint injections performed in a median of 3 joints (interquartile range [IQR] 0–8.5) at some time point between baseline and the 8-year follow-up visit. Between the 8- and the 18-year follow-up visits, 146 of 373 participants (39.1%) had joint injections performed, also with a median of 3 joints (IQR 2–7).

Among those who did not receive any medication irrespective of the JIA categories, we grouped together the participants off medication at the 18-year follow-up in the oligoarticular (≤4 cumulative joints) and polyarticular courses (>4 cumulative joints). Among participants with an oligoarticular course, 123 of 136 (90.4%) were off medication at the 8-year follow-up, and 132 of 237 (55.7%) with a polyarticular course were off medication. At the 18-year follow-up, 120 of 136 participants (88.2%) were off medication in the oligoarticular group, and 148 of 237 (62.4%) with a polyarticular course were off medication. Of the 139 participants who never received any DMARDs during their disease course, 85.5% received ≥1 intraarticular glucocorticoid injections, and the remaining participants received NSAIDs alone.

Autologous bone marrow transplantation was performed in 1 participant with recalcitrant systemic JIA. Allogeneic transplantation was completed in 1 participant with ERA due to aplastic anemia during a period of clinical remission of JIA.

Changes in disease status. The median number of active and cumulative joints at baseline was 3 (range 1–30). At the 8-year follow-up visit, the median number of active joints was 0 (range 0–13), and the median number of cumulative joints affected was 6 (range 1–41). Similarly, the median numbers of active and cumulative joints at the 18-year follow-up were 0 (range 0–5) and 7 (range 1–47), respectively.

Of the 151 participants whose disease was in remission off medication 8 years after disease onset, disease in 32% of participants did not remain in remission (Table 3). At the 18-year follow-up, the median duration of remission off medication was 11.5 years (IQR 6.3–15.5 years) (data not shown).

Overall, data regarding remission status for at least 1 of the follow-up visits were missing in 147 participants. To estimate the likely impact of the missing data in these participants on the results, we made an assumption that all missing data were for patients who were either in "remission off medication" or "not in remission" at the 8- and 18-year follow-up visits. If disease in all participants lost to follow-up was in remission off medication, disease in 298 of 510 participants (58% [95% CI 54–63%]) at the 8-year follow-up visit and in 302 of 510 participants (59% [95% CI 55–63%]) at the 18-year follow-up visit would be in remission off medication. Conversely, if disease in all participants lost to follow-up was not in remission, the corresponding numbers would be 30% (95% CI 26–34%) and 30% (95% CI 27–35%) at the 8-year and 18-year follow-up visits, respectively. Among the 34 participants whose disease was in remission on medication 8 years after onset, disease in the majority of participants (18 of 34 [53%]) was in remission on or off medication at the 18-year follow-up (RR 0.5, $P < 0.02$).

Changes in ILAR category. The distribution of patients among the ILAR categories continued to change throughout the study period of 18 years (Figure 2). Altogether, 289 of 422 participants (68.5%) were categorized into the same ILAR category throughout the study period. From baseline to the 8-year follow-up, 30 patients (7%) changed ILAR categories; furthermore,

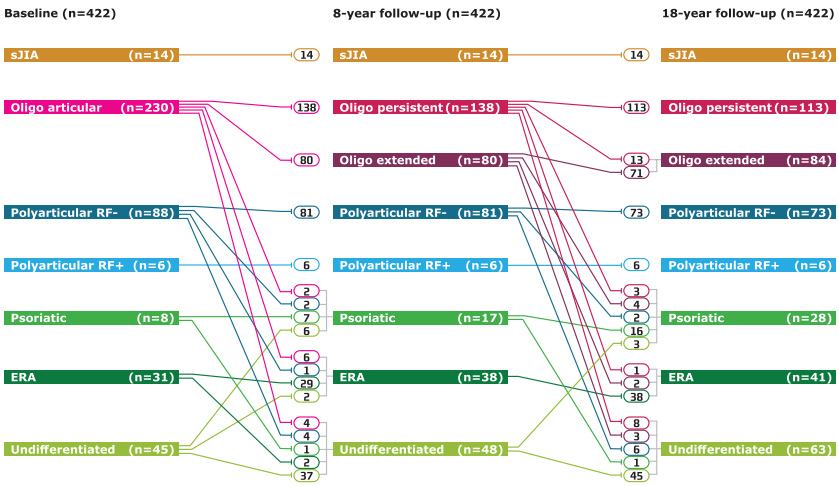

Figure 2. Changes in International League of Associations for Rheumatology category at 3 time points of follow-up in 422 participants with juvenile idiopathic arthritis (JIA). sJIA = systemic JIA; RF- = rheumatoid factor negative JIA; RF+ = rheumatoid factor positive JIA; ERA = enthesitis-related arthritis.

Table 4. Changes in ILAR category at baseline, 8-year, and 18-year follow-up in participants with JIA (n = 422)*

ILAR category	Baseline	8-year follow-up	18-year follow-up	Risk ratio (95% CI)	*P*
Systemic	14	14	14		
Oligoarticular	230	218	197	0.9 (0.7–1.0)	0.02†
Persistent		138	113		
Extended		80	84		
Polyarticular RF negative	88	81	73	0.8 (0.6–1.1)	0.19
Polyarticular RF positive	6	6	6		
Psoriatic	8	17	28	3.5 (1.6–7.6)	<0.00†
Enthesitis-related	31	38	41	1.3 (0.8–2.1)	0.22
Undifferentiated	45	48	63	1.4 (1.0–2.0)	0.06

* Values are the number (%) of patients. Risk ratios and *P* values are for the International League of Associations for Rheumatology (ILAR) category at baseline compared to 18-year follow-up. 95% CI = 95% confidence interval; JIA = juvenile idiopathic arthritis; RF = rheumatoid factor.
† Statistically significant difference between the prevalence at baseline compared to 18-year follow-up.

46 patients (11%) changed categories between the 8- and 18-year follow-up visits. During the study period, there was a significant decrease in the combined persistent and extended oligoarticular categories and a significant increase in the psoriatic group (Table 4). Of the 230 participants with oligoarticular disease at baseline, 84 (36.5%) developed an extended course at the 18-year follow-up (Table 4).

Approximately half of the shifts between ILAR categories (in 30 of 63 participants [47.6%]) were due to updated information on ankylosing spondylitis, psoriasis, or acute anterior uveitis in a first-degree relative. In the psoriatic group, 22 participants were added from other ILAR categories during the disease course, of whom 18 (81.8%) developed psoriasis. In 4 cases, the reasons for shifts between ILAR categories were a first-degree relative with psoriasis combined with 1 additional finding, including nail pitting, onycholysis, or dactylitis. Two patients changed from the psoriatic category because of new information of a first-degree relative with ankylosing spondylitis. Among the 29 participants who changed to the undifferentiated group during the disease course, 25 (86.2%) were added because of new information on a first-degree relative with psoriasis or ankylosing spondylitis, 1 participant developed RF positivity (2 tests greater than 3 months apart), 2 participants fulfilled both ERA and juvenile PsA categories, and 1 male participant who later developed psoriasis was excluded from the juvenile PsA group because he was HLA–B27 positive (with disease onset after his 6th birthday).

The ERA category also increased during 18-year follow-up. An additional 12 participants fulfilled the criteria for this category during disease course due to enthesitis (5), sacroiliitis (4), a first-degree relative with ankylosing spondylitis (2) or acute anterior uveitis (1), and 2 participants left the category (1 participant fulfilled 2 categories and 1 HLA–B27–positive male participant, with disease onset at 7.2 years, later developed psoriasis).

DISCUSSION

In this 18-year prospective study in a population-based setting, we found no further improvement of the remission rate from 8 years of disease and frequent change of ILAR categories during the disease course. Significantly more patients were off medication after 18 years of follow-up (70%) compared to after 8 years of follow-up (60%), but the number of patients whose disease was in remission off medication did not increase (52% versus 51%). The distribution of patients among the ILAR categories, as defined after 8 years, was not sustained and changed significantly by 11% during the follow-up period from 8 to 18 years. The oligoarticular arthritis category decreased significantly, the psoriatic arthritis group increased significantly, and there was an increasing trend for the undifferentiated group. In almost half of the cases (47.6%), the change was due to heredity of ankylosing spondylitis, psoriasis, or acute anterior uveitis in a first-degree relative.

Few studies have investigated longitudinal drug use in JIA with a follow-up of >15 years. New treatment options have changed dramatically since the introduction of biologics and, not surprisingly, we found an increase in the use of biologics during the observation period of 18 years. Furthermore, we found an increased RR of being off any systemic treatment 18 years after disease onset of 1.3 in participants who were off medication at the 8-year follow-up. Almost 60% of participants were off all medication after 8 years, which increased to 70% after 18 years. For comparison, in a single-center study from Norway, Selvaag et al (18) reported that 56% of their cohort (n = 176) used no systemic treatment at the 15-year follow-up, which increased to 87% after 30 years of disease duration. In contrast, a Swedish population-based cohort study by Bertilssson et al (19) found that the 85% of patients who were off csDMARDs after 5 years (n = 129 patients) decreased to 75% at the 17-year follow-up (n = 86). Even though these 2 cohorts were collected almost 2 decades before our cohort (1980–1985), and the treatment strategies have

changed noticeably since then, the rates of patients off systemic treatment were comparable to what we found in the present study (70% of participants after 18 years). This similarity in rates might indicate that the drugs used improve the sequelae of the disease but not the disease course by its very nature. A recent retrospective, 6-year follow-up study on 247 patients from 2 Canadian centers demonstrated that 47% were in remission off medication and 25% were in remission on medication at the last follow-up (22), which is in line with our findings.

Despite the increased chance of being off medication at 8 and 18 years of disease, the remission rate did not increase similarly. Given the counterintuitive finding that the remission rate was stable, but the withdrawal of medication was increased, one can speculate that some participants neglect mild disease activity by not taking the medication and thus may not want affiliation with an outpatient clinic and regular follow-up visits (17). In 2 cases, the disease activity was unknown uveitis activity found due to participation in the study. We have previously described that 19% of the participants in this cohort had >15 minutes of morning stiffness as the only sign of disease activity.

In 68.5% of cases, there was stability throughout the disease course from 8 to 18 years after disease onset. Of the 151 patients whose disease was in remission off medication after 8 years of disease, disease in 69% was still in remission after 18 years. Similarly, the study by Selvaag et al demonstrated that, altogether, 70% of patients had a stable disease course between 15 and 30 years of follow-up (18), and of those whose disease was in remission while off medication after 15 years, 87% remained in that category. In contrast, Bertilsson et al (19) described a stable remission course in 61% of patients between years 5 to 17 of the disease; however, the authors applied their own definition of remission as no evidence of active synovitis and/or active extraarticular features and without drugs for ≥2 years, hampering the means of comparison. Our results suggest that, despite some continued individual shifts between active disease and remission in >30% of participants, the overall disease status of the majority of the cohort remained unchanged between 8 and 18 years. This means that if an individual's disease is in remission off medication 8 years after disease onset, this status is likely to persist for the following 10 years. Likewise, if an individual's disease is not in remission after 8 years, the disease is more prone to remain so.

The use of the ILAR criteria for JIA has been an object of fierce criticism over time (2,23–27), and although the revisions have addressed some weaknesses, several challenges endure (2,28). Beyond many concerns about simplification and lack of validation in large cohorts, there remain inclusion criteria that are not assessed in clinical practice and exceedingly restrictive exclusion criteria that appear to be too rigid for some categories. Furthermore, inclusion/exclusion criteria may induce changes in the distribution of patients among the ILAR categories over time due to supplementary information regarding heredity or development of newly onset rheumatic diseases among first-degree relatives (4). Evidently, it can be difficult to obtain a reliable history of HLA–B27–associated disease or psoriasis among relatives (28), and recorded family history may not be applied repeatedly and strictly enough to keep it continuously updated, leading to an inappropriate, unadjusted classification.

Altogether, the aforementioned concerns regarding criteria strain the credibility of the ILAR classification over time. We found that information about heredity continued to change over time in our cohort; hence, the exclusion criteria accounted for almost half of the changes in ILAR categories. In our study, 32.8% of participants changed ILAR category during the observation period and, in comparison, Bertilsson et al (19) observed that 44% of individuals changed categories during their study's follow-up of 17 years. Nevertheless, they used the European Alliance of Associations for Rheumatology criteria for categorizing juvenile chronic arthritis, which hampers the means of comparability.

The psoriatic group increased significantly over time, but 2 participants left the category because of ankylosing spondylitis diagnosed in a first-degree relative. Although not significantly, the group of undifferentiated JIA increased over time mainly because of development of psoriasis or ankylosing spondylitis in a first-degree relative, which is in line with other studies (29–31). Originally, the undifferentiated category was merely intended to be a temporary group that would conceivably decrease over time (24,26,28); however, this decrease was not confirmed in our study, and this category, by virtue of its lack of homogeneity, is most likely to be unsuitable for research. Our findings support the idea to withdraw, or at least modify, the exclusion criteria in the new data-driven classification criteria proposed by the Pediatric Rheumatology International Trials Organization (26), to ensure more pristine and homogeneous entities.

Several strengths of this study must be emphasized. To our knowledge, this is the only long-term, multicenter, population-based cohort study to include ongoing disease activity with a follow-up of >15 years using the ILAR classification of JIA. The proportion of participants who were lost to follow-up is acceptable, with reliable data on 83.7% of the cohort. The use of validated definitions of disease activity and ILAR classification facilitate comparison with other studies.

A limitation of the present study is that the inclusion period was at the very beginning of the introduction of biologic medicines, which might impede comparison with outcome studies of today. The small sample size in some categories limit the conclusions to be drawn. Additionally, 137 participants did not attend ≥1 follow-up visits, which might bias the results. The distribution of the ILAR categories among the missing and the included participants were similar except for the polyarticular RF+ category; however, caution must be applied when interpreting the changes in ILAR category over time due to missing data. Regarding remission off medication, we calculated the sensitivity of the worst and the best scenario. One could speculate that the participants with a disease no longer playing a prominent part of their daily lives

are more prone not to attend a follow-up visit, which may skew the outcomes in a more severe direction.

In summary, in this population-based setting significantly more patients were off medication at the 18-year follow-up compared to 8 years after disease onset; however, the number of patients whose disease was in remission off medication had not increased. The ILAR categories as defined at baseline were not sustained but changed significantly even beyond 8 years after disease onset.

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. Glerup 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. Glerup, Arnstad, Rypdal, Peltoniemi, Aalto, Rygg, Nielsen, Fasth, Berntson, Nordal, Herlin.

Acquisition of data. Glerup, Arnstad, Rypdal.

Analysis and interpretation of data. Glerup, Rygg, Fasth, Berntson, Nordal, Herlin.

REFERENCES

1. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
2. Fink CW. Proposal for the development of classification criteria for idiopathic arthritides of childhood. J Rheumatol 1995;22:1566–9.
3. Petty RE, Southwood TR, Baum J, Bhettay E, Glass DN, Manners P, et al. Revision of the proposed classification criteria for juvenile idiopathic arthritis: Durban, 1997. J Rheumatol 1998;25:1991–4.
4. Nordal E, Zak M, Aalto K, Berntson L, Fasth A, Herlin T, et al. Ongoing disease activity and changing categories in a long-term nordic cohort study of juvenile idiopathic arthritis. Arthritis Rheum 2011;63:2809–18.
5. Ringold S, Angeles-Han ST, Beukelman T, Lovell D, Cuello CA, Becker ML, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the treatment of juvenile idiopathic arthritis: therapeutic approaches for non-systemic polyarthritis, sacroiliitis, and enthesitis. Arthritis Care Res (Hoboken) 2019;71:717–34.
6. Beukelman T, Patkar NM, Saag KG, Tolleson-Rinehart S, Cron RQ, DeWitt EM, et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. Arthritis Care Res (Hoboken) 2011;63:465–82.
7. Ringold S, Weiss PF, Beukelman T, Dewitt EM, Ilowite NT, Kimura Y, et al. 2013 update of the 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: recommendations for the medical therapy of children with systemic juvenile idiopathic arthritis and tuberculosis screening among children receiving biologic medications. Arthritis Care Res (Hoboken) 2013;65:1551–63.
8. Ringold S, Weiss PF, Colbert RA, DeWitt EM, Lee T, Onel K, et al. Childhood Arthritis and Rheumatology Research Alliance consensus treatment plans for new-onset polyarticular juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2014;66:1063–72.
9. Rezaei E, Hogan D, Trost B, Kusalik AJ, Boire G, Cabral DA, et al. Associations of clinical and inflammatory biomarker clusters with juvenile idiopathic arthritis categories. Rheumatology (Oxford) 2020;59:1066–75.
10. Ravelli A, Consolaro A, Horneff G, Laxer RM, Lovell DJ, Wulffraat NM, et al. Treating juvenile idiopathic arthritis to target: recommendations of an international task force. Ann Rheum Dis 2018;77:819–28.
11. Consolaro A, Giancane G, Alongi A, van Dijkhuizen EH, Aggarwal A, Al-Mayouf SM, et al. Phenotypic variability and disparities in treatment and outcomes of childhood arthritis throughout the world: an observational cohort study. Lancet Child Adolesc Health 2019;3:255–63.
12. Giancane G, Muratore V, Marzetti V, Quilis N, Benavente BS, Bagnasco F, et al. Disease activity and damage in juvenile idiopathic arthritis: methotrexate era versus biologic era. Arthritis Res Ther 2019;21:168.
13. Lovell DJ, Giannini EH, Reiff A, Cawkwell GD, Silverman ED, Nocton JJ, et al. Etanercept in children with polyarticular juvenile rheumatoid arthritis. Pediatric Rheumatology Collaborative Study Group. N Engl J Med 2000;342:763–9.
14. Ruperto N, Lovell DJ, Cuttica R, Wilkinson N, Woo P, Espada G, et al. A randomized, placebo-controlled trial of infliximab plus methotrexate for the treatment of polyarticular-course juvenile rheumatoid arthritis. Arthritis Rheum 2007;56:3096–106.
15. Lovell DJ, Ruperto N, Goodman S, Reiff A, Jung L, Jarosova K, et al. Adalimumab with or without methotrexate in juvenile rheumatoid arthritis. N Engl J Med 2008;359:810–20.
16. Brunner HI, Ruperto N, Tzaribachev N, Horneff G, Chasnyk VG, Panaviene V, et al. Subcutaneous golimumab for children with active polyarticular-course juvenile idiopathic arthritis: results of a multicentre, double-blind, randomised-withdrawal trial. Ann Rheum Dis 2018;77:21–9.
17. Glerup M, Rypdal V, Arnstad ED, Ekelund M, Peltoniemi S, Aalto K, et al. Long-term outcomes in juvenile idiopathic arthritis: eighteen years of follow-up in the population-based Nordic juvenile idiopathic arthritis cohort. Arthritis Care Res (Hoboken) 2020;72:507–16.
18. Selvaag AM, Aulie HA, Lilleby V, Flato B. Disease progression into adulthood and predictors of long-term active disease in juvenile idiopathic arthritis. Ann Rheum Dis 2016;75:190–5.
19. Bertilsson L, Andersson-Gare B, Fasth A, Petersson IF, Forsblad-D'elia H. Disease course, outcome, and predictors of outcome in a population-based juvenile chronic arthritis cohort followed for 17 years. J Rheumatol 2013;40:715–24.
20. Berntson L, Andersson Gare B, Fasth A, Herlin T, Kristinsson J, Lahdenne P, et al. Incidence of juvenile idiopathic arthritis in the Nordic countries. A population based study with special reference to the validity of the ILAR and EULAR criteria. J Rheumatol 2003;30:2275–82.
21. Wallace CA, Ruperto N, Giannini E, Childhood Arthritis and Rheumatology Research Alliance, Pediatric Rheumatology International Trials Organization, Pediatric Rheumatology Collaborative Study Group. Preliminary criteria for clinical remission for select categories of juvenile idiopathic arthritis. J Rheumatol 2004;31:2290–4.
22. Chhabra A, Robinson C, Houghton K, Cabral DA, Morishita K, Tucker LB, et al. Long-term outcomes and disease course of children with juvenile idiopathic arthritis in the ReACCh-Out cohort: a two-centre experience. Rheumatology (Oxford) 2020;59:3727–30.
23. Rumsey DG, Laxer RM. The challenges and opportunities of classifying childhood arthritis. Curr Rheumatol Rep 2020;22:4.
24. Colbert RA. Classification of juvenile spondyloarthritis: enthesitis-related arthritis and beyond. Nat Rev Rheumatol 2010;6:477–85.
25. Duffy CM, Colbert RA, Laxer RM, Schanberg LE, Bowyer SL. Nomenclature and classification in chronic childhood arthritis: time for a change? Arthritis Rheum 2005;52:382–5.
26. Martini A, Ravelli A, Avcin T, Beresford MW, Burgos-Vargas R, Cuttica R, et al. Toward new classification criteria for juvenile

idiopathic arthritis: first steps, Pediatric Rheumatology International Trials Organization International Consensus. J Rheumatol 2019;46: 190–7.

27. Eng SW, Aeschlimann FA, van Veenendaal M, Berard RA, Rosenberg AM, Morris Q, et al. Patterns of joint involvement in juvenile idiopathic arthritis and prediction of disease course: a prospective study with multilayer non-negative matrix factorization. PLoS Med 2019;16:e1002750.
28. Manners P, Lesslie J, Speldewinde D, Tunbridge D. Classification of juvenile idiopathic arthritis: should family history be included in the criteria? J Rheumatol 2003;30:1857–63.
29. Merino R, de Inocencio J, Garcia-Consuegra J. Evaluation of revised International League of Associations for Rheumatology classification criteria for juvenile idiopathic arthritis in Spanish children (Edmonton 2001). J Rheumatol 2005;32:559–61.
30. Chan MO, Petty RE, Guzman J, ReACCh-Out Investigators. A family history of psoriasis in a first-degree relative in children with JIA: to include or exclude? J Rheumatol 2016;43:944–7.
31. Berntson L, Fasth A, Andersson-Gare B, Herlin T, Kristinsson J, Lahdenne P, et al. The influence of heredity for psoriasis on the ILAR classification of juvenile idiopathic arthritis. J Rheumatol 2002;29: 2454–8.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 728–732
DOI 10.1002/acr.24827

BRIEF REPORT

Epidemiology of Childhood-Onset Systemic Lupus Erythematosus: A Population-Based Study

Maria O. Valenzuela-Almada,[1] Mehmet Hocaoglu,[2] Jesse Y. Dabit,[1] Shirley-Ann Osei-Onomah,[1] Matthew L. Basiaga,[1] Amir B. Orandi,[1] Rachel E. Giblon,[1] Kamil E. Barbour,[3] Cynthia S. Crowson,[1] and Alí Duarte-García[1]

Objective. To characterize the incidence and prevalence of childhood-onset systemic lupus erythematosus (SLE), and to estimate the proportion of patients who are diagnosed with SLE during childhood.

Methods. A cohort of patients with incident childhood-onset SLE from 1976 to 2018 from an 8-county region in the US were identified based on comprehensive medical record review. All patients met the European Alliance of Associations for Rheumatology (EULAR)/American College of Rheumatology (ACR) classification criteria for SLE or the ACR SLE classification criteria from 1997 at or before age 18 years. Incidence rates were estimated using Poisson methods. We estimated the childhood-onset SLE point prevalence for January 1, 2015. Results were sex and age adjusted to the US 2000 population. Among all the SLE patients living in the 8-county region on January 1, 2015, the proportion of patients diagnosed at ≤18 years was estimated.

Results. A total of 13 children were diagnosed with childhood-onset SLE during the study period (using the EULAR/ACR definition; mean age at diagnosis 15.1 years, 85% female, 69% White). Childhood-onset SLE overall adjusted incidence rate was 0.7 (95% confidence interval [95% CI] 0.2–1.1) per 100,000 children. The incidence rate in girls was 1.2 (95% CI 0.5–1.9) per 100,000 children, while in boys it was 0.2 (95% CI 0.0–0.5) per 100,000. The adjusted prevalence of childhood-onset SLE was 1.1 (95% CI 0.0–3.1) per 100,000 children. The proportion of patients with SLE diagnosed as children was 9% (95% CI 6–13%).

Conclusion. In this population-based study, both the incidence and prevalence rates of childhood-onset SLE were ~1 per 100,000 children. One in 10 adults with SLE was diagnosed in childhood. More studies are needed to further characterize the epidemiology of childhood-onset SLE in minorities.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic systemic autoimmune disease that causes significant morbidity and mortality. Although adults are more commonly affected, higher burden of disease and a different disease expression with a more aggressive course have been observed in patients with childhood-onset systemic lupus erythematosus (SLE) (1). In youth, quality of life is particularly affected (2). It is estimated that the 5-year survival rate for patients with childhood-onset SLE is 94–100%; however, the 10-year survival rate drops to 81–92% (3,4). Additionally, childhood-onset SLE has been shown to be a strong predictor of mortality in adults with SLE (1).

Despite the impact that this disease has in children and adolescents, the epidemiology of childhood-onset SLE has not been

The findings and conclusions herein are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention, the National Institutes of Health, the Department of Health and Human Services, or the US government.

The Lupus Midwest Network (LUMEN) project is supported by the CDC and the Department of Health and Human Services (grant U01-DP-006491). The Rochester Epidemiology project was supported by the NIH (National Institute on Aging) grant R01-AG-034676 and National Center for Advancing Translational Sciences grant UL1-TR-002377.

[1]Maria O. Valenzuela-Almada, MD, Jesse Y. Dabit, MD, MS, Shirley-Ann Osei-Onomah, MPH, Matthew L. Basiaga, DO, MSCE, Amir B. Orandi, MD, Rachel E. Giblon, MS, Cynthia S. Crowson, PhD, Alí Duarte-García, MD, MSc: Mayo Clinic, Rochester, Minnesota; [2]Mehmet Hocaoglu, MD: Mayo Clinic, Rochester, Minnesota, and University of Maryland Medical Center, Baltimore; [3]Kamil E. Barbour, PhD, MPH, MS: Centers for Disease Control and Prevention, Atlanta, Georgia.

Drs. Valenzuela-Almada and Hocaoglu contributed equally to this work.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24827&file=acr24827-sup-0001-Disclosureform.pdf.

Address correspondence to Alí Duarte-García, MD, MSc, Mayo Clinic, 200 First Street SW, Rochester, MN 55905. Email: duarte.ali@mayo.edu.

Submitted for publication June 16, 2021; accepted in revised form November 23, 2021.

SIGNIFICANCE & INNOVATIONS

- This is the first population-based study investigating the epidemiology of childhood systemic lupus erythematosus using the European Alliance of Associations for Rheumatology/American College of Rheumatology classification criteria for case ascertainment.
- Incidence and prevalence rates of childhood-onset systemic lupus erythematosus were 0.7 (95% confidence interval [95% CI] 0.2–1.1) and 1.1 (95% CI 0.0–3.1) per 100,000 children, respectively.
- The proportion of SLE patients diagnosed as children (≤18 years) was 9%.

well characterized. Prior studies from the US have reported an incidence of SLE in the range of 0.4–2.2 per 100,000 children (5,6) and a prevalence of 9.73–24 per 100,000 children (6,7). Studies from different parts of the world have reported an incidence of 0.28–0.9 per 100,000 children and a prevalence of 3.3–8.8 per 100,000 children (1).

Similar to other pediatric diseases, there is no consensus on the cutoff age to differentiate childhood SLE from adult SLE. Moreover, all previous studies have been either clinic- or reference–center-based (which can suffer from referral [Berkson's] bias, and thus miss milder cases), used International Classification of Diseases (ICD) codes for case definition (which may introduce misclassification bias), and estimated period prevalence rather than point prevalence (which may result in overestimation) (8–10). Additionally, it is commonly reported across the literature that the proportion of patients with SLE diagnosed in childhood is between 10% and 20%; however, these data are not backed by any published epidemiologic study (1).

Given the lack of population-based studies evaluating childhood-onset SLE epidemiology, we aimed to characterize the incidence (from 1976 to 2018) and the prevalence of childhood-onset SLE and to estimate the proportion of patients who are diagnosed with SLE during childhood leveraging the resources of a record-linkage system, which minimizes misclassification and misdiagnosis. In addition, we used the novel and rigorously developed European League Against Rheumatism/American College of Rheumatology (EULAR/ACR) SLE classification criteria (11), which classify more patients in epidemiologic studies compared to older criteria (12).

PATIENTS AND METHODS

Study design. The Lupus Midwest Network (LUMEN) is a population-based registry of a 27-county region in southeast Minnesota and southwest Wisconsin nested in the Rochester Epidemiology Project (REP), a record-linkage system. The REP included patients from Olmsted County, Minnesota since 1966, and in 2010, it was expanded to 26 additional counties. Eight counties (Olmsted, Mower, Freeborn, Waseca, Steele, Dodge, Wabasha, and Goodhue) with a >95% capture of the census population were used for this study. According to US Census data, Olmsted County, the county included in the REP for the longest period of time, had a population of 92,006 in 1980, 29.5% being <18 years of age, and 98% White. In 2010, the population was 144,248, with 25.3% being <18 years of age. The ethnic distribution in 2010 was 85.7% White, 4.2% Hispanic, 4.8% African American, 5.5% Asian/Native Hawaiian/Pacific Islander, and 0.2% American Indian/Alaska Native (13). Data from all patients meeting EULAR/ACR criteria for SLE from 1976 to 2018 living within these 8 counties in Minnesota were collected. The EULAR/ACR criteria were selected as the primary definition due to their better performance classifying patients in epidemiologic studies than the other commonly used SLE classification criteria (12). The ACR criteria from 1997 (ACR 97) were used as a secondary definition for comparison purposes (14). The REP allows for an adequate investigation of the epidemiology of childhood-onset SLE because comprehensive medical records for all residents seeking medical care are available. The REP provides ready access to the medical records from all health care providers (including pediatric care providers, both primary care and subspecialty care) for the local population, including the Mayo Clinic, the Olmsted Medical Center, and their affiliated hospitals, local nursing homes, and a few private practitioners. This system ensures virtually complete ascertainment of all clinically recognized cases of SLE of children living in the 8 counties. The characteristics and strengths of the REP, as well as its generalizability, have been described elsewhere (15).

Case finding and ascertainment. Potential childhood-onset SLE cases were identified through 2 different strategies: 1) Hospital International Classification of Disease Adaptation, ICD-9, and ICD-10 codes for SLE, cutaneous lupus erythematosus, and other associated diseases, and 2) laboratory measures associated with SLE (antinuclear antibodies [≥1:80]), low complement, anti–double-stranded DNA, anti-Sm, lupus anticoagulant anticardiolipin (IgG, IgM, and IgA), and anti–β_2-glycoprotein I (IgG, IgM, and IgA) antibodies (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24827/abstract). Then, individual chart reviews were performed, and data were abstracted by extensively trained reviewers. Data extraction was done using a standardized REDCap data capture tool. Demographic characteristics and clinical and laboratory data included in the classification criteria were abstracted from the electronic medical record and/or paper records. Patients meeting the EULAR/ACR or ACR 97 classification criteria who were ≤18 years of age at the time of diagnosis and living in 1 of the 8 counties prior to the earliest date of criteria fulfillment were considered as having incident SLE. The date of criteria fulfillment was considered the incidence date. We considered prevalent

Table 1. Demographic characteristics of 13 patients with incident childhood systemic lupus erythematosus*

Characteristic	Olmsted county, 1976–2009 (n = 8)	8 counties, 2010–2018 (n = 5)†	Total (n = 13)	*P*‡
Age, mean ± SD years	15.3 ± 2.72	14.9 ± 4.84	15.1 ± 3.49	0.66
Sex, female	7 (88)	4 (80)	11 (85)	1.0
Race/ethnicity§				0.32
Asian	3 (38)	0 (0)	3 (23)	
Black	0 (0)	1 (20)	1 (8)	
White	5 (63)	4 (80)	9 (69)	

* Values are the number (%) unless indicated otherwise.
† Eight counties were included because the Rochester Epidemiology Project captured >95% of their population: Olmsted, Mower, Freeborn, Waseca, Steele, Dodge, Wabasha, Goodhue, and Minnesota.
‡ Differences between proportions among Olmsted County (1976–2009) versus the 8-county region (2010–2018) were evaluated using Fisher's exact test, and mean differences were evaluated using Kruskal-Wallis exact test.
§ There were no Hispanic patients.

SLE patients those who were age ≤18 years on January 1, 2015, and who met the EULAR/ACR classification criteria, or if they migrated to the region after diagnosis (and therefore were under treatment) if they had 7 EULAR/ACR points and a physician diagnosis. For the ACR 97 criteria, a patient was considered as having prevalent SLE if he or she fulfilled at least 3 criteria and had a physician diagnosis. If a clinical manifestation was better explained by something other than SLE, it was not considered for the criteria. A rheumatologist with expertise in SLE (AD) reviewed all the cases that were classified to confirm the attribution of clinical manifestation to SLE.

Statistical analysis. Descriptive statistics (percentages, means, etc.) were used to summarize the demographic characteristics. Fisher's exact test was used to assess differences across categorical variables; mean differences in continuous variables were evaluated using the Kruskal-Wallis exact test. Incidence rates were calculated using the number of incident cases as the numerator and population counts for residents age ≤18 years from the REP census as the denominator (16). Overall incidence rates were age adjusted and/or age and sex adjusted to the US total population in 2000 (17). In order to compute 95% confidence intervals (95% CIs) for incidence rates, it was assumed that the number of incident cases followed a Poisson distribution. The prevalence of childhood-onset SLE on January 1, 2015, was calculated in those patients age ≤18 years living within 8 counties in Minnesota. Among all the SLE patients living in the 8-county region on January 1, 2015, the proportion of patients diagnosed at ≤18 years of age was obtained. Statistical analyses were performed using SAS, version 9.4. This study was approved by institutional review boards of the Mayo Clinic (IRB 20-006485) and Olmsted Medical Centre (IRB 036-OMC-20).

RESULTS

Thirteen patients ≤18 years of age, living in the 8-county region fulfilled the EULAR/ACR criteria for childhood-onset SLE during the study period (11 [85%] girls and 2 [15%] boys. Eleven cases were classified by the ACR 97 criteria, and all of them were

Table 2. Incidence rates for 1976–2018 and point prevalence rates on January 1, 2015, of childhood systemic lupus erythematosus*

	EULAR/ACR criteria			ACR 97 criteria, overall (1976–2018)
	Olmsted county (1976–2009)	8 counties (2010–2018)	Overall (1976–2018)	
Incidence rate per 100,000 children (95% CI)				
Total	0.7 (0.2–1.2)	0.6 (0.1–1.2)	0.7 (0.2–1.1)	0.59 (0.24–0.94)
Female	1.3 (0.3–2.3)	1.0 (0.0–2.1)	1.2 (0.5–1.9)	0.98 (0.34–1.63)
Male	0.2 (0.0–0.5)	0.3 (0.0–0.8)	0.2 (0.0–0.5)	0.21 (0.0–0.51)
Prevalence per 100,000 children (95% CI)†				
Total	–	–	1.1 (0.0–3.1)	1.1 (0.0–3.1)
Female	–	–	2.2 (0.0–6.4)	2.2 (0.0–6.4)
Male	–	–	0.0 (0.0–0.0)	0.0 (0.0–0.0)

* Sex-specific rates are age adjusted, and total rates are age and sex adjusted to the US total 2000 population. 95% CI = 95% confidence interval; ACR = American College of Rheumatology; EULAR = European Alliance of Associations for Rheumatology.
† January 1, 2015.

also classified by the EULAR/ACR criteria. The female to male ratio was 5.5:1. The mean age at diagnosis was 15.1 years, and the majority (69%) were non-Hispanic White; there were no Hispanic children with SLE in this population. Only 1 child was diagnosed before age 10 years (Table 1).

The childhood-onset SLE overall incidence rate was 0.7 (95% CI 0.2–1.1) per 100,000 children. The incidence rate in girls was 1.2 (95% CI 0.5–1.9) per 100,000 children, while in boys it was 0.2 (95% CI 0.0–0.5) per 100,000 children. The prevalence of childhood-onset SLE was 1.1 (95% CI 0.0–3.1) per 100,000 children. When subdivided between the 2 time periods, from 1976 to 2009 (only Olmsted county) and from 2009 to 2018 (8 counties), the incidence rates were similar. The incidence using the ACR 97 definition was lower, with an overall incidence of 0.59 (95% CI 0.24–0.94) (Table 2).

We found only 1 prevalent case of childhood-onset SLE meeting both the EULAR/ACR and ACR 97 criteria on January 1, 2015; the estimated prevalence was 1.1 (95% CI 0.0–3.1) per 100,000 children. A total of 251 patients with SLE lived in 1 of the 8 counties on January 1, 2015, and 23 (9% [95% CI 6–13]) were diagnosed as children (≤18 years).

The most common clinical manifestations at the time of classification fulfillment were arthritis/synovitis, leukopenia (31% each), and thrombocytopenia and nephritis (23% each). None of the patients had neurologic manifestations. All the patients were antinuclear antibody positive; 92% were either positive for anti–double-stranded DNA or anti-Sm antibodies, and 85% had hypocomplementemia (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24827/abstract).

DISCUSSION

In this population-based study, we found that the incidence and prevalence of childhood-onset SLE using the EULAR/ACR criteria were 0.7 and 1.1 per 100,000 children, respectively. The incidence using the ACR 97 definition was slightly lower at 0.6 per 100,000 children, while the prevalence was unchanged. In addition, we estimated that of all patients meeting criteria for SLE in the 8-county population, 9% were diagnosed in childhood. Two prior studies in the US have estimated the incidence of childhood-onset SLE. The first one was based on pediatric referral centers in the New England region of the US; the childhood-onset SLE incidence rate was 0.4 per 100,000 children, which was similar to the estimate from our study (5). While the second study, by Hiraki et al, observed an annual incidence of childhood-onset SLE of 2.2 per 100,000 children using data from Medicaid-enrolled US children 3–18 years of age (6). Although it is difficult to compare the studies because the present study is population based and the others were not, and because the case ascertainment and definitions were done differently, it is possible that the New England study, by including all the pediatric rheumatology clinics, was able to identify most or all of the childhood-onset SLE cases in the population. The Medicaid study, however, had an incidence rate twice as high. The relatively higher incidence rate in that study could be attributed to the unique characteristics of the population that is covered under Medicaid, such as the higher proportion of minorities, in addition to differences in case definition and case ascertainment. While in our study all the cases were verified by thorough chart reviews and access to all the medical records, the Medicaid study used an administrative database, with an ICD code–based rule that has been demonstrated to have positive predictive value of 75% (8), which could explain the higher rates.

We estimated a childhood-onset SLE prevalence of 1.1 per 100,000 children. In comparison, 2 studies in the US have shown higher prevalence estimates. A prevalence of 9.73 per 100,000 was seen in the Medicaid claims–based study in the US, although they reported a lower prevalence of childhood-onset SLE (8.03) in the US Midwest when compared to the rest of the country (6). An even higher prevalence of 24.0 per 100,000 children was reported in a clinic-based study in Hawaii (7). There are notable differences in the way prevalence was estimated in these studies in addition to ascertainment differences and case definitions (ICD code based, etc.). Our study used a point prevalence on January 1, 2015, while the Medicaid study used a 4-year period prevalence, and the Hawaii study used a 10-year period prevalence. Period prevalence has been shown to produce estimates that are many times higher than point prevalence estimates (10). The low prevalence in our study, relative to our incidence rate, is explained by the fact that children grow to adulthood, and thus were not counted in the study population. The mean age of diagnosis in our study was 15 years; therefore, most of the children identified in the study were adults by the time of the point prevalence estimation. More than 40% of the children included in the Medicaid study were age ≥15 years, thus reaching adulthood during the period when prevalence was calculated. Similarly, the Hawaii study prevalence period was 10 years; therefore, many children grew to adulthood in the time included in the prevalence estimate.

Finally, we found that 9% of patients with SLE are diagnosed during childhood. It is commonly reported in the literature that 10–20% of patients with SLE are diagnosed in childhood; however, there are no population-based epidemiologic studies supporting these estimates (1).

Our study does have limitations. First, the small size of our population could decrease the accuracy of our estimates. Second, the retrospective nature of data collection through review of medical records could potentially lead to missing information or misclassification and under-ascertainment of cases, as it depended on the documentation and the workup

completeness of the treating clinicians. Third, the number of undiagnosed patients in the community who have not yet been seen and evaluated appropriately for the disease could lead to underestimation of our estimates. And fourth, because our population is mostly non-Hispanic White, our results may not be generalizable to other populations. However, the study also has strengths. We had a >95% capture of the census population of the 8 counties included, minimizing the likelihood of underestimating childhood-onset SLE incidence and prevalence. Having access to all the medical records and laboratory findings from multiple health care facilities through the REP and performing a thorough review and abstraction allowed us to decrease misclassification of patients and missing data. Additionally, in having identified all the clinically detected cases of childhood-onset SLE in the community, we minimized referral bias.

In conclusion, in this population-based study, the incidence and prevalence rates of childhood-onset SLE and the proportion of patients diagnosed with SLE during childhood were lower than has been reported previously in the US. Moreover, this information could potentially aid in the appropriate allocation of resources for early diagnosis and management of the disease. Nevertheless, given the differences of childhood-onset SLE across racial and ethnic groups, it is important to conduct further population-based studies in diverse populations.

ACKNOWLEDGMENT

The study team thanks Barbara Abbott for her support in accessing the Rochester Epidemiology Project data.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Duarte-García 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. Crowson, Duarte-García.

Acquisition of data. Valenzuela-Almada, Hocaoglu, Dabit, Osei-Onomah, Crowson, Duarte-García.

Analysis and interpretation of data. Valenzuela-Almada, Hocaoglu, Basiaga, Orandi, Giblon, Barbour, Crowson, Duarte-García.

REFERENCES

1. Kamphuis S, Silverman ED. Prevalence and burden of pediatric-onset systemic lupus erythematosus. Nat Rev Rheumatol 2010;6:538–46.
2. Falcini F, Nacci F. Systemic lupus erythematosus in the young: the importance of a transition clinic. Lupus 2007;16:613–7.
3. Ravelli A, Ruperto N, Martini A. Outcome in juvenile onset systemic lupus erythematosus. Curr Opin Rheumatol 2005;17:568–73.
4. Appenzeller S, Marini R, Costallat LT. Damage did not independently influence mortality in childhood systemic lupus erythematosus. Rheumatol Int 2005;25:619–24.
5. Denardo BA, Tucker LB, Miller LC, Szer IS, Schaller JG. Demography of a regional pediatric rheumatology patient population. Affiliated Children's Arthritis Centers of New England. J Rheumatol 1994;21:1553–61.
6. Hiraki LT, Feldman CH, Liu J, Alarcón GS, Fischer MA, Winkelmayer WC, et al. Prevalence, incidence, and demographics of systemic lupus erythematosus and lupus nephritis from 2000 to 2004 among children in the US Medicaid beneficiary population. Arthritis Rheum 2012;64:2669–76.
7. Kurahara DK, Grandinetti A, Fujii LL, Tokuda AA, Galario JA, Han MJ, et al. Visiting consultant clinics to study prevalence rates of juvenile rheumatoid arthritis and childhood systemic lupus erythematosus across dispersed geographic areas. J Rheumatol 2007;34:425–9.
8. Barnado A, Casey C, Carroll RJ, Wheless L, Denny JC, Crofford LJ. Developing electronic health record algorithms that accurately identify patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69:687–93.
9. Westreich D. Berkson's bias, selection bias, and missing data. Epidemiology 2012;23:159–64.
10. Spronk I, Korevaar JC, Poos R, Davids R, Hilderink H, Schellevis FG, et al. Calculating incidence rates and prevalence proportions: not as simple as it seems. BMC Public Health 2019;19:512.
11. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71:1400–12.
12. Duarte-García A, Hocaoglu M, Osei-Onomah SA, Dabit JY, Giblon RE, Helmick CG, et al. Population-based incidence and time to classification of systemic lupus erythematosus by three different classification criteria. Rheumatology (Oxford) 2021. doi: 10.1093/rheumatology/keab807. E-pub ahead of print.
13. Rochester Epidemiology Project. REP population overview. URL: https://rochesterproject.org/for-researchers/population-overview/.
14. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.
15. Rocca WA, Grossardt BR, Brue SM, Bock-Goodner CM, Chamberlain AM, Wilson PM, et al. Data resource profile: expansion of the Rochester Epidemiology Project medical records-linkage system (E-REP). Int J Epidemiol 2018;47:368–j.
16. St Sauver JL, Grossardt BR, Yawn BP, Melton LJ III, Rocca WA. Use of a medical records linkage system to enumerate a dynamic population over time: the Rochester epidemiology project. Am J Epidemiol 2011;173:1059–68.
17. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People 2010 Stat Notes 2001:1–10.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 733–740
DOI 10.1002/acr.24837

American College of Rheumatology
Empowering Rheumatology Professionals

Medication Interruptions and Subsequent Disease Flares During the COVID-19 Pandemic: A Longitudinal Online Study of Patients With Rheumatic Disease

Tiffany Dharia,[1] Shilpa Venkatachalam,[2] Joshua F. Baker,[1] Shubhasree Banerjee,[1] David Curtis,[2] Maria I. Danila,[3] Kelly Gavigan,[2] Jennifer Gordon,[4] Peter A. Merkel,[1] Dianne G. Shaw,[5] Kalen Young,[5] Jeffrey R. Curtis,[3] William B. Nowell,[2] and Michael D. George[1]

Objective. We aimed to assess trends in anxiety and interruptions in disease-modifying antirheumatic drug (DMARD) use among patients with rheumatic diseases during the COVID-19 pandemic and to evaluate whether DMARD interruptions were associated with disease flares.

Methods. ArthritisPower, the Vasculitis Patient-Powered Research Network, and other patient organizations invited members to join a 52-week longitudinal study, with baseline surveys completed March 29 to June 30, 2020, with follow-up through May 2021. Logistic regression incorporating generalized estimating equations evaluated associations between interruptions in DMARD use and self-reported disease flares at the next survey, adjusting for demographic characteristics, medications, disease, and calendar time.

Results. Among 2,424 patients completing a median of 5 follow-up surveys, the mean age was 57 years, 87% were female, and the most common conditions were rheumatoid arthritis, vasculitis, and psoriatic arthritis. Average Patient-Reported Outcomes Measurement Information System (PROMIS) anxiety T scores decreased from April 2020 (58.7) to May 2021 (53.7) ($P < 0.001$ for trend). Interruptions in DMARD use decreased from April (11.2%) to December 2020 (7.5%) ($P < 0.001$) but increased through May 2021 (14.0%) ($P < 0.001$). Interruptions in DMARD use were associated with a significant increase in severe flares (rated ≥6 of 10) at the next survey (12.9% versus 8.0% [odds ratio (OR) 1.71 (95% confidence interval [95% CI 1.23, 2.36]) although not any flare (OR 1.18 [95% CI 0.89, 1.58])].

Conclusion. Anxiety and interruptions in DMARD use initially decreased over time, but DMARD interruptions increased during 2021, possibly related to an increase in COVID-19 cases or vaccine availability. Interruptions in DMARD use were associated with increased rates of severe disease flares, highlighting the importance of avoiding unnecessary DMARD interruptions.

INTRODUCTION

SARS–CoV-2, a novel coronavirus, is a highly pathogenic virus that causes COVID-19 and rapidly led to a global pandemic (1). The COVID-19 pandemic has been a particular concern for patients with autoimmune rheumatic diseases (ARDs), who are known to be at a higher risk for infections due to their autoimmune conditions, comorbidities, and use of immunosuppressive therapies (2–4). Despite ongoing research, the risk of severe COVID-19 due to use of different immunosuppressive therapies

The Arthritis and Rheumatic Disease COVID-19 Project has received sponsorship support specifically from the Patient-Centered Outcomes Institute (PCORI), Eli Lilly and Company, and Janssen Pharmaceutical. The Vasculitis Patient-Powered Research Network was supported through a PCORI award (grant PPRN-1306-04758), through the Vasculitis Foundation, and through the Vasculitis Clinical Research Consortium (VCRC).

Supported by the VCRC, which is part of the Rare Diseases Clinical Research Network, an initiative of the Office of Rare Diseases Research, National Center for Advancing Translational Science (NCATS). The VCRC is funded through collaboration between NCATS and the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant U54-AR-057319). Dr. George's work was supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant K23-AR-073931).

[1]Tiffany Dharia, MD, Joshua F. Baker, MD, MSCE, Shubhasree Banerjee, MD, Peter A. Merkel, MD, MPH, Michael D. George, MD, MSCE: University of Pennsylvania, Philadelphia; [2]Shilpa Venkatachalam, PhD, MPH, David Curtis, BA, Kelly Gavigan, MPH, William B. Nowell, PhD: Global Healthy Living Foundation, Upper Nyack, New York; [3]Maria I. Danila, MD, MSc, MSPH, Jeffrey R. Curtis, MD, MPH: University of Alabama at Birmingham; [4]Jennifer Gordon, PhD: Temple University, Philadelphia, Pennsylvania; [5]Dianne G. Shaw, MA, Kalen Young, MA: Vasculitis Foundation, Kansas City, Missouri.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24837&file=acr24837-sup-0001-Disclosureform.pdf.

Address correspondence to Michael D. George, MD, MSCE, Division of Rheumatology, 5 White Building, 3400 Spruce Street, Philadelphia, PA 19104. Email: michael.george@pennmedicine.upenn.edu.

Submitted for publication September 7, 2021; accepted in revised form December 9, 2021.

SIGNIFICANCE & INNOVATIONS

- Among patients with autoimmune rheumatic disease, concerns regarding interruptions in disease-modifying antirheumatic drug (DMARD) use because of COVID-19 declined by >30%, and anxiety scores improved by December 2020.
- Between December 2020 and May 2021, interruptions in DMARD use increased by >80% despite continued improvements in anxiety, perhaps related to concerns about vaccine efficacy.
- Patients who reported interruptions in DMARD use had higher rates of self-reported severe disease flares on subsequent surveys, demonstrating the importance of avoiding unnecessary medication interruptions.

remains uncertain (5–9). The impact of immunosuppressive therapies on vaccination response has also emerged as a concern for patients with ARDs (10). There is an urgent need to determine how patient care has been affected by the pandemic to better understand barriers to effective care during this and future public health crises.

Prior studies have shown that patients with ARDs had frequent health care disruptions and interruptions in the use of their disease-modifying antirheumatic drugs (DMARDs) early in the COVID-19 pandemic (11,12). Little is known, however, about how these disruptions affected patient health or how patient concerns and behaviors changed over time. The goals of this study were to use longitudinal data contributed by an online sample of patients to examine trends in anxiety, as a proxy for mental health, and interruptions in DMARD use throughout the pandemic and to evaluate whether interruptions in DMARD use were associated with disease flares.

PATIENTS AND METHODS

Study population. Adults >18 years of age in the ArthritisPower and vasculitis patient-powered research networks (PPRNs), the CreakyJoints patient community, and partnering patient organizations were sent email invitations. The ArthritisPower PPRN (13,14) is a patient-led online registry of patients with inflammatory arthritis and other rheumatic conditions created as a joint venture of the patients and patient advocates of the Global Healthy Living Foundation, the CreakyJoints patient community, and researchers at the University of Alabama at Birmingham. The Vasculitis PPRN, a collaboration between the Vasculitis Clinical Research Consortium and the Vasculitis Foundation, is an online research registry utilizing patient-reported data. Participants of partnering patient organizations (the Vasculitis Foundation, the Relapsing Polychondritis Foundation, American Bone Health, the Lupus and Allied Diseases Association, Myositis Support and Understanding Association, and the International Foundation for Autoimmune and Autoinflammatory Arthritis) were directed to a landing page (https://autoimmunecovid.org). Patients were invited to complete a survey at baseline (week 0), weeks 2, 4, 6, and 8, monthly until week 28, and then at weeks 38 and 52 (for a maximum of 11 follow-up surveys) (see Covid registry surveys in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24837/abstract). We analyzed patients with ARDs who completed baseline surveys between March 29 and June 30, 2020, and who completed at least 1 follow-up survey, with follow-up captured through May 2021.

Data collection. At baseline, patients were asked to report demographic data, including country, city, and zip code, as well as rheumatologic conditions, comorbidities, and medications used to treat their ARD. For participants indicating multiple rheumatologic conditions, a hierarchical approach was taken. Patients from the ArthritisPower PPRN and CreakyJoints were classified as having systemic lupus erythematosus (SLE) > psoriatic arthritis (PsA) > ankylosing spondylitis (AS) > rheumatoid arthritis (RA) > vasculitis > myositis, similarly to prior studies (15,16). Patients in the Vasculitis PPRN were thought to be more likely to have vasculitis as a primary diagnosis and so were preferentially categorized as having vasculitis. Patients were considered to have antineutrophil cytoplasmic antibody–associated vasculitis if they reported diagnoses of eosinophilic granulomatosis with polyangiitis, granulomatosis with polyangiitis, or microscopic polyangiitis or were otherwise characterized as having "other vasculitis."

County rural versus urban status was defined using National Center for Health Statistics classification (17). Zip code–based median household income and education from the 5-year estimates of the American Community Survey 2014–2018 were also divided into tertiles (18).

At baseline and at each follow-up survey, patients reported respiratory illnesses within the prior 2 weeks, any COVID-19 testing/diagnosis since the prior survey, use and availability of telemedicine, current DMARD use, interruptions in DMARD use because of COVID-19 concerns ("Have you stopped or temporarily paused any of your medications for your rheumatic/autoimmune disease because of concerns about coronavirus/COVID-19?"), and anxiety in the prior week using the 4-question Patient-Reported Outcomes Measurement Information System (PROMIS) anxiety short form. PROMIS anxiety responses were converted to T scores, where a score of 50 represents the general population mean with a 10-point SD (19). In addition, at each follow-up visit, patients reported whether they were currently experiencing a flare of their autoimmune condition and, if so, how severe the flare had been the prior week on a scale ranging from 0 (no flare) to 10 (extremely bad). As agreement between physicians and patients on the presence of flare is associated with greater patient-reported flare severity (20), we a priori created a definition of "severe flare" defined as a severity ≥6 of 10 based on the distribution of flare severities observed, hypothesizing that

these severe flares would be more likely to represent true inflammatory disease flares.

Statistical analysis. Characteristics of patients who completed at least 1 follow-up survey versus those who completed only the baseline survey were compared descriptively. Among patients with ≥1 follow-up survey, changes over time in PROMIS anxiety T scores were assessed using generalized estimating equation (GEE) models to account for within-person correlations, with PROMIS anxiety scores as the

Table 1. Baseline characteristics of study participants*

Characteristic	Completed ≥1 follow-up survey (n = 2,424)	Baseline survey only (n = 619)
Age, mean ± SD years	56.8 ± 12.0	52.6 ± 12.6
Female	2,098 (86.6)	534 (86.3)
Hispanic	612 (4.0)	35 (5.6)
Race		
White	2,200 (90.8)	531 (85.8)
Black	57 (2.4)	21 (3.4)
Asian	23 (1.0)	11 (1.8)
Other/multiracial	144 (5.9)	56 (9.1)
Autoimmune disease		
Rheumatoid arthritis	1,012 (41.8)	232 (37.5)
ANCA-associated vasculitis	359 (14.8)	82 (13.3)
Psoriatic arthritis	300 (12.4)	89 (14.4)
Ankylosing spondylitis	183 (7.6)	46 (7.4)
Other vasculitis	176 (7.3)	62 (10.0)
Lupus	123 (5.1)	37 (6.0)
Myositis	61 (2.5)	14 (2.3)
Other†	210 (8.7)	57 (9.2)
Patient organization		
ArthritisPower	1,162 (47.9)	363 (58.6)
Vasculitis PPRN	521 (21.5)	142 (22.9)
CreakyJoints	567 (23.4)	90 (14.5)
Partnering patient organizations	174 (7.2)	24 (3.9)
Rural residence	276 (12.6)	78 (13.7)
Region		
South	844 (37.0)	220 (37.2)
West	545 (23.9)	138 (23.3)
Midwest	492 (21.6)	136 (22.97)
Northeast	401 (17.6)	98 (16.6)
Medications		
Biologic/JAK inhibitor	1,274 (52.6)	300 (48.5)
Methotrexate	727 (30.0)	161 (26.0)
Hydroxychloroquine	518 (21.4)	121 (19.6)
Glucocorticoids <10 mg/day	581 (24.0)	155 (25.0)
Glucocorticoids ≥10 mg/day	112 (4.6)	41 (6.6)
Illness reported at baseline		
No respiratory illness	2,160 (89.1)	520 (84.0)
Respiratory illness without COVID-19 diagnosis	239 (9.9)	92 (14.9)
COVID-19 diagnosis	25 (1.0)	7 (1.1)
PROMIS anxiety, T score‡	58.9 (8.6)	60.1 (8.5)
Health-related behaviors (baseline visit)		
Avoided an office visit	1,430 (59.0)	380 (61.4)
Avoided getting laboratory tests	1,015 (41.9)	293 (47.3)
Avoided getting an infusion	322 (13.3)	94 (15.2)
Interrupted use of a DMARD because of COVID-19 concerns, no./total no. (%)§	191/1,748 (10.9)	75/405 (18.5)
Any flare during follow-up	1,414 (58.3)	NA

* Values are the number (%) unless indicated otherwise. ANCA = antineutrophil cytoplasmic antibody; DMARDs = disease-modifying antirheumatic drugs; NA = not applicable; PPRN = patient-powered research network; PROMIS = Patient-Reported Outcomes Measurement Information System.
† "Other" includes patients with other autoimmune conditions (most commonly inflammatory bowel disease, Sjögren's syndrome, or psoriasis) or patients who reported a non-listed autoimmune condition.
‡ From the PROMIS anxiety short form, with a range of 1–100 and a mean ± SD US adult population of 50 ± 10.
§ Interruptions in the use of a DMARD among patients receiving DMARDs who did not report a respiratory illness.

dependent variable, and month (as a categorical variable) as the independent variable, adjusting for baseline PROMIS anxiety scores. Marginal predictions of scores by month were used to create graphical representations of trends. Similar GEE logit models were used to assess trends in interruptions in DMARD use over time (among patients receiving a DMARD who did not report a respiratory illness or COVID-19 diagnosis) and the cumulative use of telemedicine. For these models, DMARD interruption or use of telemedicine (at the current or any prior survey) were the dependent variables, and month (categorical) was the independent variable. Statistical evaluations of trends over time were assessed by modeling month as a continuous variable, using a linear spline with a knot at December 2020 for interruptions in DMARD use because of changes in trends visualized at this time point. Differences in use of telemedicine by region or by rural status over time were assessed in models including month and either region or rural status, with differences in rates over time assessed with models that included region, month (continuous), and month–region interaction terms. Differences in reasons for interruptions in DMARD use between 2020 and 2021 were compared with chi-square or Fisher's exact testing.

To evaluate whether interruptions in DMARD use were associated with disease flares, we examined patients who were receiving a DMARD, not currently in a flare, and not reporting a current respiratory illnesses or diagnosis of COVID-19. Among this population, we examined whether the DMARD interruptions were associated with the frequency of self-reported flares of any severity or severe flares (severity ≥6 of 10) at the next survey using GEE models with a logit link, adjusting for age, sex, race (White versus non-White), medication type, disease type, and calendar time (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24837/abstract). Marginal predictions were used to estimate flare rates in patients who interrupted versus those who had not interrupted DMARDs at the mean of all covariates in the model.

In sensitivity analyses, we evaluated associations between interruptions in DMARD use and flares only among patients with RA, PsA, AS, or SLE, adjusted for season instead of calendar time, and repeated analyses among patients with at least 6 months of follow-up. Analyses were performed using Stata, version 15.1. All patients provided informed consent and participated without compensation. The study protocol was approved by the Advarra institutional review board.

RESULTS

Between April and June 2020, 45,977 patients opened emails with information about the study, 6,065 (13.2%) opened a link to the study, and a total of 3,338 (55.0%) of these patients completed the baseline survey. Among these patients, 295 did not report having an autoimmune disease and were excluded from analysis. Of the remaining 3,043 patients, 2,424 completed at least 1 follow-up survey (a median of 5 follow-up surveys completed [interquartile range (IQR) 2–8]) and are included in this analysis. The mean age was 57 years (range 18–93), 87% were female, and the most common ARDs were RA, vasculitis, and PsA (Table 1). At baseline, 53% of patients were taking biologics or JAK inhibitors, 30% were taking methotrexate, and 29% were taking glucocorticoids.

Longitudinal trends in anxiety, telemedicine use, and interruptions in DMARD use. Average PROMIS anxiety T scores decreased significantly from 58.7 in April 2020 to 53.7 in May 2021 ($P < 0.001$ for trend) (Figure 1). Anxiety was greater in patients who reported avoiding a doctor's office visit on their baseline survey (coefficient 3.2 [95% confidence interval (95% CI) 2.6, 3.9], $P < 0.001$), with no differences among patients who reported versus who did not report use of telemedicine (coefficient –0.2 [95% CI –0.5, 0.1], $P = 0.17$).

On the baseline survey, 42% of patients reported completing a telemedicine visit. Over time, the proportion of patients reporting ever using telemedicine increased to >80% of patients from all US regions by May 2021 (Figure 2). The Midwestern and Southern US showed lower telemedicine usage overall compared to the West (both $P < 0.01$), and increases in telemedicine over time were slower in the Midwest ($P < 0.001$), South ($P < 0.01$), and Northeast ($P = 0.04$) versus the West. Rural areas were also associated with lower telemedicine use ($P < 0.001$). Rates of telemedicine visits did not vary significantly by tertiles of household income or education (data not shown).

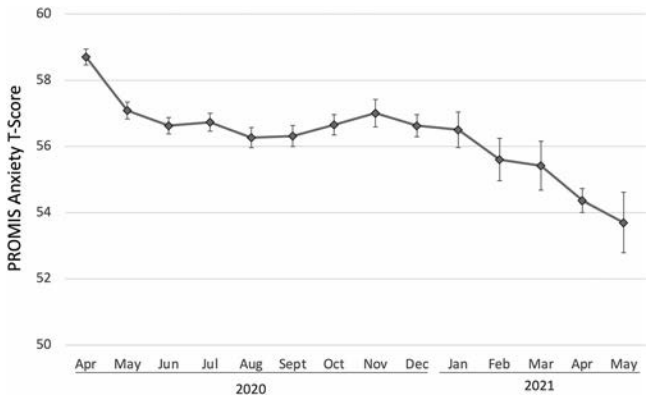

Figure 1. Changes in anxiety over time. Results were graphed using predictions from a generalized estimating equation model adjusted for baseline anxiety. There was a significant reduction in scores over time ($P < 0.001$ for trend; mean ± SD US adult population 50 ± 10). PROMIS = Patient-Reported Outcomes Measurement Information System. Diamonds represent predictions at each month. Bars show the 95% confidence intervals.

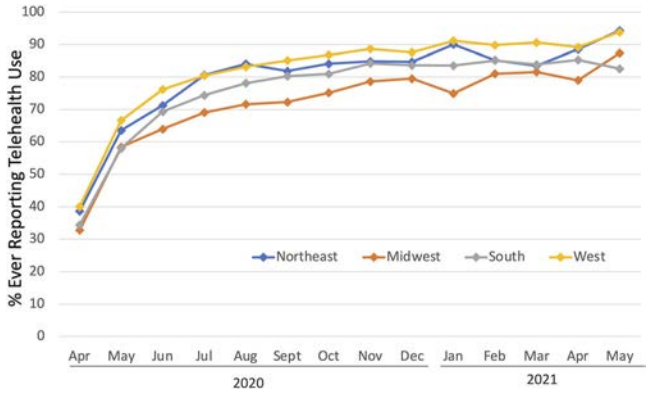

Figure 2. Telemedicine use over time. Results were graphed using predictions from generalized estimating equation logit models. The Midwest region was associated with significantly lower telemedicine use ($P < 0.04$) overall versus the Northeast. Time interactions demonstrate significantly slower increase in telemedicine use in the Midwest ($P < 0.001$) and faster increase in the West ($P = 0.04$) versus the Northeast. Diamonds represent predictions at each month.

Among patients who were not currently reporting a respiratory illness, who had not been diagnosed with COVID-19, and who were taking a DMARD, 191 of 1,748 (10.9%) reported stopping a DMARD at baseline because of concerns regarding COVID-19. Baseline rates of discontinuation of a DMARD were higher (75 of 405 [18.5%]) among patients who did not complete follow-up surveys than among those who had at least 1 additional response (Table 1). Interruptions in DMARD use decreased significantly from April (11.3%) to December 2020 (7.5%) ($P < 0.001$ for trend) but increased from December through May 2021 (14.0%) ($P < 0.001$ for trend) (Figure 3). Discontinuation of a biologic/JAK inhibitor DMARD demonstrated a similar trend, decreasing from April (10.7%) to December (4.8%) ($P < 0.001$), and then increasing through May 2021 (12.2%) ($P < 0.001$) (see Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24837/abstract). A minority of these interruptions in DMARD use was reported to have been recommended by a physician: 214 of 864 (25%) in 2020 versus 53 of 147 (36%) in 2021 ($P < 0.01$). Patients were more likely to report worry about getting sick as a reason for stopping in 2020 versus 2021 (59% versus 28%; $P < 0.001$) and were less likely to give their reason as "other" (16% versus 31%; $P < 0.001$) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24837/abstract). Interruptions in DMARD use were more common in those with higher anxiety, occurring in 10.6% with PROMIS anxiety T score ≥60 versus 8.4% with PROMIS score <60 (odds ratio [OR] 1.29 [95% CI 1.12, 1.49], $P < 0.001$). Telemedicine use was not associated with interruptions in DMARD use (OR 1.15 [95% CI 0.97, 1.36], $P = 0.12$), although interruptions were more likely in patients who reported avoiding an office visit at baseline (10.5% versus 7.6% [OR 1.43 (95% CI 1.13, 1.79)], $P = 0.002$).

Influence of interruptions in DMARD use on subsequent disease flares. A total of 1,464 patients with an ARD had at least 1 survey response during which they were receiving a DMARD, did not report a current flare, and had a subsequent survey available (5,800 total responses). Among this population, 925 (16.0%) reported any flare (median severity 6 [IQR 5–7]), while 516 (8.9%) reported a severe flare (≥6 of 10) at their next survey. In adjusted models, interruptions in DMARD use were associated with a significant increase in severe flares at the next survey (OR 1.71 [95% CI 1.23, 2.36]), with predicted severe flare incidence of 12.9% versus 8.0% (Figure 4). Differences in flares of any severity were not statistically significant (OR 1.18 [0.89, 1.58]) (Figure 4). Both severe flares and flares of any severity were more common in patients receiving glucocorticoids, in younger patients, and in those with PsA or AS (versus RA) but were less common in patients with vasculitis; flares were not associated with the type of DMARDs that patients were receiving (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24837/abstract). Results were similar in analyses restricted to patients with RA, PsA, AS, or SLE (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24837/abstract), in analyses restricted to patients with at least 6 months of follow-up (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24837/abstract), and in analyses adjusted for season (not shown).

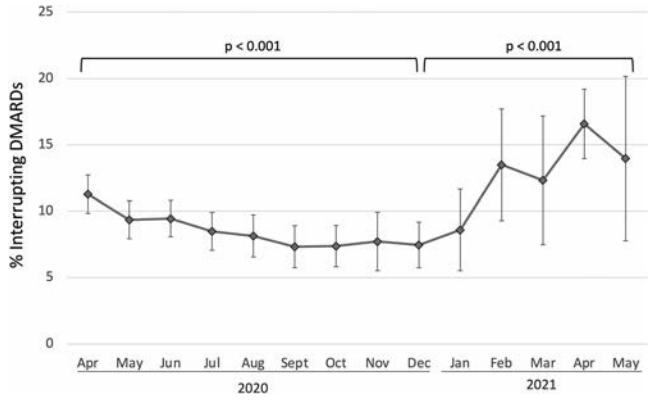

Figure 3. Frequency of interruption in disease-modifying antirheumatic drugs (DMARDs) among patients who reported that they were not sick, did not have COVID-19, and who were taking DMARDs. Results were graphed using predictions from a generalized estimating equation logit model. There was a significant trend for reduction in interruptions from April 2020 to December 2020 ($P < 0.001$) and a significant trend for increase from December 2020 to May 2021 ($P < 0.001$). Diamonds represent predictions at each month. Bars show the 95% confidence intervals.

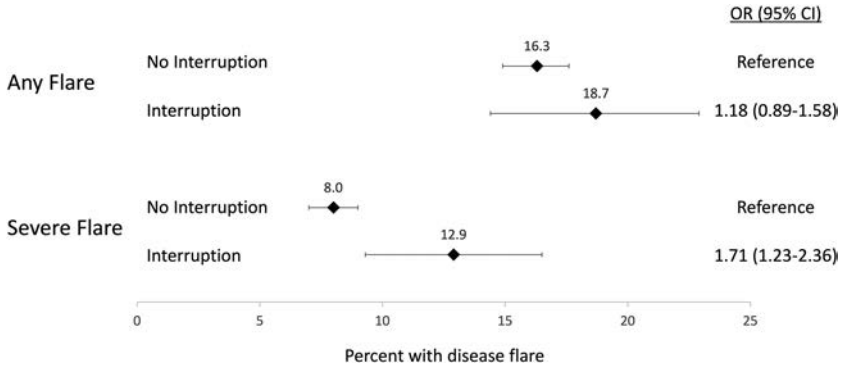

Figure 4. Association between interruptions in disease-modifying antirheumatic drug (DMARD) use and disease flares. Results are from generalized estimating equation logit models assessing the frequency of any flare or of severe flares (rated ≥6 of 10) at the subsequent visit among patients receiving immunomodulatory medications who did not report a respiratory illness or COVID-19 and who did not report currently having a flare. Models also included age, sex, race, autoimmune disease type, glucocorticoid use, DMARD type, and month (full models in Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24837/abstract). 95% CI = 95% confidence interval; OR = odds ratio.

DISCUSSION

In this study of patients with ARDs followed over the course of 14 months, we found frequent interruptions in DMARD use. Importantly, interruptions in DMARD use were associated with self-reported severe disease flares, demonstrating the impact of these DMARD interruptions on patient health. While both interruptions in DMARD use and anxiety decreased during 2020, interruptions in DMARD use increased in 2021, even though anxiety among participants continued to decline.

Several studies have shown that patient concerns early in the pandemic led to frequent health care disruptions, as well as patient interruptions in DMARD use, because of concerns about COVID-19 (11,12,16). Little was known, however, about how rates of interruptions in DMARD use changed as the pandemic progressed. We found that interruptions in DMARD use decreased by >30% by December 2020, with a >50% decrease in biologic/JAK inhibitor interruptions. This reduction may have been due in part to greater comfort with social distancing, discussions with health care providers, and guidance released from the American College of Rheumatology (ACR) in late April 2020 recommending that DMARDs be continued unless patients developed COVID-19 or had a close exposure to an infected person (21). Anxiety, which was strongly associated with interruptions in DMARD use, also decreased substantially over this time period.

Surprisingly, we found a substantial increase in interruptions in DMARD use after December 2020. This increase could be related to a surge of COVID-19 cases during this time period, but because anxiety continued to decrease, it seems possible that interruptions in DMARD use were related to patient or physician concerns about the effect of medications on vaccine efficacy, which were not fully captured in this study. Supporting this assertion, during January to May 2021, patients were less likely to cite "worry about getting sick" as a reason for stopping medications and were more likely to note the reason as "other" (which may include vaccine-related interruptions). Additionally, the proportion of interruptions in DMARD use that were recommended by a physician increased in 2021, although still only accounting for one-third of interruptions. It is possible that some interruptions in DMARD use may have been appropriate even when patients did not discuss with their physician; patients may have independently followed guidance released by the ACR in February 2021 (10,22) to briefly interrupt use of some DMARDs around the time of vaccination, although the intention of this guidance was to encourage shared decision-making. Given the low rate of interruptions directed by a physician, however, it seems likely that some patients stopped DMARDs for longer periods of time because of general concerns about vaccine efficacy.

Few studies have evaluated how interruptions in DMARD use directly affect patient health. One study from Iran found that ~10% of patients with medication nonadherence during the pandemic experienced a worsening of disease symptoms (23). A cross-sectional study from Saudi Arabia also reported that patients with worse medication adherence and more self-titration of medication had worsening of disease activity (24). The longitudinal nature of our study allowed us to examine patients not currently reporting a flare and to evaluate predictors of subsequent flares. We found that interruptions in DMARD use were associated with a substantial increase in self-reported severe disease flares at the next survey. We did not find an association between interruptions in DMARD use and more mild flares, perhaps because of challenges for patients in distinguishing whether a mild increase in symptoms is related to a true inflammatory disease flare or to other causes. Although patient-reported flares may not always match physician assessments (25), previous studies have shown better agreement between patients and physician measures of flares when patient report of flare severity was higher (20), suggesting that flares rated

more severe by patients are more likely to represent true increases in inflammatory disease activity. Additionally, flares rated as more severe by patients presumably have a greater effect on patient health and are more clinically important. As expected, we also found that glucocorticoid use was associated with flares, presumably because patients receiving glucocorticoids are less likely to start in states of lower disease activity.

Some interruptions in DMARD use may be appropriate, and guidance from the ACR recommends brief interruptions in DMARD use in several situations (10). We excluded interruptions in DMARD use due to COVID-19 or to illness but did not capture whether DMARDs were interrupted because of vaccination. We were not able to examine the duration of interruptions in DMARD use, but previous work has shown that short interruptions in DMARDs are less likely to lead to flares than longer interruptions (26). Because our results combine interruptions of different lengths in use of a DMARD, it seems likely that the actual risk of flare with prolonged interruptions is higher than what we found. These results highlight the importance of maintaining continuity of care and avoiding unnecessary or prolonged interruptions in DMARD use during the pandemic and in future public health crises.

We previously found that patients who avoided office visits early in the pandemic were the most likely group to stop a DMARD, but that patients who replaced these missed visits with telemedicine were less likely to stop a DMARD (11,16). In the current study, avoiding office visits at baseline was associated with increased future interruptions in DMARD use, but there was no association between telemedicine and interruptions in DMARD use. This difference may be due to the rapid uptake of telemedicine in this cohort, with >80% of patients reporting telemedicine use within 4 months. In addition, while telemedicine may help maintain continuity of care, use of telemedicine may reflect higher local rates of COVID-19 or greater patient anxiety. Additionally, patients with the highest likelihood of stopping DMARDs (those who lost contact with the health care system) may have been less likely to participate in this study; notably, we found that patients who did not complete follow-up surveys were substantially more likely to have stopped a DMARD than patients who completed follow-up surveys. Methods to proactively identify patients who lose contact with the health care system may help prevent unnecessary interruptions in DMARD use.

Several limitations are important to note. The majority of the survey population were White, female, and of higher socioeconomic status, and responses may not reflect those of underrepresented populations. Given that patients who answered only the baseline survey were more likely to stop medications than those that were followed over several months, our results likely underestimate rates of interruptions in DMARD use in the general population. Not all patients answered every survey, but results were similar in analyses restricted to patients with at least 6 months of follow-up. Flares were based on patient self-report; we did not include the Rheumatoid Arthritis Flare Questionnaire because many patients in the study did not have RA, although a key question in this questionnaire is a patient self-report of flare (27). We also did not capture details on the duration of interruptions in DMARD use and could not compare the effects of brief interruptions in DMARD use (as might occur around the time of vaccination) to longer duration interruptions. Lack of time anchors could also have led some patients to continue to report that a medication had been interrupted even once it had been resumed, but this misclassification would tend to bias time trends and associations with disease flares toward the null.

In conclusion, the COVID-19 pandemic led to substantial anxiety and high rates of interruptions in DMARD use among patients with ARDs, and these interruptions were associated with a higher frequency of self-reported severe disease flares. While anxiety and interruptions in DMARD use improved throughout 2020, interruptions in DMARD use increased in 2021, perhaps related to self-discontinuation or physician-directed discontinuation around the time of vaccination. Avoiding unnecessary and prolonged interruptions in DMARD use may help avoid disease flares. Ensuring that patients have continuity of care and continued communication with their physicians is critical to helping patients navigate a confusing and constantly changing landscape.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Dharia, Venkatachalam, Baker, Banerjee, D. Curtis, Danila, Gavigan, Gordon, Merkel, Shaw, Young, J. R. Curtis, Nowell, George.

Acquisition of data. Venkatachalam, D. Curtis, Gavigan, Young.

Analysis and interpretation of data. Dharia, Venkatachalam, Baker, J. R. Curtis, Nowell, George.

ROLE OF THE STUDY SPONSOR

Eli Lilly and Company and Janssen Pharmaceuticals had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Eli Lilly and Company and Janssen Pharmaceuticals.

REFERENCES

1. Li H, Liu SM, Yu XH, Tang SL, Tang CK. Coronavirus disease 2019 (COVID-19): current status and future perspectives. Int J Antimicrob Agents 2020;55:105951.
2. Listing J, Gerhold K, Zink A. The risk of infections associated with rheumatoid arthritis, with its comorbidity and treatment. Rheumatology (Oxford) 2013;52:53–61.
3. Ogdie A, Maliha S, Shin D, Love TJ, Baker J, Jiang Y, et al. Cause-specific mortality in patients with psoriatic arthritis and rheumatoid arthritis. Rheumatology (Oxford) 2017;56:907–11.

4. Danza A, Ruiz-Irastorza G. Infection risk in systemic lupus erythematosus patients: susceptibility factors and preventive strategies. Lupus 2013;22:1286–94.
5. Ye C, Cai S, Shen G, Guan H, Zhou L, Hu Y, et al. Clinical features of rheumatic patients infected with COVID-19 in Wuhan, China. Ann Rheum Dis 2020;79:1007–13.
6. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature 2020;584:430–6.
7. England BR, Roul P, Yang Y, Kalil AC, Michaud K, Thiele GM, et al. Risk of COVID-19 in rheumatoid arthritis: a National Veterans Affairs matched cohort study in at-risk individuals. Arthritis Rheumatol 2021;73:2179–88.
8. Gianfrancesco M, Hyrich KL, Al-Adely S, Carmona L, Danila MI, Gossec L, et al. Characteristics associated with hospitalisation for COVID-19 in people with rheumatic disease: data from the COVID-19 Global Rheumatology Alliance physician-reported registry. Ann Rheum Dis 2020;79:859–66.
9. Strangfeld A, Schäfer M, Gianfrancesco MA, Lawson-Tovey S, Liew JW, Ljung L, et al. Factors associated with COVID-19-related death in people with rheumatic diseases: results from the COVID-19 Global Rheumatology Alliance physician-reported registry. Ann Rheum Dis 2021;80:930–42.
10. Curtis JR, Johnson SR, Anthony DD, Arasaratnam RJ, Baden LR, Bass AR, et al. American College of Rheumatology guidance for COVID-19 vaccination in patients with rheumatic and musculoskeletal diseases: version 1. Arthritis Rheumatol 2021;73:1093–107.
11. George MD, Baker JF, Banerjee S, Busch H, Curtis D, Danila MI, et al. Social distancing, health care disruptions, telemedicine use, and treatment interruption during the COVID-19 pandemic in patients with or without autoimmune rheumatic disease. ACR Open Rheumatol 2021;3:381–9.
12. George MD, Danila MI, Watrous D, Reddy S, Alper J, Xie F, et al. Disruptions in rheumatology care and the rise of telehealth in response to the COVID-19 pandemic in a community practice–based network. Arthritis Care Res (Hoboken) 2021;73:1153–61.
13. Patient-Centered Outcomes Research Institute. ArthritisPower (Arthritis Partnership With Comparative Effectiveness Researchers). URL: https://www.pcori.org/research-results/2015/arthritispower-arthritis-partnership-comparative-effectiveness-researchers.
14. Nowell WB, Curtis D, Thai M, Wiedmeyer C, Gavigan K, Venkatachalam S, et al. Digital interventions to build a patient registry for rheumatology research. Rheum Dis Clin North Am 2019;45:173–86.
15. Ogdie A, Yu Y, Haynes K, Love TJ, Maliha S, Jiang Y, et al. Risk of major cardiovascular events in patients with psoriatic arthritis, psoriasis and rheumatoid arthritis: a population-based cohort study. Ann Rheum Dis 2015;74:326–32.
16. George MD, Venkatachalam S, Banerjee S, Baker JF, Merkel PA, Gavigan K, et al. Concerns, healthcare use, and treatment interruptions in patients with common autoimmune rheumatic diseases during the COVID-19 pandemic. J Rheumatol 2021;48:603–7.
17. National Center for Health Statistics. Urban–rural classification scheme for counties. 2019. URL: https://www.cdc.gov/nchs/data_access/urban_rural.htm.
18. US Census Bureau. American Community Survey 2014–2018 (5-year estimates). URL: http://www.socialexplorer.com/pub/reportdata/HtmlResults.aspx?reportid=R12543316.
19. Patient-Reported Outcomes Measurement Information System. PROMIS anxiety scoring manual. URL: http://www.healthmeasures.net/images/PROMIS/manuals/PROMIS_Anxiety_Scoring_Manual.pdf.
20. Bykerk VP, Bingham CO, Choy EH, Lin D, Alten R, Christensen R, et al. Identifying flares in rheumatoid arthritis: reliability and construct validation of the OMERACT RA Flare Core domain set. RMD Open 2016;2:e000225.
21. Mikuls TR, Johnson SR, Fraenkel L, Arasaratnam RJ, Baden LR, Bermas BL, et al. American College of Rheumatology guidance for the management of rheumatic disease in adult patients during the COVID-19 pandemic: version 1. Arthritis Rheumatol 2020;72:1241–51.
22. American College of Rheumatology. COVID-19 vaccine clinical guidance summary for patients with rheumatic and musculoskeletal diseases. URL: https://www.rheumatology.org/Portals/0/Files/COVID-19-Vaccine-Clinical-Guidance-Rheumatic-Diseases-Summary.pdf.
23. Khabbazi A, Kavandi H, Paribanaem R, Khabbazi R, Malek Mahdavi A. Adherence to medication in patients with rheumatic diseases during COVID-19 pandemic. Ann Rheum Dis 2020 doi: 10.1136/annrheumdis-2020-218756. E-pub ahead of print.
24. Hassen LM, Almaghlouth IA, Hassen IM, Daghestani MH, Almohisen AA, Alqurtas EM, et al. Impact of COVID-19 outbreak on rheumatic patients' perceptions and behaviors: a cross-sectional study. Int J Rheum Dis 2020;23:1541–9.
25. Barton JL, Imboden J, Graf J, Glidden D, Yelin EH, Schillinger D. Patient-physician discordance in assessments of global disease severity in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2010;62:857–64.
26. Park JK, Kim MJ, Choi Y, Winthrop K, Song YW, Lee EB. Effect of short-term methotrexate discontinuation on rheumatoid arthritis disease activity: post-hoc analysis of two randomized trials. Clin Rheumatol 2020;39:375–9.
27. Bartlett SJ, Barbic SP, Bykerk VP, Choy EH, Alten R, Christensen R, et al. Content and construct validity, reliability, and responsiveness of the Rheumatoid Arthritis Flare Questionnaire: OMERACT 2016 workshop report. J Rheumatol 2017;44:1536–43.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 741–747
DOI 10.1002/acr.24831

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Increased Risk of COVID-19 in Patients With Rheumatoid Arthritis: A General Population-Based Cohort Study

Yilun Wang,[1] Kristin M. D'Silva,[2] April M. Jorge,[2] Xiaoxiao Li,[1] Houchen Lyv,[3] Jie Wei,[1] Chao Zeng,[1] Guanghua Lei,[1] and Yuqing Zhang[2]

Objective. Patients with rheumatoid arthritis (RA) are at an increased risk of acquiring infections owing to immunologic dysfunction and use of potent immunomodulatory medications; however, few data are available on their risk of COVID-19. We estimated the rate of COVID-19 among RA participants and compared it with that of the general population.

Methods. Using the Health Improvement Network, we identified RA patients before February 2020 and followed them to September 2020. We calculated the rate of COVID-19 among participants with RA and compared it with that of the general population using a Cox proportional hazards model, adjusting for potential confounders using overlap weighting of exposure score. We repeated the same analysis among participants with osteoarthritis, a nonautoimmune rheumatic disease, as a negative control exposure.

Results. We identified 225 cases of suspected and confirmed COVID-19 among 17,268 RA patients, and 14,234 cases among 1,616,600 participants in the general population (1.4 versus 0.9/1,000 person-months), with the adjusted hazard ratio (HR_{adj}) being 1.19 (95% confidence interval [95% CI] 1.04–1.36). Confirmed COVID-19 cases developed in 46 RA participants and in 2,249 in the general population (0.3 versus 0.1/1,000 person-months), with the HR_{adj} being 1.42 (95% CI 1.01–1.95). No statistically significant difference was observed for suspected and confirmed (HR 1.00 [95% CI 0.93–1.07]) or confirmed (HR 1.08 [95% CI 0.92–1.27]) COVID-19 rates between participants with osteoarthritis and the general population.

Conclusion. RA, but not osteoarthritis, was associated with an increased risk of COVID-19. Our findings provide timely evidence to support recommendations that booster vaccines and priority access to anti–SARS–CoV-2 monoclonal antibody treatments should be encouraged for RA patients.

INTRODUCTION

COVID-19 has become a global health crisis. By the end of 2020, >100 million cases have been diagnosed worldwide. To date, several studies have reported that patients with rheumatic diseases were at an increased risk of hospitalization (1) and had more severe sequalae after COVID-19 (2–4). However, the risk of developing COVID-19 may vary among the different rheumatic diseases (5), and most prior studies have evaluated heterogeneous rheumatic diseases in aggregate (1–4).

Rheumatoid arthritis (RA) is a common systemic autoimmune disorder, and patients with RA are at an increased risk of acquiring infections owing to immunologic dysfunction and use of potent immunomodulatory medications. To date, 3 studies have specifically compared the risk of COVID-19 between patients with RA and those without RA (6–8); the findings,

The interpretation of the data herein is the sole responsibility of the authors.

Supported by the National Natural Science Foundation of China (grants 81772413, 81930071, 81902265, and 82072502), the National Key Research and Development Project (grant 2018YFB1105705), the Project Program of the National Clinical Research Center for Geriatric Disorders (Xiangya Hospital; grant 2020LNJJ03), the Science and Technology Program of Hunan Province (grant 2019RS2010), and the Key Research and Development Program of Hunan Province (grant 2018SK2070).

[1]Yilun Wang, MD, PhD, Xiaoxiao Li, MS, Jie Wei, PhD, Chao Zeng, MD, PhD, Guanghua Lei, MD, PhD: Xiangya Hospital, Central South University, Changsha, China; [2]Kristin M. D'Silva, MD, April M. Jorge, MD, Yuqing Zhang, DSc: Massachusetts General Hospital, Harvard Medical School, Boston; [3]Houchen Lyv, MD, PhD: Xiangya Hospital, Central South University, Changsha, China, and General Hospital of Chinese PLA, Beijing, China.

Drs. Lei and Zhang contributed equally to this work.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24831&file=acr24831-sup-0001-Disclosureform.pdf.

Address correspondence to Guanghua Lei, MD, PhD, Department of Orthopaedics, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, Hunan 410008, China. Email: lei_guanghua@csu.edu.cn.

Submitted for publication June 10, 2021; accepted in revised form December 2, 2021.

SIGNIFICANCE & INNOVATIONS

- This study shows that the risk of COVID-19 is higher among patients with rheumatoid arthritis (RA) than the general population.
- Patients with osteoarthritis, a nonautoimmune rheumatic disease, do not present a higher risk of COVID-19 than the general population.
- Our findings provide timely evidence to support recommendations that booster vaccines and priority access to anti–SARS–CoV-2 monoclonal antibody treatments should be encouraged for patients with RA.

however, were inconclusive. Two studies found that patients with RA had a higher risk of COVID-19 than individuals without RA (6,7), but another failed to confirm it (8).

To our knowledge, no study has been conducted to describe the incidence of COVID-19 among patients with RA and compare it with that among the general population. To address this knowledge gap, we conducted 2 cohort studies to estimate the risk of COVID-19 among patients with RA and those with osteoarthritis (OA), a common nonautoimmune rheumatic disease, and compared it with that of the general population.

MATERIALS AND METHODS

Data source. The Health Improvement Network (THIN) is an electronic medical record database including general practitioner (GP) records in the UK and represents the UK population regarding demographic characteristics and medical conditions (THIN is a registered trademark of Cegedim in the UK and other countries). Reference made to the THIN database herein is intended to be descriptive of the data asset licensed by IQVIA. This study uses deidentified data provided by patients as a part of their routine primary care. It was approved by the THIN Scientific Review Committee (20SRC003_A1).

Study design and cohort definition. We conducted a cohort study to compare the risks of COVID-19 between patients with RA with the general population without RA. RA diagnosis was made using Read codes (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24831/abstract) (9) that have been previously validated in the UK General Practice Research Database, with a positive predictive value of ~80%. Eligible participants included those between 18 and ~90 years of age and had at least 1 year of continuous enrollment with a general practice before January 29, 2020 (i.e., the index date when the first COVID-19 case was diagnosed in the UK). Participants were followed until the middle of September 2020. Individuals were excluded if they had missing information on body mass index, smoking status, alcohol use, or socioeconomic deprivation index score.

As there has been no purported association between OA and the risk of COVID-19, we further conducted a cohort study to compare the risks of COVID-19 between the patients with OA (i.e., a negative comparison group) and the general population without OA. OA diagnosis was made using Read codes according to previous studies using the THIN database (10). This approach has been preferred in validation studies, as opposed to other approaches, such as medical visits, referrals, or prescription records. Eligible participants included those age ≥40 years and who had at least 1 year of continuous enrollment with a general practice before the index date. Participants were also followed until the middle of September 2020. Exclusion criteria were in line with the RA cohort. This study received approval from the Medical Ethics Committee at Xiangya Hospital (2018091077), with waiver of informed consent.

Assessment of outcomes. The primary outcome was a composite of suspected and confirmed diagnoses of COVID-19 (suspected/confirmed COVID-19) based on Read codes recommended in national guidelines (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24831/abstract) (11). The secondary outcome was a confirmed diagnosis of COVID-19 (11). According to National Health Service guidance and standard operating procedures for primary care and UK Faculty of Clinical Informatics guidelines, confirmed COVID-19 codes represent a positive reverse transcriptase polymerase chain reaction (RT-PCR) test result, while a suspected COVID-19 code represents a symptomatic presentation of COVID-19 and/or contact history with a confirmed patient. A recent study on suspected COVID-19 codes recorded in primary care suggested that clinical diagnosis of COVID-19 by physicians followed a similar trend to test positive cases confirmed by the UK National Testing Service.

Assessment of covariates. Sociodemographic, anthropometric, and lifestyle factors were assessed using the nearest available data prior to the index date; comorbidities were assessed before the index date; and medication use as well as health care utilization were assessed within 1 year prior to the index date (Table 1). These covariates were chosen, as they are potentially causal for RA, OA, and COVID-19.

Statistical analysis. Person-months of follow-up for each participant were calculated as the amount of time from the index date to the first of the following events: COVID-19, death, age 90 years, transferring out of the THIN GP practice, or the end of study follow-up on September 16, 2020. We calculated the incidence rates of suspected/confirmed and confirmed COVID-19, respectively. We estimated rate differences (RDs) between the

Table 1. Baseline characteristics of patients with rheumatoid arthritis (RA) and the general population without RA*

Variable	RA (n = 17,268)	Non-RA (n = 1,616,600)	Stand. diff. before overlap weighting	Stand. diff. after overlap weighting
Demographic characteristic				
Age, mean ± SD years	64.9 ± 13.5	53.3 ± 16.8	0.762	<0.001
Socioeconomic deprivation index score, mean ± SD†	2.8 ± 1.3	2.8 ± 1.3	0.026	<0.001
Women	71.2	54.0	0.362	<0.001
BMI, mean ± SD kg/m^2	28.3 ± 6.4	27.7 ± 6.0	0.098	<0.001
Region			0.079	0.007
England	27.7	31.2		
Northern Ireland	10.9	10.7		
Scotland	33.8	32.3		
Wales	27.6	25.8		
Lifestyle factors				
Drinking			0.221	<0.001
None	26.8	19.0		
Past	5.0	3.1		
Current	68.2	77.9		
Smoking			0.223	<0.001
None	48.6	56.5		
Past	35.8	25.5		
Current	15.6	18.0		
Comorbidity				
Hypertension	44.3	25.4	0.404	<0.001
Diabetes mellitus	19.2	12.5	0.187	<0.001
Chronic kidney disease	12.6	4.6	0.286	<0.001
Pneumonia or infection	11.3	5.9	0.196	<0.001
Chronic obstructive pulmonary disease	10.0	3.6	0.257	<0.001
Influenza	5.1	3.2	0.097	<0.001
Cancer	12.5	7.8	0.156	<0.001
Venous thromboembolism	5.5	2.2	0.173	<0.001
Atrial fibrillation	6.4	3.2	0.150	<0.001
Ischemic heart disease	11.6	5.7	0.212	<0.001
Congestive heart failure	3.8	1.6	0.136	<0.001
Stroke	3.9	2.1	0.107	<0.001
Medication‡				
Antihypertensive	50.2	29.2	0.440	<0.001
Antidiabetic medicine	10.3	7.0	0.117	<0.001
Statin	37.2	20.2	0.383	<0.001
Loop diuretics	9.7	3.3	0.262	<0.001
Thiazide diuretics	6.4	3.7	0.124	<0.001
Health care utilization, mean ± SD				
Hospitalizations‡	0.6 ± 1.5	0.3 ± 0.9	0.247	<0.001
General practice visits‡	7.1 ± 7.1	3.4 ± 4.5	0.624	<0.001
Specialist referrals‡	0.7 ± 1.2	0.4 ± 0.9	0.274	<0.001

* Values are the percentage unless indicated otherwise. BMI = body mass index; stand. diff. = standard difference.
† The socioeconomic deprivation index score was measured by the Townsend Deprivation Index, which was grouped into quintiles from 1 (least deprived) to 5 (most deprived).
‡ Frequency during the past 1 year.

RA group and the comparison group. We performed a Cox proportional hazards model to examine the relation of RA to the risk of suspected/confirmed (or confirmed) COVID-19. We used the Fine and Gray approach to account for the competing risk of death. Specifically, if a person was diagnosed with COVID-19 after the index date, the outcome variable was defined as "Yes" (code 1), and the follow-up time was calculated from the index date to the date of COVID-19 diagnosis. If a person lost follow-up before she/he was diagnosed with COVID-19, the outcome variable was defined as "No" (code 0), and the follow-up time was calculated from the index date to the date of lost follow-up. If a person died before she/he developed COVID-19, the outcome variable was defined as "No" but coded as 2 (a competing risk), and the follow-up time was calculated from the index date to the date of death. We tested the proportional hazards assumption by using the Kolmogorov-type supremum test. We used exposure score, analogous to propensity score, overlap weighting to balance baseline characteristics between the compared groups. Specifically, the exposure score for RA was calculated using the logistic regression model with the

covariates described previously. Individuals with RA were weighted by the probability of not having RA (i.e., 1-exposure score), and individuals without RA were weighted by the probability of having RA (i.e., exposure score). Overlap weights were bounded and smoothly reduced the influence of individuals at the tails of the exposure score distribution without making any exclusions. We repeated the same analysis among patients with OA as a negative control exposure.

We further performed 4 sensitivity analyses to assess the robustness of our findings. First, considering the potential role of disease-modifying antirheumatic drug (DMARD) use in COVID-19, we compared RA with and RA without DMARD use to the risk of COVID-19, respectively. Second, because individuals with missing values were not included in our analyses, we performed imputation analyses to account for missing data. Specifically, missing values of the variables listed previously were imputed by a sequential regression method based on a set of covariates as predictors. Third, we compared the risk of COVID-19 between patients with RA and the general population as well as patients with OA and the general population after late April 2020, when COVID-19 testing was available to the general population in the UK (12). Finally, we conducted a cohort study to compare the risk of COVID-19 between patients with RA and patients with OA.

P values less than 0.05 (2-tailed) were considered significant for all tests. All statistical analyses were performed with SAS software, version 9.4, and RStudio, version 1.1,456 (R Foundation). Full references for the statistical analysis are shown in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24831/abstract.

RESULTS

The flow chart depicting the participant selection process is shown in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/

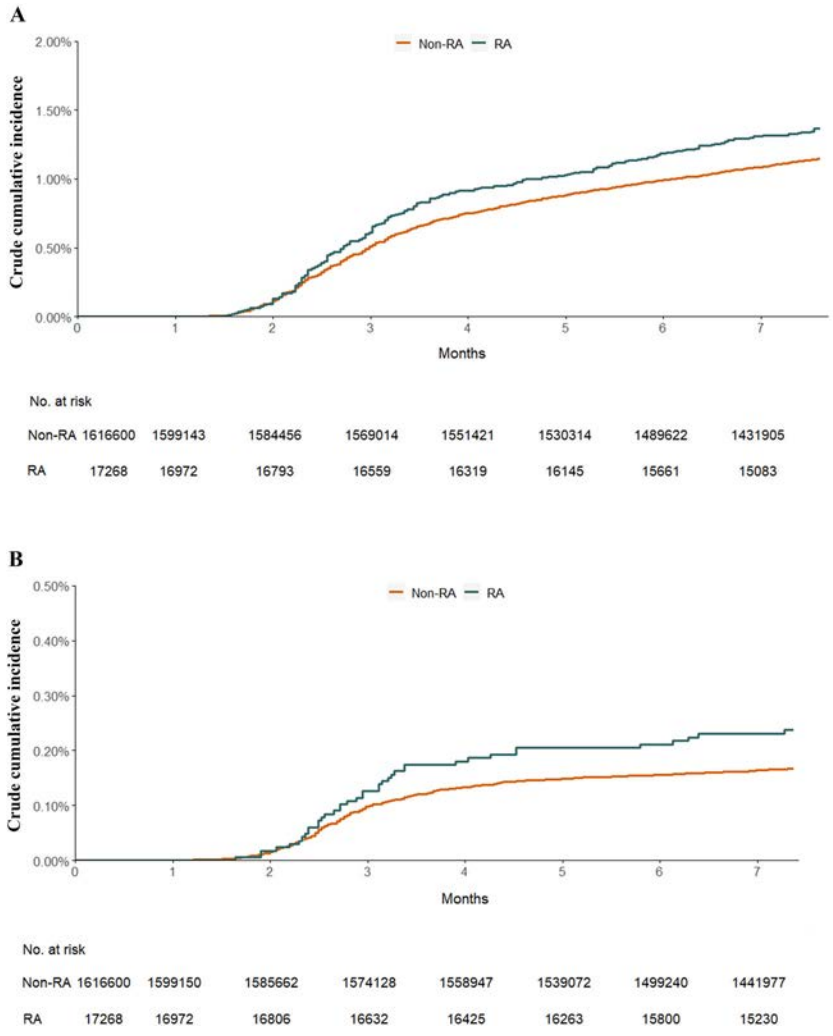

Figure 1. Crude cumulative incidence of suspected and confirmed COVID-19 (**A**) and confirmed COVID-19 (**B**) in 17,268 patients with rheumatoid arthritis (RA) as compared with 1,616,600 individuals without RA.

Table 2. Association between rheumatoid arthritis (RA) and the risk of COVID-19*

	RA (n = 17,268)	Non-RA (n = 1,616,600)
Suspected and confirmed COVID-19		
Event, no.	225	14,234
Mean follow-up, months	7.3	7.3
Incidence rate, per 1,000 person-months	1.4	0.9
Overlap weighted RD (95% CI), per 1,000 person-months	0.3 (0.1–0.5)	0.0 (ref.)
Crude HR (95% CI)	1.49 (1.31–1.69)	1.00 (ref.)
Overlap weighted HR (95% CI)	1.19 (1.04–1.36)	1.00 (ref.)
Overlap weighted HR (95% CI)†	1.20 (1.03–1.44)	1.00 (ref.)
Missing data imputation HR (95% CI)	1.19 (1.04–1.37)	1.00 (ref.)
Confirmed COVID-19		
Event, no.	46	2,249
Mean follow-up, months	7.3	7.4
Incidence rate, per 1,000 person-months	0.3	0.1
Overlap weighted RD (95% CI), per 1,000 person-months	0.1 (0.0–0.2)	0.0 (ref.)
Crude HR (95% CI)	1.93 (1.44–2.58)	1.00 (ref.)
Overlap weighted HR (95% CI)	1.42 (1.01–1.95)	1.00 (ref.)
Overlap weighted HR (95% CI)†	1.53 (1.08–2.28)	1.00 (ref.)
Missing data imputation HR (95% CI)	1.34 (1.07–1.63)	1.00 (ref.)

* The number (rate) of deaths (i.e., competing event) in the RA cohort and the non-RA cohort was 291 (2.3/1,000 person-months) and 9,754 (0.8/1,000 person-months), respectively. 95% CI = 95% confidence interval; HR = hazard ratio; RD = rate difference; ref. = reference.
† Considering April 30, 2020, as the index date.

acr.24831/abstract. In total, we assembled 17,268 participants with RA and 1,616,600 participants without RA. As shown in Table 1, participants with RA were older on average and more likely to be women. However, the baseline characteristics were well balanced between the compared groups after exposure score overlap weighting (all standard differences <0.1).

The risk of suspected/confirmed COVID-19 was higher among the RA group than among the comparison group (Figure 1). As shown in Table 2, 225 cases of suspected/confirmed COVID-19 occurred in the RA group (1.4/1,000 person-months), and 14,234 cases occurred in the non-RA group (0.9/1,000 person-months). The overlap weighted RD of suspected/confirmed COVID-19 between the 2 groups was 0.3/1,000 person-months (95% confidence interval [95% CI 0.1–0.5]), and the adjusted hazard ratio (HR_{adj}) was 1.19 (95% CI 1.04–1.36). Of suspected/confirmed COVID-19 cases, 46 in the RA group (0.3/1,000 person-months) and 2,249 (0.1/1,000 person-months) in the comparison group were diagnosed with confirmed COVID-19. The HR of confirmed COVID-19 for patients with RA versus the general population was 1.42 (95% CI 1.01–1.95).

We assembled 161,065 participants with OA and 779,300 participants without OA from the general population (see Supplementary Figure 2 and Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24831/abstract). No apparent association was observed between OA and risk of COVID-19 (see Supplementary Figure 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24831/abstract). The HRs of suspected/confirmed and confirmed COVID-19 for OA were 1.00 (95% CI 0.93–1.07) and 1.08 (95% CI 0.92–1.27), respectively (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24831/abstract).

Similar results were observed from the sensitivity analyses when we stratified participants with RA according to DMARD use (see Supplementary Tables 4 and 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24831/abstract), performed imputation analyses to account for missing data (Table 2), analyzed the data obtained after April 30, 2020 (Table 2), and compared the risk of COVID-19 between participants with RA and those with OA (see Supplementary Tables 6 and 7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24831/abstract).

DISCUSSION

Using a population-based representative sample from the UK, we found that the rate of COVID-19 was higher among patients with RA, but not OA, than that among the general population. There are limited data regarding the risk of COVID-19 among patients with rheumatic diseases, and few studies that have assessed the risk of COVID-19 infection and its sequalae. A meta-analysis of 6 case–control studies reported that the odds of COVID-19 among patients with rheumatic diseases were 60% higher than that among those without rheumatic diseases (odds ratio [OR] 1.60 [95% CI 1.13–2.25]) (1). Several observational studies also compared the sequalae (e.g., hospitalization, intensive care unit admission, and death) after COVID-19 among

patients with rheumatic diseases (2,3). However, these studies may be susceptible to potential collider bias by conditioning upon having a confirmed COVID-19 diagnosis, leading to biased conclusions (12). To date, 3 studies have specifically compared the risk of COVID-19 between patients with RA to those without RA (6–8). A retrospective study conducted in 7 hospitals in Spain reported that patients with RA had a similar prevalence of hospital PCR-confirmed COVID-19 as those without rheumatic diseases (crude OR 0.98 [95% CI 0.76–1.26]) (8). However, 1 age-, sex-, and Veterans Affairs site–matched cohort study using American Veterans Affairs data showed that veterans with RA were at a higher incidence of PCR-confirmed COVID-19 than non-RA veterans (HR 1.25 [95% CI 1.13–1.39]) after adjusting for demographic information, comorbidities, health behaviors, and county level COVID-19 incidence rates (7). Another cohort study conducted in Denmark found that the risk of COVID-19 hospitalization among patients with RA was much higher than that among the general population (HR 1.72 [95% CI 1.29–2.30]) after adjusting for sex and comorbidities, with age as underlying time scale (6). Likewise, our study findings are in line with those reported in the US and Denmark suggesting that the risk of incident COVID-19 is higher among patients with RA than in the general population, independent of major potential confounders (i.e., sociodemographic characteristics, anthropometrics, lifestyle factors, comorbidities, medication use, and health care utilization).

Several mechanisms may explain the positive association between RA and risk of COVID-19. Studies have demonstrated that premature aging of the immune system in RA contributes to weakened protection against infectious organisms (13). Glucocorticoids are commonly used in the treatment of RA, and there is evidence that glucocorticoid use increases the risk of serious infections in a dose-dependent manner (14). Additionally, other chronic immunosuppressive medications commonly used to treat RA, including biologics and non-biologic DMARDs, may also increase susceptibility to respiratory infections (15), making patients with RA more susceptible to COVID-19.

Our study has several strengths. First, our results were derived from a general population sample from the UK; thus, the study findings are likely generalizable. Second, the risk of developing COVID-19 may vary among the different rheumatic diseases (5). Our study provided the timely evidence that patients with RA, a common autoimmune inflammatory rheumatic disease, but not OA, are at an increased risk of COVID-19. These findings not only could help professional organizations in updating their guidelines around COVID-19 for patients with RA, but they could also shed light on our understanding of the roles of systemic autoimmunity and inflammation in acquiring COVID-19. Third, major potential confounders were addressed and were well balanced after using overlap weights of exposure score. In addition, the sensitivity analyses did not change the results materially, suggesting that the observed associations are robust.

Several limitations of the study deserve comment. First, the sensitivity for capturing COVID-19 cases through the utilized approaches has not been assessed in THIN, and our estimate of COVID-19 incidence could be underestimated. Second, it is possible that patients with RA seek medical care more often and are therefore more likely to have COVID-19 tests than the general population during the COVID-19 pandemic period, which may lead to the higher observed risk of COVID-19 among patients with RA. Since data on the number of COVID-19 tests is unavailable in THIN, we cannot directly assess whether surveillance bias may affect our study findings. Nevertheless, when we compared the risk of COVID-19 infection between 2 comparison cohorts after late April 2020, when COVID-19 testing was available to the general population in the UK (12), RA, but not OA, was still associated with an increased risk of COVID-19. In addition, when we compared the risk of COVID-19 between participants with RA and those with OA, the results were consistent with those between RA and the general population. These findings suggest that a higher risk of COVID-19 among participants with RA than in the general population or participants with OA may not be completely explained by surveillance bias. Third, as we cannot determine the source of the case (i.e., ambulatory encounters or hospitalizations), the testing patterns, and the disease severity of COVID-19 in THIN, future studies are needed to assess the association between RA and the severity of COVID-19. Fourth, it is of great interest to examine whether severity of RA may affect susceptibility to the risk of COVID-19 infection; however, this information was unavailable in THIN. Future studies are required to test this hypothesis. Finally, while the HR generated from imputation analyses (suspected/confirmed COVID-19 HR 1.19; confirmed COVID-19 HR 1.34) was smaller than that from completed data analysis (suspected/confirmed COVID-19 HR 1.19; confirmed COVID-19 HR 1.42), the difference in these effect estimates appears small. Nevertheless, as in any observational study, we cannot rule out the potential selection bias due to missingness, and future studies are needed to verify our findings.

In conclusion, RA, but not OA, was associated with an increased risk of COVID-19. Our findings provide timely evidence to support recommendations that booster vaccines and priority access to anti–SARS–CoV-2 monoclonal antibody treatments should be encouraged for patients with RA.

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. Lei 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. Lei, Zhang.

Acquisition of data. Wang, D'Silva, Jorge, Li, Lyv, Wei, Zeng, Lei, Zhang.

Analysis and interpretation of data. Wang, D'Silva, Jorge, Li, Lyv, Wei, Zeng, Lei, Zhang.

REFERENCES

1. Akiyama S, Hamdeh S, Micic D, Sakuraba A. Prevalence and clinical outcomes of COVID-19 in patients with autoimmune diseases: a systematic review and meta-analysis. Ann Rheum Dis 2021;80:384–391.
2. D'Silva KM, Serling-Boyd N, Wallwork R, Hsu T, Fu X, Gravallese EM, et al. Clinical characteristics and outcomes of patients with coronavirus disease 2019 (COVID-19) and rheumatic disease: a comparative cohort study from a US 'hot spot'. Ann Rheum Dis 2020;79:1156–62.
3. D'Silva KM, Jorge A, Cohen A, McCormick N, Zhang Y, Wallace ZS, et al. COVID-19 outcomes in patients with systemic autoimmune rheumatic diseases compared to the general population: a US multicenter, comparative cohort study. Arthritis Rheumatol 2021;73: 914–20.
4. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature 2020;584:430–6.
5. Hyrich KL, Machado PM. Rheumatic disease and COVID-19: epidemiology and outcomes. Nat Rev Rheumatol 2021;17:71–2.
6. Cordtz R, Lindhardsen J, Soussi BG, Vela J, Uhrenholt L, Westermann R, et al. Incidence and severeness of COVID-19 hospitalisation in patients with inflammatory rheumatic disease: a nationwide cohort study from Denmark. Rheumatology (Oxford) 2021;60: SI59–67.
7. England BR, Roul P, Yang Y, Kalil AC, Michaud K, Thiele GM, et al. Risk of COVID-19 in rheumatoid arthritis: a National Veterans Affairs matched cohort study in at-risk individuals. Arthritis Rheumatol 2021;73:2179–88.
8. Pablos JL, Abasolo L, Alvaro-Gracia JM, Blanco FJ, Blanco R, Castrejon I, et al. Prevalence of hospital PCR-confirmed COVID-19 cases in patients with chronic inflammatory and autoimmune rheumatic diseases. Ann Rheum Dis 2020;79:1170–3.
9. Zhang Y, Lu N, Peloquin C, Dubreuil M, Neogi T, Avina-Zubieta JA, et al. Improved survival in rheumatoid arthritis: a general population-based cohort study. Ann Rheum Dis 2017;76:408–13.
10. Wei J, Wood MJ, Dubreuil M, Tomasson G, LaRochelle MR, Zeng C, et al. Association of tramadol with risk of myocardial infarction among patients with osteoarthritis. Osteoarthritis Cartilage 2020;28:137–45.
11. Chandan JS, Zemedikun DT, Thayakaran R, Byne N, Dhalla S, Acosta-Mena D, et al. Nonsteroidal antiinflammatory drugs and susceptibility to COVID-19. Arthritis Rheumatol 2021;73:731–9.
12. Griffith GJ, Morris TT, Tudball MJ, Herbert A, Mancano G, Pike L, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. Nat Commun 2020;11:5749.
13. Hohensinner PJ, Goronzy JJ, Weyand CM. Telomere dysfunction, autoimmunity and aging. Aging Dis 2011;2:524–37.
14. Dixon WG, Suissa S, Hudson M. The association between systemic glucocorticoid therapy and the risk of infection in patients with rheumatoid arthritis: systematic review and meta-analyses. Arthritis Res Ther 2011;13:R139.
15. Bongartz T, Sutton AJ, Sweeting MJ, Buchan I, Matteson EL, Montori V. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. JAMA 2006;295:2275–85.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 748–758
DOI 10.1002/acr.24523

One-Year Treatment Outcomes of Secukinumab Versus Tumor Necrosis Factor Inhibitors in Spondyloarthritis: Results From Five Nordic Biologic Registries Including More Than 10,000 Treatment Courses

Bente Glintborg,[1] Ulf Lindström,[2] Daniela Di Giuseppe,[3] Sella Aarrestad Provan,[4] Bjorn Gudbjornsson,[5] Merete Lund Hetland,[1] Brigitte Michelsen,[6] Johan K. Wallman,[7] Kalle Aaltonen,[8] Anna-Mari Hokkanen,[9] Dan Nordström,[9] Tanja Schjødt Jørgensen,[10] Rebekka Lund Hansen,[10] Arni Jon Geirsson,[11] Kathrine Lederballe Grøn,[12] Niels Steen Krogh,[13] Johan Askling,[3] Lars Erik Kristensen,[10] and Lennart T. H. Jacobsson,[2] on behalf of the Danish Rheumatology Database (DANBIO), Anti-Rheumatic Therapy in Sweden/Swedish Rheumatology Quality (ARTIS/SRQ), Center for Rheumatology Research (ICEBIO), Finnish Register of Biological Treatment (ROB-FIN), and Norwegian Antirheumatic Drug Register (NOR-DMARD) registries

Objective. To describe baseline characteristics and to compare treatment effectiveness of secukinumab versus tumor necrosis factor inhibitors (TNFi) in patients with spondyloarthritis (SpA) using adalimumab as the main comparator.

Methods. This was an observational, prospective cohort study. Patients with SpA (clinical ankylosing spondylitis, nonradiographic axial SpA, or undifferentiated SpA) starting secukinumab or a TNFi during 2015–2018 were identified from 5 Nordic clinical rheumatology registries. Data on comorbidities and extraarticular manifestations (psoriasis, uveitis, and inflammatory bowel disease) were captured from national registries (data available in 94% of patients) and included in multivariable analyses. We assessed 1-year treatment retention (crude survival curves, adjusted hazard ratios [HR_{adj}] for treatment discontinuation) and 6-month response rates (Ankylosing Spondylitis Disease Activity Score [ASDAS] score <2.1, Bath Ankylosing Spondylitis Disease Activity Index [BASDAI] <40 mm, crude/LUNDEX-adjusted, adjusted logistic regression analyses with odds ratios [ORs]) stratified by line of biologic treatment (first, second, and third plus).

Results. In total, 10,853 treatment courses (842 secukinumab and 10,011 TNFi, of which 1,977 were adalimumab) were included. The proportions of patients treated with secukinumab during the first, second, and third-plus lines of treatment were 1%, 6%, and 22%, respectively). Extraarticular manifestations varied across treatments, while other baseline characteristics were largely similar. Secukinumab had a 1-year retention comparable to adalimumab as a first or second line of treatment but poorer as a third-plus line of therapy (secukinumab 56% [95% confidence interval (95% CI) 51–61%] versus adalimumab 70% [95% CI 64–75%]; HR_{adj} 1.43 [95% CI 1.12–1.81]). Across treatment lines, secukinumab had poorer estimates for 6-month response rates than adalimumab, statistically significantly only for the third-plus line (adjusted analyses: ASDAS score <2.1 OR 0.56 [95% CI 0.35–0.90]; BASDAI <40 mm OR 0.62 [95% CI 0.41–0.95]). Treatment outcomes varied across the 5 TNFi.

Conclusion. Secukinumab was mainly used in biologics-experienced patients with SpA. Secukinumab and adalimumab performed similarly in patients who had failed a first biologic, although with increasing prior biologic exposure, adalimumab was superior.

Supported in part by grants from NordForsk and Foreum.

[1]Bente Glintborg, MD, Merete Lund Hetland, MD: DANBIO and Rigshospitalet, Glostrup, Denmark, and University of Copenhagen, Copenhagen, Denmark; [2]Ulf Lindström, MD, Lennart T. H. Jacobsson, MD: Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden; [3]Daniela Di Giuseppe, MSc, Johan Askling, MD: Karolinska Institutet, Stockholm, Sweden; [4]Sella Aarrestad Provan, MD: Diakonhjemmet Hospital, Oslo, Norway; [5]Bjorn Gudbjornsson, MD: University Hospital and University of Iceland, Reykjavik, Iceland; [6]Brigitte Michelsen, MD: Diakonhjemmet Hospital, Oslo, Norway, and Sørlandet Sykehus, Kristiansand, Norway; [7]Johan K. Wallman, MD: Lund University, Skane University Hospital, Lund, Sweden; [8]Kalle Aaltonen, MSc (Pharm): Ministry of Social Affairs and Health, Helsinki, Finland; [9]Anna-Mari

SIGNIFICANCE & INNOVATIONS

- In spondylarthritis, secukinumab is mainly prescribed in biologics-experienced patients and not as a first-line biologic.
- Outcomes in difficult-to-treat patients who have failed >2 prior biologics are generally poor.
- Our data did not suggest that secukinumab was superior to adalimumab or other tumor necrosis factor inhibitors (TNFi) after failure of a previous TNFi.

INTRODUCTION

The effect of tumor necrosis factor inhibitors (TNFi) for treatment of spondyloarthritis (SpA) is well established (1). Inhibition of the interleukin-17 (IL-17) signaling pathway represents a newer mode of action. In 2015, the first IL-17 inhibitor (secukinumab) was approved for ankylosing spondylitis based on phase III studies performed on biologics-naive and previously TNFi-treated patients (i.e., the MEASURE trials) (2–4).

Currently, the optimal treatment strategy for SpA in routine care remains to be established (5,6), and randomized head-to-head comparisons of secukinumab versus individual TNFi are awaited (7). Recent recommendations acknowledge these evidence gaps but state that treatment with TNFi should be preferred over secukinumab as a first-line biologic due to familiarity with long-term safety. However, in case of TNFi failure due to lack of effect, secukinumab should be favored (5). Studies applying indirect comparisons based on data acquired from phase III trials of the respective drugs found similar or superior effectiveness of secukinumab versus adalimumab (8).

Many patients treated with biologics in routine care would never be eligible for inclusion in a randomized trial due to atypical disease presentation, low disease activity, comorbidity, and advanced age (9). Thus, observational studies contribute a valuable supplement to results from randomized trials (10,11). A recent Swiss real-life study including 106 biologics-experienced axial SpA patients treated with secukinumab reported similar 1-year treatment effectiveness compared to TNFi (12). The study included no comparisons for secukinumab versus the individual TNFi, although TNFi treatment effectiveness is known to vary across individual TNFi drugs (13,14). Furthermore, concomitant extraarticular manifestations might affect TNFi prescription patterns and observed outcomes.

We have previously described a Nordic epidemiologic research collaboration in inflammatory arthritis aiming to investigate rare exposures or outcomes based on combined data sets from 5 prospective biologic registries enriched with data from national registries (15–17). Within this collaboration, we aimed to explore the following in SpA patients treated with secukinumab versus TNFi (mainly adalimumab) and followed in routine care in the Nordic countries: 1) patient characteristics at treatment start; 2) retention to treatment during the first year, including reasons for withdrawal; and 3) 6-month treatment responses.

MATERIALS AND METHODS

This was an observational cohort study based on 5 Nordic biologics registries: the Danish Rheumatology Database (DANBIO) registry (Denmark); the Anti-Rheumatic Therapy in Sweden/Swedish Rheumatology Quality (SRQ/ARTIS) registry (Sweden); the Finnish Register of Biological Treatment (ROB-FIN; Finland); the Norwegian Antirheumatic Drug Register (NOR-DMARD; Norway); and the Center for Rheumatology Research (ICEBIO) registry (Iceland) (15). In these registries, patients are followed prospectively in routine care with monitoring of treatment and outcomes, as previously described (15).

Population and treatments. Each registry identified adult SpA patients (age ≥18 years) who started a TNFi (infliximab, etanercept, adalimumab, golimumab, or certolizumab pegol) or secukinumab between January 1, 2015, and December 31,

Hokkanen, MD, Dan Nordström, MD: Helsinki University and Helsinki University Hospital, Helsinki, Finland; [10]Tanja Schjødt Jørgensen, MSc, Rebekka Lund Hansen, MD, Lars Erik Kristensen, MD: The Parker Institute, Copenhagen University Hospital, Bispebjerg and Frederiksberg, Copenhagen, Denmark; [11]Arni Jon Geirsson, MD: University Hospital, Reykjavik, Iceland; [12]Kathrine Lederballe Grøn, MD, PhD: DANBIO and Rigshospitalet, Copenhagen, Denmark; [13]Niels Steen Krogh, MSc: Zitelab, Copenhagen, Denmark.

Drs. Glintborg and Lindström contributed equally to this work. Drs. Kristensen and Jacobsson contributed equally to this work.

Dr. Glintborg has received research support from AbbVie, Pfizer, and Bristol Myers Squibb. Dr. Provan has received consulting fees from Novartis (less than $10,000). Dr. Gudbjornsson has received speaking fees from Amgen and Novartis (less than $10,000 each). Dr. Hetland has received consulting fees and/or speaking fees from Eli Lilly and Company, Orion Pharma, Biogen, Pfizer, CellTrion, Merck, and Samsung Bioepis (less than $10,000 each) and research support from Bristol Myers Squibb, MSD, AbbVie, Roche, Novartis, Biogen, and Pfizer. Dr. Michelsen has received consulting fees from Novartis (less than $10,000) and research support from Novartis. Dr. Wallman has received consulting fees from Celgene, Eli Lilly and Company, and Novartis (less than $10,000 each). Dr. Hokkanen has received research support from MSD. Dr. Nordström has received consulting fees from AbbVie, Bristol Myers Squibb, Eli Lilly and Company, MSD, Novartis, Pfizer, Roche, and UCB (less than $10,000 each). Dr. Jørgensen has received speaking fees from AbbVie, Pfizer, Roche, Novartis, UCB, Biogen, and Eli Lilly and Company (less than $10,000 each). Dr. Grøn has received research support from Bristol Myers Squibb. Dr. Kristensen has received speaking fees from Pfizer, AbbVie, Amgen, UCB, Sanofi, Gilead, Bristol Myers Squibb, Biogen, MSD, Novartis, Eli Lilly and Company, and Janssen Pharmaceuticals (less than $10,000 each). Dr. Jacobsson has received consulting fees, speaking fees, and/or honoraria from Pfizer, AbbVie, Novartis, Eli Lilly and Company, and Janssen (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Bente Glintborg, MD, DANBIO and Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Centre of Head and Orthopedics, Rigshospitalet, Copenhagen, Denmark. Email: glintborg@dadlnet.dk.

Submitted for publication August 24, 2020; accepted in revised form November 24, 2020.

2018, regardless of previous treatment with biologic disease-modifying antirheumatic drugs (bDMARDs). SpA was defined as a clinical diagnosis of ankylosing spondylitis, undifferentiated SpA, nonradiographic axial SpA, axial SpA, or by International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) codes M45, M46.0, M46.1, M46.8, M46.9, and M07.2. Treatments with other biologics (e.g., ixekizumab) or JAK inhibitors were excluded due to there being few treatment courses, but each patient could contribute with >1 bDMARD treatment course.

Patient characteristics. Patient characteristics upon start of individual bDMARDs (= baseline) were retrieved from the biologics registries and included sex, age, disease duration, and concomitant treatment with conventional synthetic DMARDs (csDMARDs) (e.g., methotrexate or sulfasalazine). Concomitant treatment with nonsteroidal antiinflammatory drugs is not routinely registered and was not reported.

For each treatment course, comorbidities 0–5 years prior to baseline (for main or contributory diagnoses, see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract) were identified as previous hospital contacts (inpatient or outpatient care) through linkage to the national patient registries of individual countries. The following 8 comorbidities were included: malignancy, infection, congestive heart disease, chronic obstructive or interstitial pulmonary disease, chronic kidney disease, diabetes mellitus, myocardial infarction, and hip or knee prosthesis. In the national patient registries, hospital contacts, including discharge diagnoses, are coded according to the ICD-10. Furthermore, extraarticular SpA manifestations (uveitis, inflammatory bowel disease, and psoriasis) were identified. Not all countries were able to deliver comorbidity linkage data for the full observation period (Iceland [not available]; Denmark [until August 2018]; Sweden, Norway, and Finland [until data censoring, December 31, 2018]).

Disease activity and treatment outcomes at 6 months. Each registry extracted individual-level data on disease activity from –60 to +540 days (if available) from the start of the respective treatment with bDMARDs. Exploratory analyses were performed to select the optimal time windows for baseline and 6-month follow-up data in order to reduce missingness. Thus, baseline data were captured within the time window of –30 to +14 days and during the 6-month follow-up visit from 90 to 270 days. In case >1 visit occurred within a given time window, the one closest to the given time point was selected. No imputation of missing data was performed. Disease activity was reported as Bath Ankylosing Spondylitis Disease Activity Index (BASDAI, mm) score, Bath Ankylosing Spondylitis Functional Index score, Bath Ankylosing Spondylitis Metrology Index score, C-reactive protein (CRP, mg/liter) level, Ankylosing Spondylitis Disease Activity Score (ASDAS), patient global score on a visual analog scale (VAS, 0–100 mm), and patient's score for pain on a VAS.

For all treatment courses with secukinumab and each of the TNFi, drug type and start date were retrieved. In case of treatment withdrawal, date of withdrawal and reason for withdrawal (pregnancy, remission, insufficient response, adverse events, death, or other) were registered. Switching from originator to corresponding biosimilar bDMARD (etanercept, infliximab, or adalimumab) or vice versa was not considered a discontinuation if the respective biosimilar/originator was started within 3 months. Reexposure to the same TNFi was counted as a separate treatment course if the time interval between stop and start was >3 months.

Statistics. All country-specific data were pooled for analyses. Analyses were performed using SAS, version 9.4, and Stata, version 16.1, and followed a predefined study protocol approved by all coauthors.

Baseline characteristics and retention to treatment. Baseline characteristics were evaluated based on percentages and mean ± SDs for patients treated with secukinumab and each of the 5 TNFi. Treatment retention (= duration of treatment) and withdrawals stratified by drug type (secukinumab and each of the 5 TNFi) and line of treatment (first, second, and third plus) were evaluated using Kaplan-Meier analysis, overall and by sex. The main outcome was treatment withdrawal irrespective of reason. The person-time to calculate retention rates was defined as the number of days from treatment start until treatment withdrawal, death, censoring, or 365 days of follow-up, whichever came first.

Similarly, Cox proportional hazards regression analyses for treatment withdrawal within 1 year were performed, crude (with 95% confidence intervals [95% CIs]) and adjusted for age and sex (model A), model A plus baseline values for BASDAI score, CRP level, patient global score on a VAS, and concomitant csDMARDs (model B), and model B plus extraarticular manifestations (uveitis [yes/no], inflammatory bowel disease [yes/no], and psoriasis [yes/no]) and number of different comorbidities (maximum 8, as described above; summed as 0/1/>1) (model C). Adalimumab was the reference drug. Analyses were stratified according to line of treatment. Additional adjustment for country did not change the results; thus, country was not included in the models.

In the multivariate analyses, age was added as a continuous variable, including a quadratic term. The following variables were added as categorical and based on quartiles of the distribution: concomitant csDMARDs (yes/no/missing), BASDAI (≤3.7/>3.7; ≤5.9/>5.9; ≤7.5/>7.5/missing), CRP level (mg/liter) (≤2/>2; ≤4.2/>4.2; ≤11/≥11/missing), and patient global score on a VAS (≤45/>45; ≤65/>65; ≤80/>80/missing).

For model C, only patients with available linkage to national patient registries were included. In all multivariable models, robust SEs were used to adjust for patients contributing ≥1 treatment course.

Treatment response at 6-month follow-up. The proportions of patients with at least 6 months of follow-up and available data who were in low disease activity (ASDAS score <2.1) or who had a BASDAI score of <40 mm at 6 months were identified. Furthermore, the corresponding LUNDEX-corrected response rates (crude) were calculated, thus taking early withdrawal (<6 months) into account (18). Logistic regression analyses for an ASDAS score of <2.1 and a BASDAI score of <40 mm (yes/no) at 6 months among patients still receiving treatment were performed and adjusted as described for models A, B, and C, with adalimumab as the reference drug.

Sensitivity analyses. The following post hoc analyses were performed: multivariate analyses (model C) that 1) compared patients receiving secukinumab to the combined TNFi group, 2) included only the subgroup of patients (second or third-plus treatment courses) for whom the reason for termination of prior TNFi was lack of effect, and 3) added the exact number of previous bDMARD courses to the third-plus treatment group as a further adjustment factor.

Missing data. In order to assess potential bias due to missing baseline data, additional multivariable Cox regression analyses including only patients with complete data on all covariates were performed (i.e., complete case analyses). Furthermore, baseline characteristics of patients with available versus missing measures of treatment response at 6 months were explored.

Ethics approval. This study was approved by the appropriate ethics committees and/or data protection committees in each country (Denmark: RH-2015-209, I-suite 04145; Finland: 73/13/03/00/2014; Iceland: VSNb2017010049/03.01; Norway: 2011/1339 and 2017/243; and Sweden: 2015/1844-31/2). Individual patient consent was not required for the reporting of anonymized registry data for research purposes.

RESULTS

In total, 10,853 treatment courses (842 secukinumab [8%] and 10,011 TNFi [92%]) in 8,050 unique patients were included. Country-specific numbers for secukinumab/TNFi were as follows: Sweden 497/5,286; Denmark 195/2,866; Finland 73/806; Norway 67/796; and Iceland 10/257. Secukinumab was used mostly in biologics-experienced patients (proportions treated with secukinumab during first/second/third-plus treatment courses: 1%/6%/22%) (Table 1), with a similar pattern for all 5 countries (Table 1). History of extraarticular manifestations varied across treatments, with inflammatory bowel disease being more common in patients treated with adalimumab, golimumab, or infliximab (adalimumab was favored in patients with a history of anterior uveitis, and secukinumab was favored in patients with psoriasis). History of other comorbidities was overall similar across different bDMARD and treatment courses, except for heart disease, which was more common in patients starting secukinumab as the first treatment course. Other baseline characteristics were largely similar, including sex distribution, age, concomitant csDMARDs, disease activity, and number of comorbidities (Table 1). The numbers of patients contributing baseline data are shown in Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract. Contributing linkage data from the national registries (n = 10,180; 94%) are shown in Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract.

Retention to treatment. The 1-year treatment retention rates varied across treatments (Table 2 and Figure 1), with secukinumab displaying a drug retention comparable to adalimumab as a first or second line of treatment, but poorer as a third-plus line of therapy: secukinumab 56% (95% CI 51%–61%) versus adalimumab 70% (95% CI 64%–75%) (Table 2). Similar results were found in multivariable Cox regression analyses adjusted for comorbidities and extraarticular manifestations (Table 2 and Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract).

Reasons for withdrawals within the first year of follow-up were mainly lack of effectiveness or adverse events (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract), with a tendency toward higher secukinumab withdrawal due to lack of effectiveness. Secukinumab contributed few withdrawals during first-line therapy (n = 14); thus, these data should be interpreted with caution.

Patients receiving secukinumab had a similar treatment retention compared to the combined TNFi group: first-line hazard ratio (HR) 0.78 (95% CI 0.45–1.36); second-line HR 0.94 (95% CI 0.69–1.28); third-plus–line HR 1.06 (95% CI 0.91–1.24) (model C). Sensitivity analyses restricted to patients in whom the prior TNFi failed due to lack of effect showed results similar to those presented in Table 2. Thus, during the third treatment course, the HR for 1-year treatment withdrawal was 1.60 (95% CI 1.14–2.24) (model C, secukinumab versus adalimumab; details not shown).

Sensitivity analyses adding the number of previous bDMARDs did not change the results markedly (HR 1.46 [95% CI 1.14–1.86] for secukinumab versus adalimumab [model C, third-plus line]). Furthermore, complete case analyses showed similar results (data not shown).

Treatment response at 6-month follow-up. Treatment response differed between secukinumab and the 5 TNFi with LUNDEX-adjusted response rates for an ASDAS score of <2.1 (first line: 27% for secukinumab versus 27–42% for TNFi; second line: 14% versus 23–31%; third-plus line: 12% versus 11–24%) (Figure 2). Similar rates and similar internal relations between the 6 evaluated drugs were seen for a BASDAI score of <40 mm (Figure 2). Baseline characteristics in patients with available

Table 1. Patient characteristics and disease activity at start of treatment, stratified by drug and line of treatment*

Characteristic	First line (n = 5,325)						Second line (n = 2,803)						Third-plus line (n = 2,725)					
	ADA	SEC	CZP	ETN	GOL	IFX	ADA	SEC	CZP	ETN	GOL	IFX	ADA	SEC	CZP	ETN	GOL	IFX
No.	858	70	396	1,720	499	1,782	662	160	247	890	304	540	457	612	322	500	372	462
Sex, men, no. (%)	478 (56)	38 (54)	207 (52)	892 (52)	312 (63)	1,010 (57)	334 (50)	70 (44)	118 (48)	464 (52)	176 (58)	302 (56)	226 (49)	280 (46)	124 (39)	213 (43)	183 (49)	224 (48)
Age, years	42 ± 13	44 ± 14	39 ± 12	41 ± 14	39 ± 13	42 ± 14	44 ± 13	45 ± 13	43 ± 13	44 ± 13	42 ± 13	44 ± 13	46 ± 13	47 ± 12	45 ± 13	46 ± 13	46 ± 13	44 ± 13
Disease duration, years	12 ± 12	12 ± 11	10 ± 10	12 ± 12	11 ± 11	11 ± 11	15 ± 12	15 ± 12	14 ± 11	15 ± 12	14 ± 11	15 ± 11	19 ± 12	18 ± 12	18 ± 12	18 ± 12	18 ± 12	16 ± 11
HAQ score	0.7 ± 0.6	0.8 ± 0.6	0.9 ± 0.7	0.8 ± 0.5	0.7 ± 0.5	0.8 ± 0.6	0.8 ± 0.6	0.9 ± 0.5	0.8 ± 0.6	0.9 ± 0.6	0.8 ± 0.6	0.9 ± 0.6	1.0 ± 0.6	1.1 ± 0.7	1.0 ± 0.6	1.1 ± 0.6	1.1 ± 0.6	1.0 ± 0.6
CRP, mg/liter	10 ± 15	7 ± 8	12 ± 18	11 ± 17	12 ± 16	12 ± 23	10 ± 24	10 ± 18	9 ± 16	9 ± 15	10 ± 19	10 ± 18	11 ± 18	13 ± 27	9 ± 16	10 ± 18	12 ± 23	10 ± 16
Pain score, VAS, mm	53 ± 25	57 ± 28	56 ± 24	58 ± 23	57 ± 23	58 ± 25	54 ± 27	59 ± 23	60 ± 24	59 ± 25	55 ± 29	52 ± 29	59 ± 26	66 ± 23	61 ± 25	63 ± 25	61 ± 24	61 ± 27
Global score, VAS, mm	52 ± 26	55 ± 26	58 ± 25	57 ± 23	53 ± 27	62 ± 25	56 ± 27	60 ± 25	63 ± 24	62 ± 25	56 ± 30	54 ± 30	62 ± 25	68 ± 24	63 ± 25	66 ± 24	62 ± 25	63 ± 27
BASDAI score, mm	48 ± 24	47 ± 28	50 ± 22	52 ± 21	48 ± 23	55 ± 21	51 ± 24	52 ± 22	53 ± 22	55 ± 22	50 ± 62	49 ± 27	56 ± 23	63 ± 22	58 ± 23	59 ± 23	58 ± 24	57 ± 25
BASMI score	11 ± 16	22 ± 23	25 ± 21	13 ± 16	18 ± 18	20 ± 19	18 ± 20	14 ± 22	22 ± 19	22 ± 21	22 ± 24	27 ± 22	26 ± 22	25 ± 19	20 ± 20	25 ± 22	23 ± 24	25 ± 23
BASFI score	34 ± 25	44 ± 31	46 ± 25	39 ± 25	34 ± 23	45 ± 24	39 ± 26	43 ± 23	45 ± 24	46 ± 26	40 ± 27	42 ± 27	47 ± 28	55 ± 26	48 ± 26	50 ± 25	49 ± 27	49 ± 27
ASDAS score	2.9 ± 1.0	3.2 ± 1.1	3.1 ± 1.1	3.1 ± 1.0	3.1 ± 1.0	3.3 ± 1.0	2.9 ± 1.1	3.1 ± 0.9	3.1 ± 1.0	3.1 ± 1.0	3.1 ± 1.2	2.9 ± 1.2	3.1 ± 1.1	3.5 ± 1.1	3.3 ± 1.0	3.2 ± 1.1	3.2 ± 1.1	3.2 ± 1.2
SJC (range 0–28)	1 ± 2	0 ± 1	0 ± 1	1 ± 1	0 ± 1	0 ± 1	1 ± 1	1 ± 2	0 ± 2	1 ± 2	1 ± 2	0 ± 1	1 ± 2	1 ± 2	1 ± 2	1 ± 2	1 ± 2	1 ± 2
Concomitant MTX, %†	6	2	11	4	8	14	7	10	8	6	10	12	5	8	9	8	14	14
Concomitant SSZ, %†	7	2	7	4	10	8	2	5	4	2	4	3	1	3	3	4	3	4
Comorbidities, yes, %‡																		
Malignancy	1.9	1.6	1.1	1.5	1.3	1.6	2.0	4.7	1.2	2.7	1.8	2.2	5.1	4.0	4.5	3.9	3.1	4.0
Pulmonary disease	0.1	1.6	1.3	0.6	0.2	1.1	0.7	2.0	1.6	1.5	1.1	0.8	1.5	1.9	1.0	0.9	0.6	0.9
Congestive heart failure	0.1	7.8	0.3	0.7	0.4	0.3	0.3	2.0	0.8	1.0	0.4	1.0	1.0	0.7	0.6	1.1	0.6	1.1
Diabetes mellitus	1.7	4.7	1.3	2.4	1.3	1.9	2.0	4.7	2.9	2.4	2.1	2.6	2.4	2.8	4.5	2.4	3.4	3.4
Myocardial infarction	0.6	4.7	0.3	0.5	0.4	0.4	0.8	0.7	0.8	1.2	0.7	1.2	1.0	1.8	1.3	1.1	1.1	0.9
Chronic kidney disease	0.4	1.6	0.5	0.3	0.0	0.6	0.2	0.0	0.0	0.6	1.4	1.4	0.2	0.9	1.0	1.3	1.4	0.9
Knee or hip prosthesis	1.0	4.5	0.0	1.4	0.4	0.8	1.3	2.7	0.4	1.2	1.1	1.6	0.5	2.7	2.6	1.3	2.3	1.6
Any infection	19.4	15.6	14.9	20.3	18.6	16.6	26.6	33.1	23.7	24.6	25.0	18.0	33.5	32.7	31.2	34.2	33.6	31.7
No. of comorbidities, %																		
1	19.7	21.9	14.7	21.7	19.4	17.9	27.8	35.1	22.9	27.0	25.7	21.6	32.5	30.9	32.8	35.0	31.6	31.9
≥2	2.6	7.8	2.3	2.7	1.5	2.6	2.6	6.8	3.3	3.6	3.9	3.0	5.5	7.6	5.7	4.9	6.6	5.8
Prior extraarticular manifestations‡																		
IBD	11.4	1.6	2.5	1.6	4.0	8.6	10.9	2.7	2.5	3.9	11.8	9.1	11.3	3.0	7.3	6.4	10.0	10.1
Psoriasis	5.2	10.9	5.8	4.2	2.9	3.7	6.6	13.5	8.6	8.2	4.6	6.7	8.7	13.4	10.5	9.2	8.3	7.0
Uveitis	28.0	3.1	8.6	8.2	15.5	11.7	21.5	10.1	13.1	9.5	15.7	12.5	25.3	12.5	14.3	12.2	17.1	11.2

* Values are the mean ± SD unless indicated otherwise. The number of patients starting secukinumab in total (first/second/third-plus line) per country were: Sweden 497 (36/106/355), Denmark 195 (18/29/148), Finland 73 (15/18/40), Norway 67 (1/4/62), and Iceland 10 (0/3/7). ADA = adalimumab; ASDAS = Ankylosing Spondylitis Disease Activity Score; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; BASMI = Bath Ankylosing Spondylitis Metrology Index; CRP = C-reactive protein; CZP = certolizumab pegol; ETN = etanercept; GOL = golimumab; HAQ = Health Assessment Questionnaire; IBD = inflammatory bowel disease; IFX = infliximab; MTX = methotrexate; SEC = secukinumab; SJC = swollen joint count; SSZ = sulfasalazine; VAS = visual analog scale.
† Based on patients with baseline visit and available data.
‡ Zero to 5 years prior to baseline; for availability, see Supplementary Tables 2 and 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract.

Table 2. Treatment withdrawals and retention after 1 year, stratified by drug type and line of treatment*

Line and drug	No.	No. of events	Person-years	1-year retention rate (95% CI)	HR (95% CI) of withdrawal		
					Model A†	Model B‡	Model C§
First line							
ADA	858	182	625.58	0.74 (0.70–0.77)	1	1	1
SEC	70	14	48.71	0.76 (0.63–0.85)	0.93 (0.54–1.61)	0.99 (0.57–1.71)	0.89 (0.50–1.57)
CZP	396	126	316.12	0.66 (0.61–0.71)	1.40 (1.11–1.75)¶	1.33 (1.03–1.70)¶	1.27 (0.99–1.64)
ETN	1,720	402	1,245.08	0.72 (0.69–0.74)	1.10 (0.92–1.31)	1.08 (0.91–1.30)	1.03 (0.85–1.24)
GOL	499	93	435.69	0.78 (0.74–0.82)	0.78 (0.61–1.00)	0.79 (0.61–1.02)	0.79 (0.61–1.02)
IFX	1,782	585	1,325.62	0.62 (0.59–0.64)	1.53 (1.29–1.80)¶	1.42 (1.19–1.70)¶	1.39 (1.16–1.67)¶
Second line							
ADA	662	167	451.61	0.69 (0.64–0.73)	1	1	1
SEC	160	45	105.24	0.67 (0.58–0.74)	1.08 (0.78–1.50)	1.07 (0.77–1.50)	1.05 (0.75–1.47)
CZP	247	101	178.26	0.55 (0.48–0.61)	1.50 (1.17–1.92)¶	1.39 (1.07–1.80)¶	1.36 (1.05–1.77)¶
ETN	890	288	617.78	0.63 (0.59–0.66)	1.26 (1.04–1.53)¶	1.21 (1.00–1.47)¶	1.18 (0.96–1.44)
GOL	304	80	243.78	0.69 (0.62–0.74)	0.91 (0.70–1.19)	0.93 (0.71–1.22)	0.89 (0.67–1.17)
IFX	540	169	414.41	0.64 (0.59–0.68)	1.12 (0.91–1.39)	1.20 (0.96–1.51)	1.18 (0.94–1.49)
Third-plus line							
ADA	457	107	315.90	0.70 (0.64–0.75)	1	1	1
SEC	612	214	402.51	0.56 (0.51–0.61)	1.53 (1.22–1.93)¶	1.47 (1.16–1.86)¶	1.43 (1.12–1.81)¶
CZP	322	148	203.74	0.47 (0.41–0.53)	2.11 (1.64–2.72)¶	2.14 (1.66–2.76)¶	2.04 (1.57–2.64)¶
ETN	500	156	349.47	0.62 (0.57–0.67)	1.30 (1.02–1.66)¶	1.27 (0.99–1.63)	1.21 (0.94–1.56)
GOL	372	117	270.77	0.64 (0.58–0.69)	1.27 (0.98–1.64)	1.30 (1.00–1.69)¶	1.26 (0.97–1.65)
IFX	462	168	320.77	0.58 (0.53–0.63)	1.55 (1.21–1.98)¶	1.54 (1.21–1.97)¶	1.41 (1.10–1.82)¶

* Results from Kaplan-Meier and Cox regression analyses (crude and adjusted hazard ratios [HRs] for withdrawal). 95% CI = 95% confidence interval; ADA = adalimumab; CZP = certolizumab pegol; ETN = etanercept; GOL = golimumab; IFX = infliximab; SEC = secukinumab.
† Adjusted by age and sex.
‡ Adjusted by sex and baseline age, C-reactive protein level, Bath Ankylosing Spondylitis Disease Activity Index score, patient global score on a VAS, and concomitant conventional synthetic disease-modifying antirheumatic drugs.
§ Model B, adding baseline comorbidity/extraarticular manifestations; see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract, for details including patient numbers.
¶ Statistically significant.

versus unavailable measures of treatment response were similar (see Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract).

In adjusted logistic regression analyses (model C), secukinumab had poorer 6-month response rates than adalimumab, but the difference was only statistically significant for the third-plus line (ASADS score <2.1, OR 0.56 [95% CI 0.35–0.90]; BASDAI score <40 mm, OR 0.62 [95% CI 0.41–0.95]) (Table 3 and Supplementary Table 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract). Results for the comparison of patients receiving secukinumab

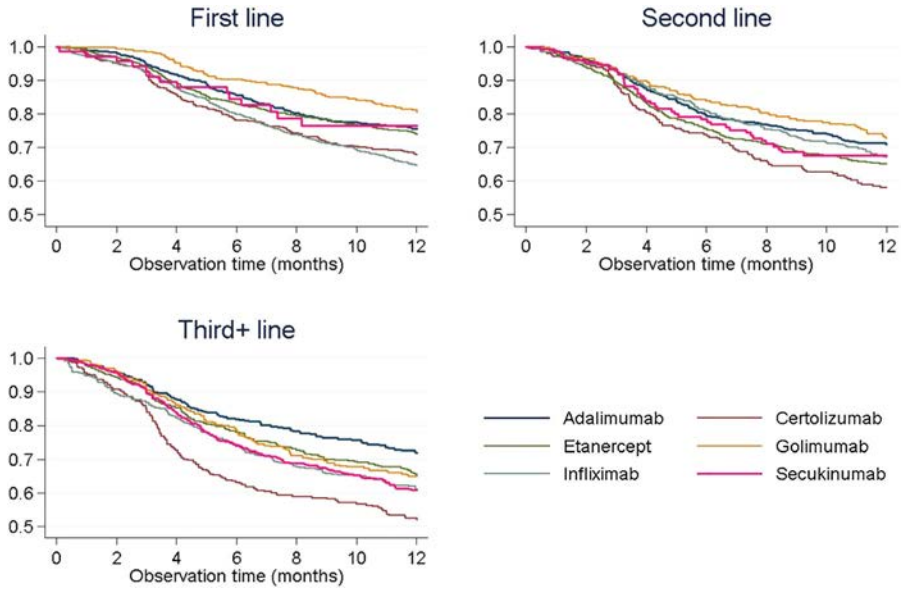

Figure 1. Survival probability curves for secukinumab and each of the 5 tumor necrosis factor inhibitors, stratified by line of treatment.

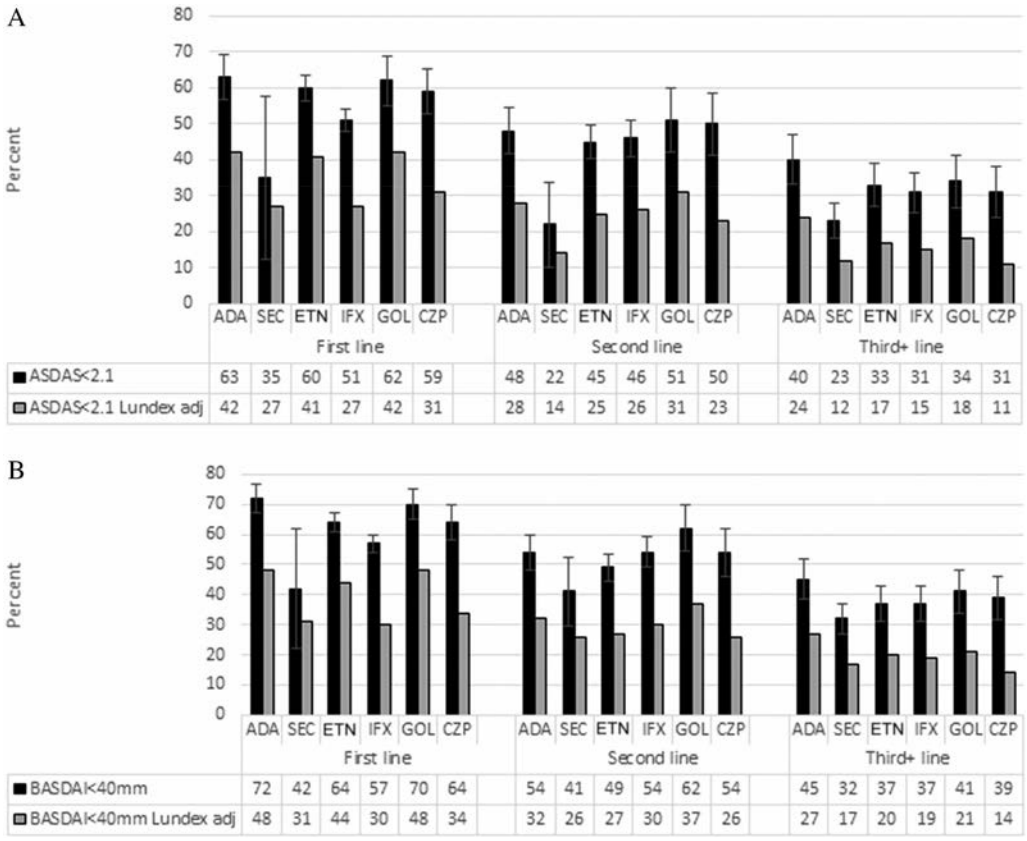

Figure 2. Treatment response after 6 months of treatment, stratified by line of treatment and drug type, for an Ankylosing Spondylitis Disease Activity Score (ASDAS) of <2.1 (**A**) and a Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) score of <40 mm (**B**). Numbers are the percentage of patients with a response (crude with 95% confidence interval [solid bars] and LUNDEX adjusted [shaded bars]). ADA = adalimumab; CZP = certolizumab pegol; ETN = etanercept; GOL = golimumab; IFX = infliximab; SEC = secukinumab.

versus the combined TNFi group were not statistically significant (model C, third-plus treatment ASDAS score <2.1, OR 0.74 [95% CI 0.51–1.07] and BASDAI score <40 mm, OR 0.79 [95% CI 0.58–1.09]) (details not shown).

Sensitivity analyses restricted to patients failing the prior TNFi due to lack of effect showed similar but not statistically significant results for secukinumab versus adalimumab (model C, third-plus line ASDAS score <2.1, OR 0.68 [95% CI 0.34–1.37] and BASDAI score <40 mm, OR 0.59 [95% CI 0.31–1.09]). Sensitivity analyses adding the number of previous bDMARDs to the third-plus line did not markedly change results (ASDAS score <2.1 OR 0.55 [95% CI 0.34–0.89] and BASDAI score <40 mm OR 0.63 [95% CI 0.41–0.98] for secukinumab versus adalimumab).

DISCUSSION

In this observational study including >10,000 patients with SpA treated in routine care from 5 prospective Nordic rheumatology registries during 2015–2018, secukinumab was mainly prescribed in biologics-experienced patients. Through linkage to national registries, we were able to identify comorbidities and extraarticular manifestations. We found marked differences across treatments, with secukinumab more often used in patients with cardiovascular disease and concomitant psoriasis, and adalimumab when there was a history of uveitis. The 6-month treatment outcomes and 1-year treatment retention showed wide variation between the 5 TNFi, with poorer outcomes for secukinumab compared to adalimumab, especially during third-line treatment. We were not able to demonstrate any superior outcomes for secukinumab versus adalimumab in patients in whom ≥1 TNFi had previously failed, neither overall nor in the subgroup that withdrew from TNFi due to lack of effect.

This study adds important knowledge to the gradually emerging evidence regarding routine care use of secukinumab in patients with SpA. Previous studies have demonstrated that secukinumab performs better in biologics-naive than biologics-experienced patients (4,19). However, in accordance with current guidelines (5,20), we found that secukinumab was mainly used in TNFi-experienced patients, and only 1% of first-line treatment courses used secukinumab. Although we demonstrated a similar performance of first-line secukinumab versus adalimumab, secukinumab-exposed patients were too few to draw firm conclusions.

Table 3. Response after 6 months of treatment, stratified by drug type and line of treatment*

Line and drug	Responses, no./total no.†	Adjusted models with OR (95% CI)		
		Model A‡	Model B§	Model C¶
ASDAS score <2.1				
First line				
ADA	107/242	1	1	1
CZP	95/234	0.85 (0.58–1.23)	0.63 (0.41–0.97)#	0.66 (0.43–1.02)
ETN	323/695	1.15 (0.85–1.56)	1.12 (0.82–1.53)	1.18 (0.86–1.64)
GOL	85/185	0.99 (0.66–1.47)	1.04 (0.69–1.56)	1.08 (0.72–1.63)
IFX	389/1,065	0.72 (0.54–0.96)#	0.75 (0.55–1.03)	0.77 (0.56–1.06)
SEC	4/17	0.43 (0.13–1.37)	0.47 (0.14–1.55)	0.57 (0.17–1.88)
Second line				
ADA	76/233	1	1	1
CZP	33/125	0.75 (0.46–1.24)	0.62 (0.36–1.07)	0.65 (0.37–1.13)
ETN	128/428	0.89 (0.63–1.27)	0.84 (0.58–1.22)	0.92 (0.62–1.35)
GOL	50/126	1.24 (0.79–1.97)	1.11 (0.68–1.80)	1.06 (0.64–1.76)
IFX	113/356	0.95 (0.66–1.36)	0.70 (0.46–1.05)	0.69 (0.46–1.06)
SEC	9/49	0.47 (0.22–1.04)	0.51 (0.23–1.16)	0.57 (0.25–1.31)
Third-plus line				
ADA	54/192	1	1	1
CZP	25/154	0.52 (0.30–0.88)#	0.52 (0.30–0.91)#	0.54 (0.30–0.96)#
ETN	56/240	0.79 (0.51–1.23)	0.87 (0.55–1.38)	0.91 (0.57–1.46)
GOL	34/155	0.72 (0.44–1.18)	0.70 (0.42–1.15)	0.68 (0.40–1.14)
IFX	53/264	0.63 (0.41–0.98)#	0.60 (0.38–0.95)#	0.65 (0.41–1.03)
SEC	48/296	0.51 (0.33–0.80)#	0.55 (0.35–0.88)#	0.56 (0.35–0.90)#
BASDAI <40 mm				
First line				
ADA	199/364	1	1#	1
CZP	122/275	0.67 (0.49–0.93)#	0.57 (0.39–0.82)#	0.61 (0.42–0.88)#
ETN	425/851	0.88 (0.68–1.13)	0.98 (0.75–1.28)	1.07 (0.81–1.41)
GOL	169/302	1.01 (0.74–1.38)	1.02 (0.73–1.41)	1.07 (0.77–1.49)
IFX	465/1,179	0.54 (0.42–0.69)#	0.61 (0.47–0.80)#	0.64 (0.48–0.84)#
SEC	8/24	0.46 (0.19–1.13)	0.56 (0.22–1.44)	0.63 (0.24–1.63)
Second line				
ADA	105/283	1	1#	1
CZP	45/146	0.76 (0.49–1.18)	0.62 (0.38–1.02)	0.64 (0.39–1.06)
ETN	157/488	0.79 (0.58–1.08)	0.74 (0.53–1.04)	0.80 (0.56–1.13)
GOL	69/154	1.29 (0.86–1.93)	1.10 (0.71–1.72)	1.05 (0.67–1.67)
IFX	143/376	1.01 (0.73–1.39)	0.78 (0.53–1.13)	0.77 (0.52–1.13)
SEC	22/70	0.80 (0.45–1.41)	0.73 (0.40–1.34)	0.85 (0.46–1.59)
Third-plus line				
ADA	67/212	1	1#	1
CZP	36/183	0.54 (0.34–0.87)#	0.53 (0.32–0.89)#	0.55 (0.33–0.93)#
ETN	66/258	0.76 (0.51–1.14)	0.81 (0.53–1.25)	0.86 (0.56–1.34)
GOL	51/187	0.80 (0.51–1.23)	0.76 (0.48–1.20)	0.76 (0.48–1.21)
IFX	70/285	0.71 (0.47–1.05)#	0.68 (0.45–1.04)	0.74 (0.48–1.13)
SEC	72/344	0.57 (0.38–0.84)#	0.59 (0.39–0.89)#	0.62 (0.41–0.95)#

* Results of crude and adjusted logistic regression analyses. 95% CI = 95% confidence interval; ADA = adalimumab; ASDAS = Ankylosing Spondylitis Disease Activity Score; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; CZP = certolizumab pegol; ETN = etanercept; GOL = golimumab; IFX = infliximab; OR = odds ratio; SEC = secukinumab.
† Only patients contributing a response measure (ASDAS or BASDAI) at 6 months were included.
‡ Adjusted by age and sex.
§ Adjusted by sex and baseline age, C-reactive protein level, Bath Ankylosing Spondylitis Disease Activity Index score, patient global score on a VAS, and concomitant conventional synthetic disease-modifying antirheumatic drugs.
¶ Model B, adding baseline comorbidity/extraarticular manifestations; see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24523/abstract, for details including patient numbers.
# Statistically significant.

It has been hypothesized that a change of mode of action from TNF inhibition to, for instance, IL-17 inhibition may be a favorable strategy in case of treatment failure, especially in patients showing lack of effect (5). To date, only a few minor observational studies (abstracts, monocenter, or <50–100 patients) (19,21–24) and no randomized trials have reported outcomes in SpA patients treated with secukinumab compared with a specific TNFi (25). A recent Swiss study reported comparable

1-year effectiveness of secukinumab versus TNFi in a TNFi-experienced real-life cohort, with results based on 106 patients treated with secukinumab, of whom 55 had available 1-year outcomes (12). The study used a comparison group comprising all TNFi combined. In that respect, it is of interest that we in the current cohort demonstrated considerable variation in effectiveness within the group of TNFi (13,14). Indeed, whereas comparisons of secukinumab versus combined TNFi outcomes were not statistically significant, results differed for secukinumab versus adalimumab, which for some comparisons were significant. It was beyond the scope of this study to further explore within-group variations between the different TNFi, but our results illustrate that pooling of TNFi may be too simplistic of an approach when it comes to disentangling the treatment strategy in SpA.

Our study included 2,725 patients in whom at least 2 TNFi had previously failed. Difficult-to-treat SpA patients in whom numerous biologics are failing are only sparsely described and usually not favored for inclusion in randomized comparative trials. Thus, little is known on how to approach this challenging patient group. The MEASURE 2 placebo-controlled study preceding the marketing of secukinumab included 72 patients with ankylosing spondylitis, of whom 39% had previously failed 1 TNFi (26). In the third-plus treatment group, we found that response rates were generally low (12–25% for ASDAS response at 6 months) regardless of drug. In this situation, patients treated with secukinumab had a 50% higher withdrawal rate and 40–45% lower odds for response than adalimumab. However, some important aspects of the study should be considered in the interpretation of our results.

Secukinumab was marketed later than the TNFi. It is possible that channeling occurred (i.e., patients in whom numerous TNFi failed ended up receiving secukinumab), leading to poorer outcomes. Thus, although baseline disease duration upon treatment start appeared similar across treatments, ≥3 prior TNFi had failed in 25% of patients treated with secukinumab. In the present study, all models were performed stratified by line of treatment. In the third-plus group, adding the number of previous biologics to the multivariate models did not change the results. Furthermore, there is always a risk of misdiagnosis or concomitant fibromyalgia in patients nonresponsive to therapy (27). We had no reason to suspect these challenges to be unevenly distributed in the secukinumab and adalimumab groups, but patients treated with secukinumab tended to have slightly higher pain and global scores at treatment start.

Comorbidities and extraarticular manifestations could have an impact due to risk of confounding. Psoriasis was more frequent in patients treated with secukinumab irrespective of treatment line, most likely because secukinumab was available for psoriasis before it was approved for SpA and due to favorable outcomes in psoriasis (28). Similarly, adalimumab was more frequently used among patients with prior uveitis and inflammatory bowel disease, manifestations for which the effectiveness of secukinumab is still unclear (29,30). Potential treatment decisions due to flares in these extraarticular manifestations are not uniformly captured in the registries and could potentially have affected the results. Although adjustment for these comorbidities in the multivariable analyses did not markedly change results, residual confounding remains a risk.

The observational design of the present study with the inclusion of patients from 5 different countries might have resulted in heterogeneous disease presentations. We have previously shown that disease presentation and the threshold for starting biologic treatment vary across the Nordic countries (15). Although adding country as a covariate in the multivariable analyses did not change the associations under study, residual confounding cannot be ruled out. Furthermore, neither TNFi nor secukinumab dosage was uniformly available, and the use of different doses (e.g., secukinumab 150 mg versus 300 SC every 4 weeks), including dose titration during follow-up, might have affected results.

This study has several strengths to consider, first and foremost, the high number of treatment courses registered prospectively in routine care and the subsequent linkage to national registries providing valid information on comorbidities. This provided us with the possibility of exploring prior extraarticular manifestations and other comorbidities as possible confounders and adjusting for these in multivariate analyses. It is a limitation that comorbidities exclusively diagnosed in primary care were not included; thus, mainly severe comorbid conditions were identified. Assessment of psoriasis-related disease activity (e.g., the Psoriasis Area and Severity Index), uveitis flares, and bowel symptoms is not routinely registered in the biologics registries contributing to this study. Thus, we could not evaluate whether flares of extraarticular manifestations affected the treatment strategy, including the decision to stop treatment early. Furthermore, enthesitis is not routinely registered. Impact of missing data was minimized by adding missingness as a separate category in the multivariable analyses but might still have affected results, although analyses from a complete case scenario provided very similar results. Reassuringly, the baseline characteristics of patients with and without 6-month outcome data were very similar, indicating no systematic differences.

In conclusion, secukinumab was mainly prescribed in biologics-experienced SpA patients in this study based on >10,000 treatment courses from 5 Nordic countries. Outcomes in difficult-to-treat patients in whom >2 prior biologics had failed were generally poor and slightly poorer for secukinumab versus adalimumab but similar for other TNFi. Our data did not suggest that secukinumab was superior to adalimumab or other TNFi after failure of a previous TNFi.

ACKNOWLEDGMENTS

The authors acknowledge patients and colleagues who contribute data to the DANBIO, SRQ, ROB-FIN, NOR-DMARD, and ICEBIO registries.

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. Glintborg 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. Glintborg, Lindström, Di Giuseppe, Gudbjornsson, Hetland, Nordström, Jørgensen, Hansen, Geirsson, Askling, Kristensen, Jacobsson.

Acquisition of data. Glintborg, Lindström, Di Giuseppe, Provan, Gudbjornsson, Hokkanen, Nordström, Grøn, Krogh, Kristensen, Jacobsson.

Analysis and interpretation of data. Glintborg, Lindström, Di Giuseppe, Gudbjornsson, Michelsen, Wallman, Aaltonen, Kristensen, Jacobsson.

ADDITIONAL DISCLOSURES

Author Krogh is an employee of Zitelab.

REFERENCES

1. Maxwell LJ, Zochling J, Boonen A, Singh JA, Veras MM, Tanjong GE, et al. TNF-alpha inhibitors for ankylosing spondylitis. Cochrane Database Syst Rev 2015:CD005468.
2. Baeten D, Baraliakos X, Braun J, Sieper J, Emery P, van der Heijde D, et al. Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. Lancet 2013;382:1705–13.
3. Baeten D, Sieper J, Braun J, Baraliakos X, Dougados M, Emery P, et al. Secukinumab, an interleukin-17A inhibitor, in ankylosing spondylitis. N Engl J Med 2015;373:2534–48.
4. Sieper J, Deodhar A, Marzo-Ortega H, Aelion JA, Blanco R, Jui-Cheng T, et al. Secukinumab efficacy in anti-TNF-naive and anti-TNF-experienced subjects with active ankylosing spondylitis: results from the MEASURE 2 study. Ann Rheum Dis 2017;76: 571–92.
5. Ward MM, Deodhar A, Gensler LS, Dubreuil M, Yu D, Khan MA, et al. 2019 update of the American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and Treatment Network recommendations for the treatment of ankylosing spondylitis and nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2019; 71:1599–613.
6. Purmonen T, Puolakka K, Mishra D, Gunda P, Martikainen J. Cost-effectiveness of secukinumab compared to other biologics in the treatment of ankylosing spondylitis in Finland. Clinicoecon Outcomes Res 2019;11:159–68.
7. Baraliakos X, Ostergaard M, Gensler LS, Poddubnyy D, Lee EY, Kiltz U, et al. Comparison of the effects of secukinumab and adalimumab biosimilar on radiographic progression in patients with ankylosing spondylitis: design of a randomized, phase IIIb study (SURPASS). Clin Drug Investig 2020;40:269–78.
8. Maksymowych WP, Strand V, Nash P, Yazici Y, Thom H, Hunger M, et al. Comparative effectiveness of secukinumab and adalimumab in ankylosing spondylitis as assessed by matching-adjusted indirect comparison. Eur J Rheumatol 2018;5:216–23.
9. Vashisht P, Sayles H, Cannella AC, Mikuls TR, Michaud K. Generalizability of patients with rheumatoid arthritis in biologic agent clinical trials. Arthritis Care Res (Hoboken) 2016;68:1478–88.
10. Yoshida K, Radner H, Kavanaugh A, Sung YK, Bae SC, Kishimoto M, et al. Use of data from multiple registries in studying biologic discontinuation: challenges and opportunities. Clin Exp Rheumatol 2013;31 Suppl 78:S28–32.
11. Askling J, Fored CM, Geborek P, Jacobsson LT, van VR, Feltelius N, et al. Swedish registers to examine drug safety and clinical issues in RA. Ann Rheum Dis 2006;65:707–12.
12. Micheroli R, Tellenbach C, Scherer A, Burki K, Niederman K, Nissen MJ, et al. Effectiveness of secukinumab versus an alternative TNF inhibitor in patients with axial spondyloarthritis previously exposed to TNF inhibitors in the Swiss Clinical Quality Management cohort. Ann Rheum Dis 2020;79:1203–9.
13. Chen C, Zhang X, Xiao L, Zhang X, Ma X. Comparative effectiveness of biologic therapy regimens for ankylosing spondylitis: a systematic review and a network meta-analysis. Medicine (Baltimore) 2016;95: e3060.
14. Heinonen AV, Aaltonen KJ, Joensuu JT, Lahteenmaki JP, Pertovaara MI, Romu MK, et al. Effectiveness and drug survival of TNF inhibitors in the treatment of ankylosing spondylitis: a prospective cohort study. J Rheumatol 2015;42:2339–46.
15. Glintborg B, Lindstrom U, Aaltonen K, Kristianslund EK, Gudbjornsson B, Chatzidionysiou K, et al. Biological treatment in ankylosing spondylitis in the Nordic countries during 2010–2016: a collaboration between five biological registries. Scand J Rheumatol 2018;1–10.
16. Chatzidionysiou K, Hetland ML, Frisell T, Di GD, Hellgren K, Glintborg B, et al. Opportunities and challenges for real-world studies on chronic inflammatory joint diseases through data enrichment and collaboration between national registers: the Nordic example. RMD Open 2018;4:e000655.
17. Lindstrom U, Glintborg B, Di GD, Nordstrom D, Aarrestad PS, Gudbjornsson B, et al. Treatment retention of infliximab and etanercept originators versus their corresponding biosimilars: Nordic collaborative observational study of 2334 biologics naive patients with spondyloarthritis. RMD Open 2019;5:e001079.
18. Kristensen LE, Saxne T, Geborek P. The LUNDEX, a new index of drug efficacy in clinical practice: results of a five-year observational study of treatment with infliximab and etanercept among rheumatoid arthritis patients in southern Sweden. Arthritis Rheum 2006;54: 600–6.
19. Michelsen B, Askling J, Codreanu C, Mann H, Loft A, Pombo-Suarez M, et al. 6 and 12-month drug retention rates and treatment outcomes in 941 patients with axial spondyloarthritis treated with secukinumab in routine clinical practice in 12 European Countries in the EuroSpA Research Collaboration Network [abstract]. Arthritis Rheumatol 2019;71 Suppl 10. URL: https://acrabstracts.org/abstract/6-and-12-month-drug-retention-rates-and-treatment-outcomes-in-941-patients-with-axial-spondyloarthritis-treated-with-secukinumab-in-routine-clinical-practice-in-12-european-countries-in-the-eurospa-re/.
20. Van der Heijde D, Ramiro S, Landewe R, Baraliakos X, van den Bosch F, Sepriano A, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis 2017; 76:978–91.
21. Elliott A, Wright G. Real-world data on secukinumab use for psoriatic arthritis and ankylosing spondylitis. Ther Adv Musculoskelet Dis 2019;11:1759720X19858510.
22. Williams T, Wadeley A, Bond D, Cavill C, Freeth M, Sengupta R. Real-world experience of secukinumab treatment for ankylosing spondylitis at the Royal National Hospital for Rheumatic Diseases, Bath. Clin Rheumatol 2020;39:1501–4.
23. Gentileschi S, Vitale A, Rigante D, Lopalco G, Emmi G, Orlando I, et al. Prompt clinical response to secukinumab in patients with axial spondyloarthritis: real life observational data from three Italian referral centers. Isr Med Assoc J 2018;20:438–41.
24. Joshi R, Latremouille-Viau D, Meiselbach MK, Xie J, Park Y, Sunkureddi P. Characterization of patients with ankylosing spondylitis receiving secukinumab and reasons for initiating treatment: a US

physician survey and retrospective medical chart review. Drugs Real World Outcomes 2019;6:1–9.

25. Blair HA. Secukinumab: a review in ankylosing spondylitis. Drugs 2019;79:433–43.
26. Deodhar A, Conaghan PG, Kvien TK, Strand V, Sherif B, Porter B, et al. Secukinumab provides rapid and persistent relief in pain and fatigue symptoms in patients with ankylosing spondylitis irrespective of baseline C-reactive protein levels or prior tumour necrosis factor inhibitor therapy: 2-year data from the MEASURE 2 study. Clin Exp Rheumatol 2019;37:260–9.
27. Macfarlane GJ, Barnish MS, Pathan E, Martin KR, Haywood KL, Siebert S, et al. Co-occurrence and characteristics of patients with axial spondyloarthritis who meet criteria for fibromyalgia: results from a UK national register. Arthritis Rheumatol 2017;69:2144–50.
28. Yang EJ, Beck KM, Liao W. Secukinumab in the treatment of psoriasis: patient selection and perspectives. Psoriasis (Auckl) 2018;8:75–82.
29. Lindstrom U, Bengtsson C, Olofsson T, Di Giuseppe D, Glintborg B, Forsblad-d'Elia H, et al. Comparison of anterior uveitis occurrence during treatment with secukinumab, adalimumab, infliximab and etanercept in spondyloarthritis [abstract]. Ann Rheum Dis 2020;79 Suppl 1:9.
30. Schreiber S, Colombel JF, Feagan BG, Reich K, Deodhar AA, McInnes IB, et al. Incidence rates of inflammatory bowel disease in patients with psoriasis, psoriatic arthritis and ankylosing spondylitis treated with secukinumab: a retrospective analysis of pooled data from 21 clinical trials. Ann Rheum Dis 2019;78:473–9.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 759–767
DOI 10.1002/acr.24517

American College of Rheumatology
Empowering Rheumatology Professionals

Secukinumab and Sustained Reduction in Fatigue in Patients With Ankylosing Spondylitis: Long-Term Results of Two Phase III Randomized Controlled Trials

Tore K. Kvien,[1] Philip G. Conaghan,[2] Laure Gossec,[3] Vibeke Strand,[4] Mikkel Østergaard,[5] Denis Poddubnyy,[6] Nicole Williams,[7] Brian Porter,[8] Abhijit Shete,[9] Isabelle Gilloteau,[9] and Atul Deodhar[10]

Objective. To investigate the longer-term effects of secukinumab 150 mg on fatigue in patients with ankylosing spondylitis (AS) in the MEASURE 1 study (up to 3 years) and the MEASURE 2 study (up to 2 years).

Methods. Patients with active AS were randomized to secukinumab or placebo in MEASURE 1 (10 mg/kg intravenous [IV] followed by 150 mg subcutaneous) and MEASURE 2 (150 mg subcutaneous). Patients were naive to treatment with anti–tumor necrosis factor (anti-TNF-naive) therapy or had an inadequate response/intolerance to anti-TNF therapy (anti-TNF-IR). Fatigue was measured using the Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) scale. Relationships between fatigue response and baseline characteristics and clinical/laboratory variables were explored.

Results. Significant improvements in FACIT-F scores from baseline were observed with secukinumab across both studies versus placebo at week 16 ($P < 0.05$). Improvements were sustained through week 156 (MEASURE 1) and week 104 (MEASURE 2). Significantly more patients reported fatigue responses (FACIT-F improvement ≥4; observed data) with secukinumab 150 mg than with placebo at week 16 in both MEASURE 1 ($P < 0.05$) and MEASURE 2 ($P < 0.01$). Fatigue responses were achieved by 75.6% of patients receiving secukinumab at week 156 (MEASURE 1) and 81.4% at week 104 (MEASURE 2); these results were consistent in patients who were anti-TNF-naive (74.3% and 84.6%, respectively) and anti-TNF-IR (81.3% and 75.0%, respectively). Baseline characteristics did not predict improvement in fatigue consistently. Fatigue responses were moderately to strongly correlated with responses in several clinical measures, including the Assessment of SpondyloArthritis international Society (ASAS) 20%/40% improvement, ASAS5/6 responses, the Ankylosing Spondylitis Disease Activity Score with C-reactive protein level, the Bath Ankylosing Spondylitis Disease Activity Index, and the Short Form 36 health questionnaire scores.

Conclusion. Secukinumab provided rapid and sustained improvements in fatigue for up to 3 years, regardless of prior anti-TNF exposure.

INTRODUCTION

Ankylosing spondylitis (AS) is a chronic inflammatory condition predominantly affecting the spine and sacroiliac joints, which, if left untreated, can lead to progressive structural and functional impairment (1).

The main goals of AS therapy are to maximize long-term health-related quality of life (HRQoL) "through control of symptoms and inflammation, prevention of progressive structural damage, and preservation/normalization of function and social participation" (2). Fatigue is reported in up to 66% of patients with AS (3) and has been identified as a key patient priority in the

ClinicalTrials.gov identifiers: NCT01358175, NCT01863732, NCT01649375.
The views expressed herein are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.
Supported by Novartis Pharma AG, Basel, Switzerland. Dr. Conaghan's work was supported by the UK NIHR Leeds Biomedical Research Centre.

[1]Tore K. Kvien, MD, PhD: Diakonhjemmet Hospital and University of Oslo, Oslo, Norway; [2]Philip G. Conaghan, MD, PhD: University of Leeds and NIHR Leeds Biomedical Research Centre, Leeds, UK; [3]Laure Gossec, MD, PhD: Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique and Pitié Salpêtrière Hospital, APHP, Paris, France; [4]Vibeke Strand, MD: Stanford University School of Medicine, Palo Alto, California; [5]Mikkel Østergaard, MD, PhD, DMSc: Rigshospitalet, Glostrup, and University of Copenhagen, Copenhagen, Denmark; [6]Denis Poddubnyy, MD, PhD: Campus Benjamin Franklin, Charité-Universitätsmedizin Berlin and German Rheumatism Research Centre, Berlin, Germany; [7]Nicole Williams, BSc: RTI Health Solutions, Durham, North Carolina; [8]Brian Porter, MD, PhD, MPH, MBA: Novartis Pharmaceuticals Corporation, East Hanover, New Jersey; [9]Abhijit Shete, MD, Isabelle Gilloteau, MSc: Novartis Pharma AG, Basel, Switzerland; [10]Atul Deodhar, MD: Oregon Health and Science University, Portland.

Dr. Kvien has received consulting and/or speaking fees from AbbVie, Biogen, Bristol Myers Squibb, Boehringer Ingelheim, Celgene, Celltrion, Eli Lilly and Company, Epirus, Hospira, Merck-Serono, MSD, Mundipharma, Novartis,

SIGNIFICANCE & INNOVATIONS

- In this exploratory analysis of 2 phase III randomized placebo-controlled trials (MEASURE 1 and MEASURE 2), treatment with secukinumab provided rapid improvements in fatigue (assessed using the Functional Assessment of Chronic Illness Therapy–Fatigue scale) that were sustained for up to 3 years in patients with active ankylosing spondylitis (AS); improvements were particularly prominent in patients who were naive to anti–tumor necrosis factor therapy.
- Fatigue responses correlated with responses in several other clinical measures, including signs and symptoms (Assessment of SpondyloArthritis international Society [ASAS] 20% improvement, ASAS 40% improvement, and ASAS5/6 responses), disease activity (Ankylosing Spondylitis Disease Activity Score with C-reactive protein level and Bath Ankylosing Spondylitis Disease Activity Index), and health-related quality of life (HRQoL; Short Form 36 health questionnaire scores), highlighting the close links between fatigue and other key aspects of disease.
- Efficacy in treating fatigue adds to the existing evidence base for secukinumab in AS, which already includes significant and sustained improvements in signs and symptoms, physical function, and HRQoL.

treatment of AS (4). Many patients report that fatigue negatively impacts HRQoL and social functioning (5–7); thus, reducing this symptom remains an important unmet need in AS.

Interleukin-17A (IL-17A) is one of the key cytokines driving the pathogenesis of AS (8). Secukinumab, a fully human monoclonal antibody that selectively binds to and inhibits IL-17A (9), is approved for the treatment of AS (10) based on the results of 2 randomized, double-blind, placebo-controlled, phase III trials, MEASURE 1 (NCT01358175 and NCT01863732) and MEASURE 2 (NCT01649375), that demonstrated significant improvements in the signs and symptoms of AS with secukinumab versus placebo (11).

Secukinumab has previously been shown to provide rapid improvements in fatigue (12). The aim of the current analysis of data from the MEASURE 1 and MEASURE 2 studies was to investigate the longer-term effects of the approved dose of secukinumab (150 mg) on fatigue in patients with AS who were naive to tumor necrosis factor (TNF) inhibitor therapy (anti-TNF-naive) or who had a previous inadequate response to or intolerance of TNF inhibitors (anti-TNF-IR). Potential predictors of fatigue response and the relationship between fatigue and other clinical response measures were also assessed.

PATIENTS AND METHODS

Participants. Detailed enrollment criteria and trial designs for MEASURE 1 and MEASURE 2 have been previously described (11). Inclusion criteria were the same for both trials. Briefly, patients were enrolled if they were age ≥18 years, had AS as defined by the modified New York criteria, had an inadequate response to or intolerance of nonsteroidal antiinflammatory drugs, and either had no previous treatment with anti-TNF therapy or an inadequate response/intolerance to not >1 anti-TNF agent. Exclusion criteria included total spinal ankylosis and active and ongoing systemic infections or inflammatory conditions, other than AS.

Trial design. Patients were randomized to secukinumab 75 mg, 150 mg, or placebo in both trials (intravenous loading, followed by subcutaneous maintenance dosing in MEASURE 1, and subcutaneous loading and maintenance dosing in MEASURE 2). At week 16 (or week 24 in MEASURE 1, depending on Assessment of SpondyloArthritis international Society [ASAS] 20% improvement response [ASAS20]), placebo patients were re-randomized to secukinumab. At week 104 of MEASURE 1, patients were invited to enter a long-term extension study for up to 3 additional years, continuing on the same treatment. Efficacy data from patients receiving secukinumab 75 mg are not presented in this article, as this dose is not approved.

Oktal, Orion Pharma, Hospira/Pfizer, Roche, Sandoz, and UCB (less than $10,000 each) and research grants to Diakonhjemmet Hospital from AbbVie, Bristol Myers Squibb, MSD, Pfizer, Roche, and UCB. Dr. Conaghan has received consulting and/or speaking fees from AbbVie, Bristol Myers Squibb, GlaxoSmithKline, Merck, Novartis, Pfizer, and Roche (less than $10,000 each) and research grants from Bristol Myers Squibb. Dr. Gossec has received consulting fees from AbbVie, Amgen, Biogen, Bristol Myers Squibb, Celgene, Janssen, Novartis, Pfizer, Roche, Sandoz, and UCB (less than $10,000 each) and research grants from UCB, Eli Lilly and Company, Pfizer, and Bristol-Myers Squibb. Dr. Strand has received consulting fees from AbbVie, Amgen, Boehringer Ingelheim, Bristol Myers Squibb, Celgene, Celltrion, CORRONA, Genentech/Roche, GlaxoSmithKline, Eli Lilly and Company, Janssen, Merck, Novartis, Pfizer, Regeneron, Samsung, Sandoz, Sanofi, and UCB (less than $10,000 each). Dr. Østergaard has received consulting fees, speaking fees, and/or travel expenses from AbbVie, Bristol Myers Squibb, Boehringer Ingelheim, Celgene, Eli Lilly and Company, Hospira, Janssen, Merck, Novartis, Novo Nordisk, Orion, Pfizer, Regeneron, Roche, and UCB (less than $10,000 each) and research grants from AbbVie, Celgene, Centocor, Merck, and Novartis. Dr. Poddubnyy has received consulting and/or speaking fees from AbbVie, Bristol Myers Squibb, Celgene, Eli Lilly and Company, MSD, Novartis, Pfizer, Roche, and UCB (less than $10,000 each) and research grants from AbbVie, MSD, Novartis, and Pfizer. Drs. Porter and Shete and Ms. Gilloteau are employees of Novartis and own Novartis stock. Dr. Deodhar has received speaking fees, consulting fees, and/or travel expenses from AbbVie, Amgen, Boehringer Ingelheim, Bristol Myers Squibb, Eli Lilly and Company, GlaxoSmithKline, Janssen, Novartis, Pfizer, and UCB (less than $10,000 each) and research grants from Bristol Myers Squibb, Eli Lilly and Company, GlaxoSmithKline, Janssen, Novartis, Pfizer, and UCB. No other disclosures relevant to this article were reported.

Address correspondence to Tore K. Kvien, MD, PhD, Department of Rheumatology, Diakonhjemmet Hospital, N-0319, Oslo, Norway. Email: t.k.kvien@medisin.uio.no.

Submitted for publication October 11, 2019; accepted in revised form November 19, 2020.

The studies were conducted in accordance with the Declaration of Helsinki (13). The institutional review board or independent ethics committee at each participating center approved the protocols. Written and informed consent was obtained for all patients.

Outcome measures. Improvements in fatigue, an exploratory endpoint of the MEASURE 1 and 2 trials, were evaluated using the Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) scale, a 13-item questionnaire that assesses self-reported fatigue and its impact on daily activities and function on a scale of 0–52, with higher scores indicating less fatigue (14). Assessments were performed at baseline and at weeks 4, 8, 12, 16, 24, 52, and 104 in both trials, and at week 156 in the MEASURE 1 extension. A fatigue response was defined as an improvement (increase) of ≥4 points in the FACIT-F score, corresponding to the minimum clinically important difference (14). Possible relationships between baseline characteristics and fatigue responses were explored using pooled data from MEASURE 1 and MEASURE 2. Prespecified subgroup analyses on the basis of anti-TNF response status (anti-TNF-naive or anti-TNF-IR) were performed for FACIT-F scores at weeks 16, 24, 52, and 104 for both trials, and at week 156 for the MEASURE 1 extension.

Statistical methods. Statistical analyses were performed using the software package SAS, version 9.4. Detailed sample size calculations have been previously reported (11). The mean change in FACIT-F score from baseline up to week 104 or 156 was assessed using a mixed model for repeated measures (MMRM) approach. Treatment regimen, analysis visit, and randomization stratum (anti-TNF-naive or anti-TNF-IR) were included as factors in the model, with weight and baseline score as continuous covariates, and treatment by analysis visit and baseline score by analysis visit as interaction terms, as well as an unstructured covariance structure. The proportions of patients with a fatigue response are presented as observed and were compared using Fisher's exact test.

A multivariate logistic regression model was used to explore the association between fatigue responses (dependent variable) and selected demographic and clinical baseline independent variables (age, sex, TNF inhibitor status, FACIT-F score, Bath Ankylosing Spondylitis Disease Activity Index [BASDAI] total score, Bath Ankylosing Spondylitis Functional Index [BASFI] score, swollen/tender 44-joint count, patient global assessment of disease [visual analog scale (VAS)], spinal pain assessment, C-reactive protein [CRP] level, Ankylosing Spondylitis Disease Activity Score [ASDAS] with CRP level, and disease duration) at weeks 16, 52, and 104. Univariate linear regression models were also conducted for absolute FACIT-F scores at week 16 and week 104, with response variable and selected baseline factors as predictors.

Possible correlations between dichotomized clinical response measures and fatigue response at week 16 (the primary endpoint for clinical responses) and at weeks 52 and 104 were examined using polychoric (used to calculate the correlation coefficients between ordinal variables) and polyserial (used for an ordinal discrete variable and a continuous variable) correlation coefficients. The following were evaluated: ASAS20 and ASAS40 response, ASAS5/6 response, ASAS partial remission, ASDAS-CRP major improvement, BASDAI 50, Work Productivity and Activity Impairment (WPAI) change from baseline in percentage of overall work impairment due to health, WPAI change from baseline in

Table 1. Demographic and baseline characteristics of patients from MEASURE 1 study and MEASURE 2 study included in the analysis*

Characteristics	MEASURE 1		MEASURE 2	
	SEC IV, 150 mg (n = 125)	Placebo (n = 122)	SEC, 150 mg (n = 72)	Placebo (n = 74)
Age, years	40.1 ± 11.6	43.1 ± 12.4	41.9 ± 12.5	43.6 ± 13.2
Female, no. (%)	41 (32.8)	37 (30.3)	26 (36.1)	18 (24.3)
Weight, kg	74.7 ± 16.2	76.7 ± 14.4	82.3 ± 18.0	80.3 ± 15.2
Time since AS diagnosis, years	6.5 ± 6.9	8.3 ± 8.9	7.0 ± 8.2	6.4 ± 8.9
Anti-TNF-naive, no. (%)	92 (73.6)	89 (73.0)	44 (61.1)	45 (60.8)
MTX use at randomization, no. (%)	17 (13.6)	16 (13.1)	8 (11.1)	9 (12.2)
hsCRP, median (range) mg/liter	7.4 (0.2–147.7)	7.9 (0.2–146.8)	7.5 (0.4–237.0)	8.3 (0.5–84.6)
PtGA VAS (0–100 mm)	64.0 ± 19.4	66.3 ± 18.6	67.5 ± 16.8	70.5 ± 15.8
BASDAI total score	6.4 ± 1.6	6.5 ± 1.5	6.6 ± 1.5	6.8 ± 1.3
BASFI score	5.6 ± 2.2	5.8 ± 2.0	6.2 ± 2.1	6.1 ± 2.0
Swollen 44-joint count at baseline	2.3 ± 4.4	2.3 ± 4.1	1.6 ± 3.3	2.0 ± 5.3
Tender 44-joint count at baseline	5.5 ± 7.7	6.5 ± 8.4	5.5 ± 7.6	4.8 ± 7.7
Subject spinal pain assessment	64.0 ± 18.6	66.7 ± 16.5	66.2 ± 16.7	69.2 ± 18.8
ASDAS-CRP	3.6 ± 0.9	3.7 ± 0.9	3.7 ± 0.9	3.8 ± 0.8
FACIT-F score	25.6 ± 10.7	24.5 ± 9.4	22.6 ± 8.8	24.3 ± 9.0

* Values are the mean ± SD unless indicated otherwise. Anti-TNF = anti–tumor necrosis factor; AS = ankylosing spondylitis; ASDAS-CRP = Ankylosing Spondylitis Disease Activity Score with C-reactive protein; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; FACIT-F = Fatigue Functional Assessment of Chronic Illness Therapy–Fatigue; hsCRP = high-sensitivity CRP; IV = intravenous; MTX = methotrexate; PtGA = patient global assessment; SEC = secukinumab; TNF = tumor necrosis factor; VAS = visual analog scale.

percentage of impairment while working due to health, hemoglobin, patient's total back pain intensity VAS, and the Short Form 36 (SF-36) health questionnaire mental component score (MCS) and physical component score (PCS). FACIT-F scores were stratified by selected clinical response measures. Degrees of association between clinical response measures were determined based on Cohen's criteria for correlation coefficients, in which 0.1–0.29 indicates a weak association, 0.3–0.49 indicates a moderate association, and ≥0.5 indicates a strong association. Analyses of fatigue are presented separately from MEASURE 1 (up to 104 weeks), MEASURE 1 extension (156 weeks), and MEASURE 2 (up to 104 weeks) for the approved dose of secukinumab (150 mg).

RESULTS

Baseline characteristics. In MEASURE 1, of the 125 patients randomized to secukinumab 150 mg at baseline, 97 (77.6%) completed the 2-year core trial; 87 of these patients entered the extension study, of which 83 patients (95.4%) completed week 156. In MEASURE 2, of the 72 patients randomized to secukinumab 150 mg at baseline, 60 patients (83.3%) completed week 104. Patient demographics and baseline characteristics were comparable across groups (Table 1). The majority of patients were male (~70%), and ~70% were anti-TNF-naive and 30% were anti-TNF-IR across both trials. Mean FACIT-F scores were low at baseline in both studies (range 22.6–25.6 across MEASURE 1 and MEASURE 2), indicating severe fatigue.

Fatigue: overall populations. In this analysis, the least mean squares improvements in FACIT-F total scores from baseline were significantly higher with secukinumab 150 mg compared with placebo at week 16 in both MEASURE 1 and MEASURE 2 (MMRM: $P < 0.001$ [MEASURE 1] and $P < 0.01$ [MEASURE 2]) for secukinumab 150 mg versus placebo (Figure 1). Signficant

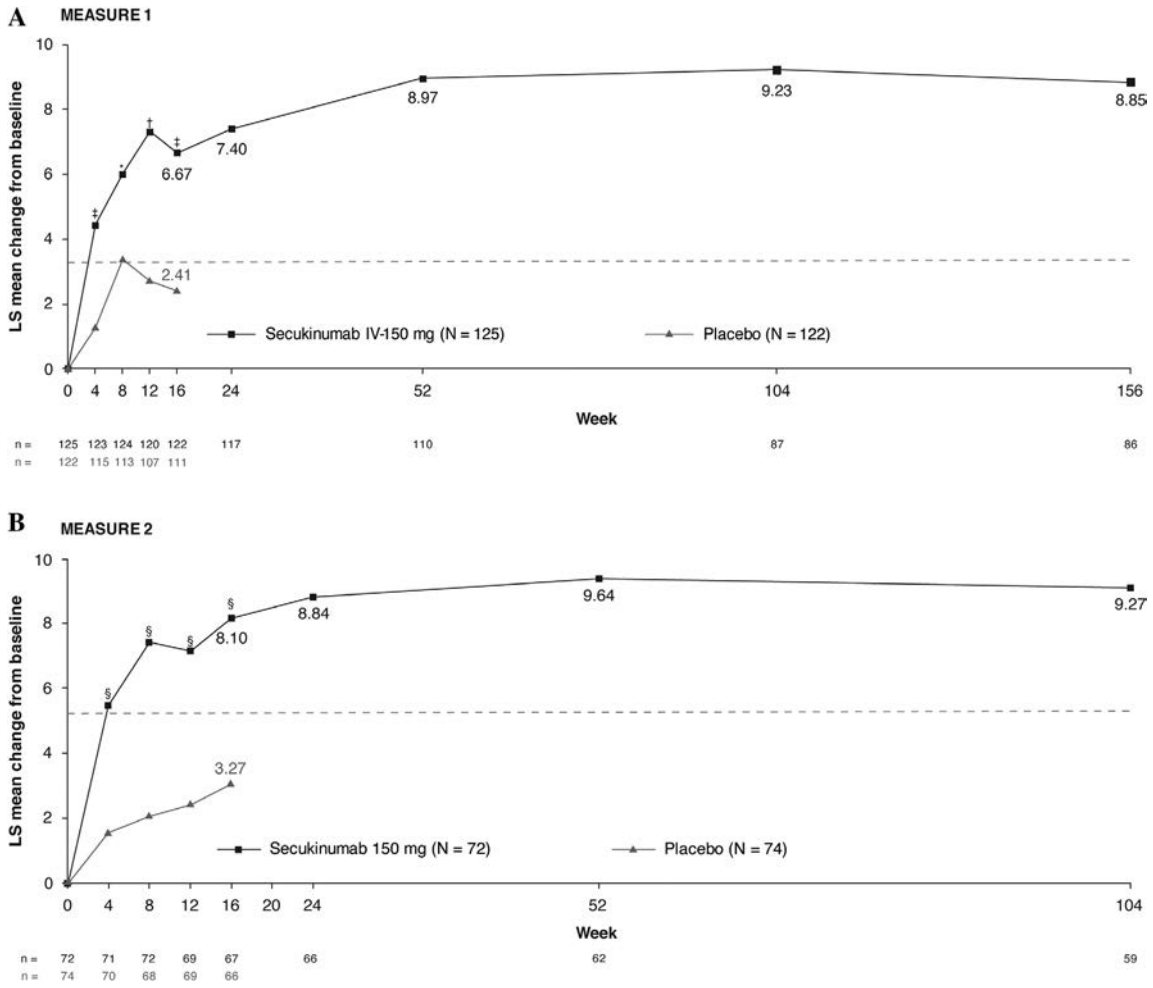

Figure 1. Least mean squares (LS) change from baseline in Functional Assessment of Chronic Illness Therapy–Fatigue scores in overall populations of **A**, MEASURE 1, and **B**, MEASURE 2 (mixed model repeated-measures analysis). Broken lines indicate minimum clinically important difference. * = $P < 0.05$; † = $P < 0.0001$; ‡ = $P < 0.001$; § = $P < 0.01$ versus placebo. IV = intravenous; N = number of randomized patients; n = number of patients in treatment group with evaluation.

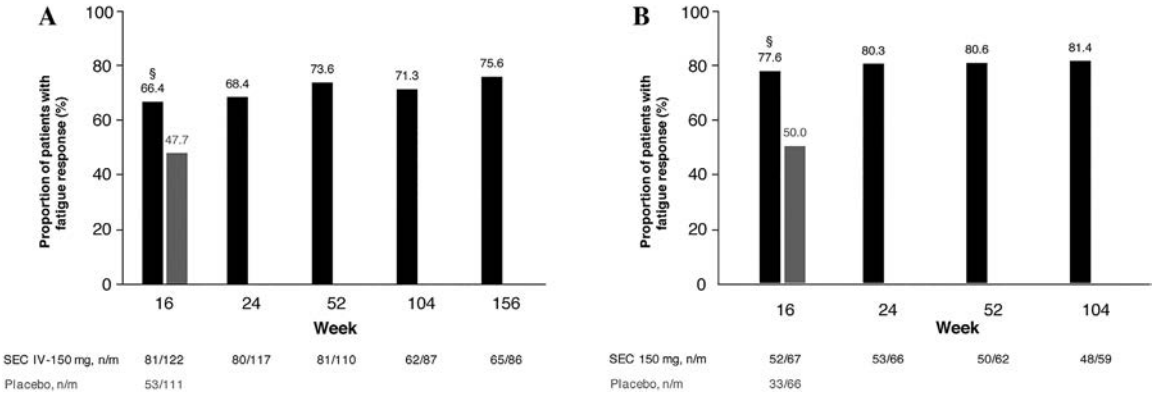

Figure 2. Fatigue responses in the overall populations of **A**, MEASURE 1, and **B**, MEASURE 2 (data shown as observed). Fatigue response was defined as an increase from baseline of ≥4 points in the Functional Assessment of Chronic Illness Therapy–Fatigue score. § = $P < 0.01$ versus placebo. IV = intravenous; m = total number of patients in treatment group with evaluation; n = number of patients who were responders; SEC = secukinumab.

improvements from baseline in FACIT-F total scores were reported as early as week 4 (the first postbaseline assessment) with secukinumab 150 mg in both studies. Improvements were sustained up to week 156 in the MEASURE 1 extension and week 104 in MEASURE 2 (observed data) (Figure 1).

Significantly more patients reported fatigue responses with secukinumab 150 mg than placebo at week 16 in both trials (observed data; both $P < 0.01$ in MEASURE 1 and MEASURE 2) (Figure 2). Fatigue responses were sustained in 75.6% of patients through week 156 in the MEASURE 1 extension and 81.4% through week 104 in MEASURE 2 (observed data) (Figure 2).

Fatigue: anti-TNF-naive and anti-TNF-IR patients. Improvements in FACIT-F scores were reported with

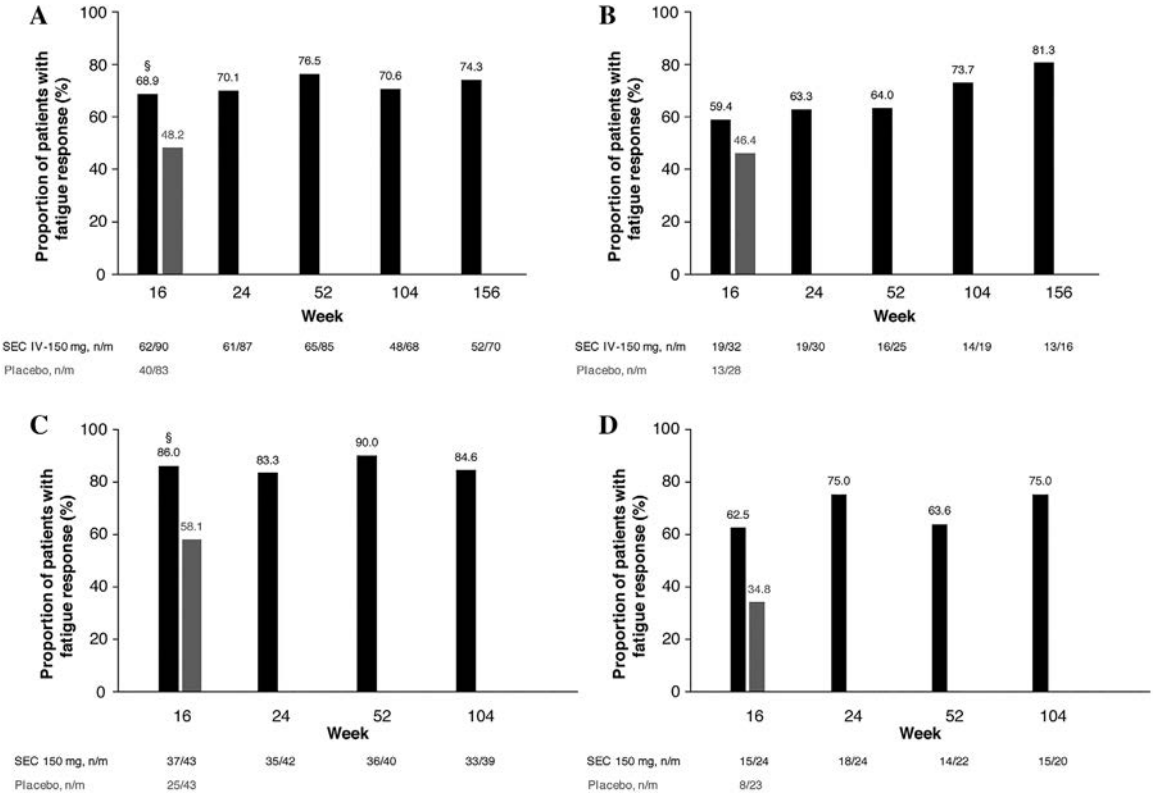

Figure 3. Fatigue responses over time. **A**, Anti–tumor necrosis factor (anti-TNF) naive, and **B**, Anti-TNF inadequate responder (IR) patients in MEASURE 1; **C**, Anti-TNF-naive, and **D**, Anti-TNF-IR patients in MEASURE 2 (data shown as observed). Fatigue response was defined as an increase from baseline of ≥4 points in the Functional Assessment of Chronic Illness Therapy–Fatigue score. § = $P < 0.01$ versus placebo. IV = intravenous; m = total number of patients in treatment group with evaluation; n = number of patients who were responders; SEC = secukinumab.

secukinumab treatment in both MEASURE 1 and MEASURE 2, regardless of anti-TNF treatment status (observed data; see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24517/abstract). In MEASURE 1, least mean squares improvements from baseline in FACIT-F scores were reported with secukinumab 150 mg at week 16 in anti-TNF-naive patients (MMRM: 8.42 versus 2.95 with placebo; $P < 0.0001$) and anti-TNF-IR patients (MMRM: 3.61 versus 2.57 with placebo; $P = 0.68$); however, these improvements only reached statistical significance in the anti-TNF-naive group. Observed improvements were sustained or further increased up to week 156 (MMRM: 9.57 and 9.28 in anti-TNF-naive and anti-TNF-IR patients, respectively). Similarly, in MEASURE 2, improvements in least mean squares FACIT-F scores from baseline were evident with secukinumab in both anti-TNF-naive (MMRM: week 16: 9.96 versus 5.15 with placebo [$P = 0.01$]; week 104: 9.85) and anti-TNF-IR patients (MMRM: week 16: 5.66 versus 0.54 with placebo [$P = 0.0560$]; week 104: 9.13), although they only reached statistical significance in the anti-TNF-naive group.

At week 16 in both MEASURE 1 and MEASURE 2, the proportions of anti-TNF-naive patients reporting fatigue responses were significantly greater with secukinumab versus placebo (observed data; $P < 0.01$); more anti-TNF-IR patients receiving secukinumab reported fatigue responses at week 16 versus placebo, although differences were not statistically significant (observed data) (Figure 3). The proportions of patients reporting fatigue responses were sustained in both studies through weeks 104 and 156 (observed data) (Figure 3). Fatigue responses were numerically higher in anti-TNF-naive patients than anti-TNF-IR patients, particularly at week 16 in the MEASURE 2 study.

Association between other variables and fatigue response. *Baseline predictors of fatigue response.* Results of univariate linear regression models demonstrated that several baseline factors, including age, sex, BASDAI score, inadequate responses to anti-TNF, and SF-36 PCS and MCS scores were predictive of absolute FACIT-F scores at week 16 (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24517/abstract) and week 104 (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24517/abstract).

In multivariate logistic regression analysis, although no baseline or disease characteristics predicted improvement in fatigue across all time points assessed, several variables were noted to be predictive at either week 16 (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24517/abstract) or week 104 (Table 2). At week 16, older patients had significantly lower odds of fatigue responses and patients with higher baseline ASDAS-CRP had significantly increased odds of fatigue responses ($P < 0.05$ for both, see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24517/abstract). No baseline or disease characteristics were significant predictors for fatigue responses at week 52 (data not shown). BASFI and FACIT-F scores at baseline were identified as significant predictive factors for achieving fatigue responses at week 104; no other predictors were identified (Table 2). Anti-TNF-IR

Table 2. Multivariate logistic regression analysis examining baseline predictors of FACIT-F response at week 104 in a pooled population from MEASURE 1 study and MEASURE 2 study*

Baseline variable	Odds ratio (95% CI)	*P*
Age (est. odds ratio for a 10-unit increase)	0.882 (0.711–1.094)	0.252
Female	0.603 (0.336–1.081)	0.089
Anti-TNF-IR	1.007 (0.552–1.834)	0.983
FACIT-F score at baseline	0.897 (0.864–0.931)	<0.0001†
BASDAI total score at baseline	1.111 (0.828–1.492)	0.482
BASFI score at baseline	0.752 (0.631–0.895)	0.0014†
Swollen 44-joint count at baseline	0.977 (0.889–1.075)	0.637
Tender 44-joint count at baseline	0.998 (0.953–1.045)	0.927
Patient global assessment VAS at baseline	0.992 (0.972–1.014)	0.479
Patient spinal pain assessment at baseline	1.008 (0.987–1.029)	0.449
CRP, mg/liter at baseline	1.001 (0.984–1.018)	0.911
ASDAS-CRP at baseline	1.690 (0.919–3.110)	0.092
Disease duration, time since diagnosis	1.007 (0.975–1.040)	0.682

* *P* values are from a multivariate logistic regression model with baseline factors including age, sex, tumor necrosis factor (TNF) inhibitor status, Fatigue Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) score at baseline, Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) total score, Bath Ankylosing Spondylitis Functional Index (BASFI) score, swollen 44-joint count, tender 44-joint count, patient global assessment (visual analog scale [VAS]), patient spinal pain assessment, C-reactive protein (CRP) level, Ankylosing Spondylitis Disease Activity Score with CRP (ASDAS-CRP), and disease duration. Patients randomized to secukinumab 75 mg or 150 mg or placebo are included. The model includes N = 402 observations. Fatigue response was defined as an increase from baseline of ≥4 points in the FACIT-F score. Parameter estimates and odds ratios apply to a 1-unit increase in all continuous variables unless otherwise specified. 95% CI = 95% confidence interval; Anti-TNF-IR = anti-TNF inadequate responder.
† *P* values < 0.05.

Table 3. Correlations between fatigue responses at weeks 16, 52, and 104 and measures of clinical response at the corresponding time points in a pooled population from MEASURE 1 study and MEASURE 2 study*

Measure of clinical response	Week 16	Week 52	Week 104
ASAS20 response	0.66	0.62	0.70
ASAS40 response	0.71	0.57	0.60
ASAS5/6 response	0.62	0.52	0.58
ASAS partial remission	0.44	0.46	0.33
ASDAS-CRP major improvement	0.56	0.60	0.49
BASDAI 50	0.55	0.52	0.52
WPAI change from baseline in percentage of overall work impairment due to health	–0.47	–0.50	–0.51
WPAI change from baseline in percentage of impairment while working due to health	–0.49	–0.50	–0.55
Hemoglobin, gm/liter	–0.04†	0.02†	–0.08†
Patient's total back pain intensity VAS	0.47	0.41	0.38
SF-36 MCS	–0.41	–0.35	–0.36
SF-36 PCS	–0.42	–0.42	–0.35

* Values are the correlation coefficient. All $P < 0.01$ except for hemoglobin (P values calculated by the chi-square likelihood ratio test). Fatigue response was defined as an increase from baseline of ≥4 points in the Fatigue Functional Assessment of Chronic Illness Therapy–Fatigue score. Polychoric correlations (used to calculate the correlation coefficients between ordinal variables) were calculated for ASAS20, ASAS40, ASAS5/6, and ASAS partial remission, Ankylosing Spondylitis Disease Activity Score with C-reactive protein ≥2.0 (ASDAS-CRP) major improvement, and 50% improvement in Bath Ankylosing Spondylitis Disease Activity Index (BASDAI 50) score. Polyserial correlations (used to calculate the correlation coefficients between an ordinal discrete variable and a continuous variable) were calculated for Work Productivity and Activity Impairment (WPAI) percentage of overall work impairment due to health and WPAI percentage of impairment while working due to health. Degrees of association between variables were determined based on Cohen's criteria for correlation coefficients (0.1–0.29 = weak association, 0.3–0.49 = moderate association, and ≥ 0.5 = strong association). ASAS20 = Assessment of SpondyloArthritis international Society 20% improvement and absolute improvement of at least 1 unit (scale 0–10) in ≥3 of the 4 main ASAS domains, with no worsening by ≥20% in the remaining domain; ASAS40 = 40% improvement and absolute improvement of ≥2 units (scale 0–10) in ≥3 of the 4 main ASAS domains, with no worsening in the remaining domain; ASAS5/6 = ≥20% improvement in 5 of the 6 ASAS response criteria; ASAS partial remission = score ≤2 units (scale 0–10) in each of the 4 core ASAS domains; MCS = mental component score; PCS = physical component score; SF-36 = Short Form 36; VAS = visual analog scale.
† P values were not < 0.05.

status was not a significant predictor of fatigue responses based on regression analysis.

Association between clinical/laboratory variables and fatigue response. Correlation analyses based on pooled data from both trials revealed moderate positive correlations (correlation coefficients 0.33–0.49) between fatigue responses and ASAS partial remission at weeks 16, 52, and 104, and ASDAS-CRP major improvements at week 104. Strong correlations (correlation coefficients 0.52–0.71) were observed between fatigue and ASAS20, ASAS40, ASAS5/6 responses, and BASDAI 50 at weeks 16, 52, and 104, and ASDAS-CRP major improvements at weeks 16 and 52 (Table 3). No correlations were observed with hemoglobin levels, while moderate negative correlations were observed between fatigue and total back pain VAS and SF-36 scores, and moderate to strong negative correlations were observed for WPAI change from baseline (in percentage of overall work impairment and impairment while working due to health). There were significant differences in FACIT-F scores between responders and nonresponders across selected clinical response measures, including ASAS20, ASAS40, ASAS5/6, ASAS partial remision, ASDAS inactive disease, and BASDAI 50 (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24517/abstract).

DISCUSSION

Secukinumab treatment provided rapid and sustained improvements in fatigue for up to 3 years in the MEASURE 1 and MEASURE 2 randomized controlled trials (RCTs). Reducing fatigue, one of the most common symptoms in patients with AS, represents an important unmet need. Patients with AS have lower employment rates and experience more absenteeism from work than the general population (15). Fatigue has also been shown to contribute to reduced work productivity (16,17) and is one of the key treatment priorities for patients with AS (4). Previous short-term studies (ranging from 12 to 24 weeks) evaluating the effects of other therapies including adalimumab, etanercept, and infliximab have reported improvements in fatigue from baseline following treatment (18–20); however, to date, long-term data on the effects of biologic treatment on fatigue have been lacking.

Baseline fatigue levels were high in both MEASURE 1 and MEASURE 2, highlighting the need for better control of fatigue, among other symptoms of AS. Previously, secukinumab has been shown to rapidly improve signs and symptoms, physical function, HRQoL, and fatigue in patients with AS (11, 12, 21), with a low long-term rate of structural progression as measured by the modified Stoke Ankylosing Spondylitis Spinal Score (22). Results presented here build upon these findings, demonstrating significant and sustained improvements in fatigue, as measured using the FACIT-F score, compared with placebo in the overall population and particularly in the anti-TNF-naive population for up to 3 years. Across both trials, although fatigue responses were numerically greater in both anti-TNF-naive and anti-TNF-IR patients treated with secukinumab 150 mg compared with placebo, statistical significance was only evident in the anti-TNF-naive subgroup. However, the small numbers of patients in the anti-TNF-IR subgroups limit interpretation of these results; furthermore, these patients had previously failed treatment with 1 anti-TNF therapy, thus reflecting a more challenging population to treat. Whether prior treatment with anti-TNF therapy changes the course of disease or whether this is simply a more refractory patient population is not known; this question remains an area of further research.

In multivariate logistic regression analyses, age, prior treatment with anti-TNF therapy and FACIT-F, BASDAI, and ASDAS-CRP scores at baseline were significant predictors of short-term fatigue responses at week 16, and BASFI and FACIT-F scores of long-term fatigue responses at week 104. Univariate regression analyses showed that several baseline factors were predictive for short- and long-term absolute FACIT-F scores, including age, sex, BASDAI score, anti-TNF status, and HRQoL (SF-36 scores). Fatigue responses were moderately to strongly correlated with responses in several clinical measures, including ASAS20, ASAS40, ASAS5/6 responses, ASDAS-CRP, BASDAI score, and SF-36 PCS and MCS scores, highlighting the link between fatigue and other measures of disease in AS. Correlations between the FACIT-F, BASDAI, and SF-36 scores are to be expected owing to conceptual overlaps between these questionnaires (14,23,24) but serve to support the clinical relevance and reliability of these metrics in assessing and monitoring fatigue in patients with AS. Moderate to strong correlations were observed between WPAI outcomes and fatigue responses. This finding would appear consistent with reports in patients with rheumatoid arthritis where correlations between work productivity and fatigue levels have been observed (25).

Potential limitations of this study include the lack of a control group after week 16, the fact that long-term fatigue response data are presented as observed (missing data were not imputed), and the potential selection bias toward patients who elected to remain on long-term secukinumab treatment. However, retention of patients receiving secukinumab was high through 3 years, and withdrawals due to lack of efficacy were rare (26). Some FACIT-F questionnaire items are known to have the potential for misinterpretation due to their phrasing or limited relevance to the concept of fatigue in the context of arthritis (27). However, this questionnaire covers a range of fatigue concepts in easy-to-understand language and has been shown to confer good internal consistency and reliability, construct and criterion validity, and sensitivity to change (27), hence its selection for this investigation. In conclusion, results of the current analyses from 2 phase III RCTs showed that secukinumab 150 mg provided rapid and sustained reductions in fatigue through up to 3 years in patients with AS that were particularly prominent in patients who were anti-TNF-naive.

ACKNOWLEDGMENTS

We thank Aurore Yocolly (Health Economics and Outcomes Research Manager), Steffen Jugl (Health Economics and Outcomes Research Director), and Kunal Gandhi (Medical Director) of Novartis. Medical writing support for the development of this manuscript, under the direction of the authors, was provided by Ashfield MedComms, an Ashfield Health company, and funded by Novartis. This publication was written in accordance with Good Publications Practice (GPP3) guidelines (http://www.ismpp.org/gpp3).

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. Kvien 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. Kvien, Conaghan, Gossec, Strand, Østergaard, Poddubnyy, Williams, Porter, Shete, Gilloteau, Deodhar.

Acquisition of data. Kvien, Conaghan, Gossec, Strand, Østergaard, Poddubnyy, Williams, Porter, Shete, Gilloteau, Deodhar.

Analysis and interpretation of data. Kvien, Conaghan, Gossec, Strand, Østergaard, Poddubnyy, Williams, Porter, Shete, Gilloteau, Deodhar.

ROLE OF THE STUDY SPONSOR

The study was sponsored by Novartis Pharma AG and was designed by the scientific steering committee and Novartis personnel. Funding for medical writing support was provided by Novartis Pharma AG. Publication of this article was not contingent upon approval by Novartis Pharma AG.

REFERENCES

1. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis 2009;68:ii1–44.
2. Van der Heijde D, Ramiro S, Landewe R, Baraliakos X, Van den Bosch F, Sepriano A, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis 2017; 76:978–91.
3. Aissaoui N, Rostom S, Hakkou J, Berrada Ghziouel K, Bahiri R, Abouqal R, et al. Fatigue in patients with ankylosing spondylitis: prevalence and relationships with disease-specific variables, psychological status, and sleep disturbance. Rheumatol Int 2012;32:2117–24.

4. Heiberg T, Lie E, van der Heijde D, Kvien TK. Sleep problems are of higher priority for improvement for patients with ankylosing spondylitis than for patients with other inflammatory arthropathies. Ann Rheum Dis 2011;70:872–3.
5. Van Tubergen A, Coenen J, Landewé R, Spoorenberg A, Chorus A, Boonen A, et al. Assessment of fatigue in patients with ankylosing spondylitis: a psychometric analysis. Arthritis Rheum 2002;47:8–16.
6. Dernis-Labous E, Messow M, Dougados M. Assessment of fatigue in the management of patients with ankylosing spondylitis. Rheumatology (Oxford) 2003;42:1523–8.
7. Gossec L, Dougados M, D'Agostino MA, Fautrel B. Fatigue in early axial spondyloarthritis: results from the French DESIR cohort. Joint Bone Spine 2016;83:427–31.
8. Raychaudhuri SP, Raychaudhuri SK. Mechanistic rationales for targeting interleukin-17A in spondyloarthritis. Arthritis Res Ther 2017; 19:51.
9. Koenders MI, van den Berg WB. Secukinumab for rheumatology: development and its potential place in therapy. Drug Des Devel Ther 2016;10:2069–80.
10. European Medicines Agency. Secukinumab summary of product characteristics. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/003729/WC500183129.pdf.
11. Baeten D, Sieper J, Braun J, Baraliakos X, Dougados M, Emery P, et al. Secukinumab, an interleukin-17A inhibitor, in ankylosing spondylitis. N Engl J Med 2015;373:2534–48.
12. Marzo-Ortega H, Sieper J, Kivitz A, Blanco R, Cohen M, Martin R, et al. Secukinumab and sustained improvement in signs and symptoms of patients with active ankylosing spondylitis through two years: results from a phase III study. Arthritis Care Res (Hoboken) 2017;69: 1020–9.
13. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310:2191–4.
14. Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. J Pain Symptom Manage 1997;13:63–74.
15. Boonen A, Chorus A, Miedema H, van der Heijde D, Landewe R, Schouten H, et al. Withdrawal from labour force due to work disability in patients with ankylosing spondylitis. Ann Rheum Dis 2001;60:1033–9.
16. Espahbodi S, Bassett P, Cavill C, Freeth M, Hole J, Sengupta R. Fatigue contributes to work productivity impairment in patients with axial spondyloarthritis: a cross-sectional UK study. Clin Exp Rheumatol 2017;35:571–8.
17. Basu N, Druce K, Aikman L, Dilleen M, Barata T, Burden A, et al. Fatigue is a determinant of reduced work productivity in ankylosing spondylitis: results from a prospective cohort study (abstract FRI0431). Ann Rheum Dis 2016;75:591.
18. Revicki DA, Luo MP, Wordsworth P, Wong RL, Chen N, Davis JC Jr, et al. Adalimumab reduces pain, fatigue, and stiffness in patients with ankylosing spondylitis: results from the Adalimumab Trial Evaluating Long-term Safety and Efficacy for Ankylosing Spondylitis (ATLAS). J Rheumatol 2008;35:1346–53.
19. Calin A, Dijkmans BA, Emery P, Hakala M, Kalden J, Leirisalo-Repo M, et al. Outcomes of a multicentre randomised clinical trial of etanercept to treat ankylosing spondylitis. Ann Rheum Dis 2004;63: 1594–600.
20. Van der Heijde D, Dijkmans B, Geusens P, Sieper J, DeWoody K, Williamson P, et al. Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum 2005;52:582–91.
21. Deodhar AA, Dougados M, Baeten DL, Cheng-Chung Wei J, Geusens P, Readie A, et al. Effect of secukinumab on patient-reported outcomes in patients with active ankylosing spondylitis: a phase III randomized trial (MEASURE 1). Arthritis Rheumatol 2016; 68:2901–10.
22. Braun J, Baraliakos X, Deodhar A, Baeten D, Sieper J, Emery P, et al. Effect of secukinumab on clinical and radiographic outcomes in ankylosing spondylitis: 2-year results from the randomised phase III MEASURE 1 study. Ann Rheum Dis 2016;76:1070–7.
23. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
24. Ware JE Jr, Snow KK, Kosinski M, Gandek B. SF-36 health survey: manual and interpretation guide. Boston: The Health Institute, New England Medical Center; 1993.
25. Zhang W, Bansback N, Boonen A, Young A, Singh A, Anis AH. Validity of the work productivity and activity impairment questionnaire: general health version in patients with rheumatoid arthritis. Arthritis Res Ther 2010;12:R177.
26. Baraliakos X, Kivitz AJ, Deodhar AA, Braun J, Wei JC, Delicha EM, et al. Long-term effects of interleukin-17A inhibition with secukinumab in active ankylosing spondylitis: 3-year efficacy and safety results from an extension of the phase 3 MEASURE 1 trial. Clin Exp Rheumatol 2017;36:50–5.
27. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. Arthritis Rheum 2007;57: 429–39.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 768–775
DOI 10.1002/acr.24515

Long-Term Association Between Disease Activity and Disability in Early Axial Spondyloarthritis: Results From a Prospective Observational Study of Inflammatory Back Pain

Pedro D. Carvalho,[1] Adeline Ruyssen-Witrand,[2] Ana Marreiros,[3] and Pedro M. Machado[4]

Objective. Our primary objective was to study the long-term association between disease activity and disability in axial spondyloarthritis (SpA). Our secondary objective was to define patient profiles according to their level of disability.

Methods. We analyzed data collected during the first 5 years of follow-up of a large early axial SpA cohort, the Devenir des Spondylarthropathies Indifférénciées Récentes (DESIR) cohort. Multivariable models were built to study the association between the Health Assessment Questionnaire for Ankylosing Spondylitis (HAQ-AS) and the Ankylosing Spondylitis Disease Activity Score with C-reactive protein (ASDAS-CRP), adjusting for potential confounders. Hierarchical multivariable analysis was conducted using the chi-square automatic interaction detector (CHAID) method, to help determine how variables best cluster to explain HAQ-AS.

Results. Data from 644 patients and 5,152 visits were analyzed. HAQ-AS was longitudinally, independently, and positively associated with ASDAS-CRP (adjusted B [adjB] 0.205 [95% confidence interval (95% CI) 0.187, 0.222]), the enthesitis score (adjB 0.011 [95% CI 0.008, 0.015]), the Bath Ankylosing Spondylitis Metrology Index (adjB 0.087 [95% CI 0.069, 0.105]), and female sex (adjB 0.172 [95% CI 0.120, 0.225]). The CHAID decision tree revealed ASDAS-CRP as the first variable with discriminative power on HAQ-AS. The cutoffs that separated different patient disability profiles were obtained.

Conclusion. Disease activity contributes longitudinally to disability and is hierarchically superior to any other variable in explaining this health domain. Enthesitis and spinal mobility are also key drivers of disability in early axial SpA. ASDAS-CRP cutoffs that separated different patient disability profiles largely mimicked the cutoffs previously defined for ASDAS-CRP disease activity states.

INTRODUCTION

The broad spectrum of manifestations in axial spondyloarthritis (SpA) has the potential to impair several health outcomes. The major contributors to this impairment are patient pain, local tenderness, joint swelling, stiffness, fatigue, muscle spasm, and mobility loss, in addition to other associated comorbidities that may interfere with the impact of the disease. To fully understand the impact of such conditions, patients, physicians, and investigators must recognize the complex and heterogeneous relationships between the several involved health domains (1–3).

ClinicalTrials.gov identifier: NCT01648907.

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

The DESIR study is conducted with Assistance Publique Hopitaux de Paris as the sponsor. The DESIR study is also under the umbrella of the French Society of Rheumatology, which financially supports the cohort. The DESIR cohort is conducted under the control of Assistance Publique Hopitaux de Paris via the Clinical Research Unit Paris Centre and under the umbrella of the French Society of Rheumatology and Institut National de la Santé et de la Recherche Médicale (Inserm). Database management is performed within the Department of Epidemiology and Biostatistics (Dr. Pascale Fabbro-Peray, DIM and Nîmes, France).

The DESIR cohort was supported by an unrestricted grant from Pfizer. Dr. Machado's work was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

[1]Pedro D. Carvalho, MD: Centro Hospitalar Universitário do Algarve and Algarve Biomedical Centre, Faro, and Academic Medical Center, Universidade de Lisboa, Lisbon, Portugal; [2]Adeline Ruyssen-Witrand, MD, PhD: UMR 2017, Inserm, Université Paul Sabatier Toulouse III, Toulouse, France; [3]Ana Marreiros, PhD: Universidade do Algarve and Algarve Biomedical Centre, Faro, Portugal; [4]Pedro M. Machado, MD, PhD: University College London, University College London Hospitals NHS Foundation Trust, and London North West University Healthcare NHS Trust, London, UK.

Dr. Machado has received consulting and/or speaking fees from AbbVie, Bristol Myers Squibb, Celgene, Eli Lilly and Company, Janssen, MSD, Novartis, Pfizer, Roche and UCB (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Pedro M. Machado, MD, PhD, Centre for Rheumatology, UCL Division of Medicine, Rayne Building, 4th Floor, Room 415, 5 University Street, London WC1E 6JF, UK. Email: p.machado@ucl.ac.uk.

Submitted for publication May 11, 2020; accepted in revised form November 12, 2020.

SIGNIFICANCE & INNOVATIONS

- This study provides further validation of the Ankylosing Spondylitis Disease Activity Score with C-reactive protein (ASDAS-CRP) in early axial spondyloarthritis (SpA), as this index showed a robust longitudinal association with disability.
- ASDAS-CRP cutoffs were able to discriminate between different disability profiles in early axial SpA, suggesting that ASDAS-CRP can be used as a surrogate measure of disability.
- Enthesitis and spinal mobility should be systematically assessed, as they are also key drivers of disability in early axial SpA.
- Comprehensive and multimodal treatment strategies that also address enthesitis and spinal mobility may contribute to decreasing the overall level of disability in early axial SpA.

In a proposed stratified model for health outcomes in axial SpA, disability was shown to be hierarchically inferior to health-related quality of life and hierarchically superior to disease activity and spinal mobility (4). The association between disability and impairment of spinal mobility (5–7) as well as with disease activity (8–10) has been reported mainly in patients with ankylosing spondylitis (AS) with established disease. Furthermore, longitudinal studies are scarce.

The Ankylosing Spondylitis Disease Activity Score with C-reactive protein (ASDAS-CRP) has been progressively replacing the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) as the main disease activity measure to assess patients with axial SpA, both in the research context as well as in clinical practice (11,12). However, further evidence is needed to demonstrate its meaningfulness regarding the longitudinal relationship with disability, particularly in the early axial SpA subgroup. The Assessment of SpondyloArthritis international Society has proposed ASDAS-CRP cutoffs to define different disease activity states in axial SpA (13). These definitions were also endorsed by the Outcome Measures in Rheumatology (14) group, and the nomenclature of ASDAS-CRP disease activity states was recently revised (15). These cutoffs were shown to perform well in the differentiation of patients with different levels of disease activity according to several external constructs, but the capacity of the cutoffs to longitudinally differentiate between distinct patients' disability profiles has not been assessed yet. Several other clinical and demographic variables such as sex (16), body mass index (BMI) (16,17), and employment status (18) have also been described as being associated with disability in patients with axial SpA.

The objective of the current study was to assess the long-term association between disease activity and disability in early axial SpA, and to determine which other cofactors influence disability. We also aimed to describe the hierarchical relationship of variables that contribute to disability and to define ASDAS-CRP cutoffs capable of defining different profiles of patients regarding their disability status.

PATIENTS AND METHODS

Study population. We analyzed 5-year follow-up data from the Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort, which is a prospective observational study of patients with recent onset (<3 years) inflammatory back pain (evaluated by either the Calin or the Berlin criteria) (19,20), suggestive of axial SpA (a physician confidence score ≥5 on a scale from 0 to 10), as previously described (21). Consecutive patients ages 18–50 years from 25 centers in France were recruited. Visits were scheduled every 6 months during the first 2 years and yearly thereafter. We selected patients with a definite diagnosis of axial SpA at the end of follow-up (month 60), according to the treating rheumatologist. The DESIR study complies with the Declaration of Helsinki, was conducted according to good clinical practice guidelines, and was approved by the appropriate local research ethics committees, and informed consent was obtained from the study subjects.

Outcome measures. Disability was assessed using the Health Assessment Questionnaire for Ankylosing Spondylitis (HAQ-AS), a validated patient-reported outcome used to assess patient disability. It encompasses 8 main categories: dressing and grooming, rising, eating, walking, hygiene, reach, grip, and activities. The HAQ-AS was adapted from the original HAQ, incorporating issues of disability and impairment specific to patients with axial SpA. The HAQ was one of the first self-report disability measures and has become the dominant instrument in many disease areas, including rheumatic and musculoskeletal diseases. The final score is calculated as the mean of the HAQ domains; the score may range from 0 to 3, with higher scores reflecting lower physical ability (more disability) (22).

Disease activity was measured using the ASDAS-CRP (13–15,23) and BASDAI. BASDAI was used as an alternative to ASDAS-CRP to explore differences between the models. The ASDAS-CRP consists of patient assessments of 4 different domains (back pain, duration of morning stiffness, global assessment of disease activity, and peripheral pain/swelling) and is assessed on a 0–10 numeric rating scale by patients, plus an inflammatory marker (CRP level) (23). The recently reviewed disease activity states defined for ASDAS-CRP (inactive, low, high, and very high) were taken into account for this study (15). BASDAI is a patient-reported outcome consisting of 6 questions, scored on a 0–10 numeric rating scale and encompassing the following domains: fatigue, spinal pain, peripheral arthritis, enthesitis, and intensity and duration of morning stiffness (24). The concise Mander Enthesitis Score (cMES) with gradation was used to evaluate enthesitis involvement. This score considers 13 entheseal sites, each one graded from 0 to 3 (0 = no pain, 1 = mild tenderness,

2 = moderate tenderness, 3 = wince or withdraw), the total range being from 0 to 39 (25).

Statistical analysis. In a first exploratory step, univariable associations were investigated using generalized estimating equations (GEEs). GEEs allow the combination of multiple measurements per patient and use all available data during follow-up, while taking into account missing values and correcting for within-patient correlation (26).

HAQ-AS was used as the dependent variable; disease activity (defined by either ASDAS-CRP or BASDAI) and clinical and demographic variables were used as independent variables in univariable models: age (years), sex (male/female), level of education (university or equivalent: yes/no), BMI (kg/m^2), HLA–B27 status (positive/negative), symptom duration (years), employment status (currently employed: yes/no), smoking status (current smoking: yes/no), past or current peripheral arthritis (yes/no), history of inflammatory bowel disease (yes/no), psoriasis (yes/no), uveitis (yes/no), dactylitis (yes/no), CRP level (mg/liter), linear version of the Bath Ankylosing Spondylitis Metrology Index (BASMI) score, enthesitis score (cMES), and current treatment with nonsteroidal antiinflammatory drugs (NSAIDs; yes/no), conventional disease-modifying antirheumatic drugs (cDMARDs; yes/no), or tumor necrosis factor inhibitors (TNFi; yes/no). Imaging data were also taken into account, namely the presence of sacroiliitis according to modified New York (mNY) criteria, the presence of active sacroiliitis on magnetic resonance imaging (MRI), the level of

Table 1. Summary of baseline clinical and demographic characteristics of the 644 patients with an axial spondyloarthritis diagnosis at the end of follow-up (60 months)*

Characteristic	Value	No. valid
Age, years	33.6 ± 8.6	644
Female, no. (%)	341 (53.0)	644
Education at baseline, university or equivalent, no. (%)	383 (59.8)	640
Body mass index, kg/m^2	23.9 ± 4.1	638
HLA–B27 positive, no. (%)	385 (59.9)	643
Symptom duration at baseline, years	1.5 ± 0.9	643
Currently employed, no. (%)	553 (86.7)	638
Current smoking, no. (%)	237 (37.1)	639
Current arthritis, no. (%)	49 (7.6)	643
History of arthritis, no. (%)	181 (28.3)	640
History of inflammatory bowel disease, no. (%)	33 (5.1)	644
History of psoriasis, no. (%)	113 (17.5)	644
History of uveitis, no. (%)	61 (9.5)	644
History of dactylitis, no. (%)	93 (14.5)	642
Enthesitis score†	4.3 ± 6.0	642
C-reactive protein, mg/liter	8.3 ± 5.1	624
ASDAS-CRP	2.7 ± 1.0	615
BASDAI	4.5 ± 2.0	638
HAQ-AS	0.7 ± 0.5	642
BASFI	3.1 ± 2.3	635
BASMI	2.4 ± 1.0	555
Sacroiliitis (mNY criteria), no. (%)	110 (17.5)	630
Active sacroiliitis on MRI (ASAS criteria), no. (%)	231 (36.8)	628
mSASSS (range 0 to 72)	0.4 ± 1.6	570
MRI spinal inflammation (Berlin score)	0.3 ± 1.3	470
NSAID intake last 6 months, no. (%)	601 (93.3)	644
cDMARD intake last 6 months, no. (%)	89 (13.8)	643
TNFi treatment last 6 months, no. (%)	0 (0.0)	644

* Values are the mean ± SD unless indicated otherwise. ASAS = Assessment of Spondyloarthritis international Society; ASDAS-CRP = Ankylosing Spondylitis Disease Activity Score with C-reactive protein; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; BASMI = Bath Ankylosing Spondylitis Metrology Index; cDMARD = conventional disease-modifying antirheumatic drug; HAQ-AS = Health Assessment Questionnaire for Ankylosing Spondylitis; mNY = modified New York; MRI = magnetic resonance imaging; mSASSS = modified Stoke Ankylosing Spondylitis Spinal Score; NSAID = nonsteroidal antiinflammatory drug; TNFi = tumor necrosis factor inhibitor.
† Concise Mander Enthesitis Score with gradation was used.

Table 2. Univariable analyses to investigate the association between HAQ-AS (dependent variable) and other clinical and demographic variables (independent variables)*

Characteristics	B (95% CI)	*P*
Age, years	0.002 (–0.002, 0.005)	0.373
Female	0.311 (0.243, 0.379)	<0.001†
Body mass index, kg/m^2	0.007 (0.001, 0.013)	0.034†
HLA–B27 positive	–0.185 (–0.259, –0.111)	<0.001†
Symptom duration, years	–0.016 (–0.023, –0.009)	<0.001†
Currently employed	–0.046 (–0.087, –0.005)	0.029†
Current smoking	0.011 (–0.025, 0.048)	0.540
Current arthritis	0.191 (0.121, 0.261)	<0.001†
History of arthritis	0.042 (–0.015, 0.099)	0.151
History of inflammatory bowel disease	0.051 (–0.028, 0.130)	0.208
History of psoriasis	–0.021 (–0.079, 0.036)	0.463
History of uveitis	0.010 (–0.076, 0.095)	0.825
History of dactylitis	0.020 (–0.050, 0.091)	0.568
ASDAS-CRP	0.233 (0.216, 0.249)	<0.001†
BASDAI	0.119 (0.111, 0.127)	<0.001†
BASMI	0.147 (0.125, 0.168)	<0.001†
Enthesitis score‡	0.022 (0.019, 0.025)	<0.001†
C-reactive protein, mg/liter	0.006 (0.004, 0.008)	<0.001†
Presence of sacroiliitis (mNY criteria)§	–0.055 (–0.115, 0.005)	0.073†
Presence of MRI-active sacroiliitis§	0.006 (–0.051, 0.063)	0.841
MRI spinal inflammation (Berlin score)§	–0.007 (–0.025, 0.011)	0.462
mSASSS§	–0.007 (–0.022, 0.007)	0.309
NSAID intake last 6 months	0.121 (0.090, 0.152)	<0.001†
cDMARD intake last 6 months	0.084 (0.034, 0.134)	0.001†
TNFi treatment last 6 months	–0.091 (–0.138, –0.043)	<0.001†

* 95% CI = 95% confidence interval (see Table 1 for other definitions).
† Statistically significant.
‡ Concise Mander Enthesitis Score with gradation was used.
§ Assessment of mSASSS and spinal Berlin score performed only at baseline and at 24 and 60 months; assessment of mNY criteria and active sacroiliitis (Assessment of Spondyloarthritis international Society criteria) on MRI performed at baseline, 12, 24, and 60 months.

Table 3. Multivariable analyses to investigate the association between HAQ-AS (dependent variable) and ASDAS-CRP and other clinical and demographic variables (independent variables)*

Characteristics	Adjusted B (95% CI)	*P*
Female	0.172 (0.120, 0.225)	<0.001
Body mass index, kg/m^2	†	–
HLA–B27 positive	†	–
Symptom duration, years	†	–
Currently employed	†	–
Current arthritis	†	–
ASDAS-CRP	0.205 (0.187, 0.222)	<0.001
Enthesitis score	0.011 (0.008, 0.015)	<0.001
BASMI	0.087 (0.069, 0.105)	<0.001
Presence of sacroiliitis (mNY criteria)	†	–
NSAID intake last 6 months	†	–
cDMARD intake last 6 months	†	–
TNFi treatment last 6 months	†	–

* All *P* values shown are statistically significant. 95% CI = 95% confidence interval; ASDAS-CRP = Ankylosing Spondylitis Disease Activity Score with C-reactive protein; BASMI = Bath Ankylosing Spondylitis Metrology Index; cDMARD = conventional disease-modifying antirheumatic drug; HAQ-AS = Health Assessment Questionnaire for Ankylosing Spondylitis; mNY = modified New York; NSAID = nonsteroidal antiinflammatory drug; TNFi = tumor necrosis factor inhibitor.
† Excluded from the best-fit model.

spinal MRI inflammation as assessed by the Berlin scoring system, and the level of structural damage as assessed by the modified Stoke Ankylosing Spondylitis Spinal Score (mSASSS).

Multivariable GEE models were then built, using a manual forward selection procedure until the best-fit model was obtained, taking confounding or contributing variables into account. This process was an intelligible method that took into account clinical relevance and statistical parameters of each variable, such as *P* value, quasi-likelihood under independence model criterion, and the influence of each of the secondary variables in the 95% confidence interval (95% CI) in the main association, when the final model was built. Variables with a *P* value less than 0.10 in the univariable analysis were tested in these multivariable models.

Subsequently, a decision tree was constructed using unbiased hierarchical multivariable analysis applying the chi-square automatic interaction detector (CHAID) method, with HAQ-AS as the dependent variable (27). The following independent variables were tested: ASDAS-CRP, enthesitis score, the presence of arthritis, employment status, sex, symptom duration, BMI, HLA–B27 status, BASMI score, the presence of sacroiliitis according to mNY criteria, and treatment with NSAIDs, cDMARDs, or TNFi. These variables were selected based on their significance in the above GEE univariable models.

The CHAID method is a technique used to study the relationship between variables, automatically building a tree model that depicts how variables best merge to explain a defined dependent outcome. A decision tree like CHAID works by recursively partitioning the data based on input field values. The data partitions are called branches. The initial/parent branch (sometimes called the root) encompasses all data records. The parent branch is split into subsets, or child branches, based on the value of a particular input field. Each child branch can be further split into subbranches, which can in turn be split again, and so on. At the lowest level of the tree are branches that have no more splits. Such branches are known as terminal branches (or leaves). We stipulated a minimum requirement of 70 cases to allow the creation of a parent branch and a minimum number of 20 cases, to allow the creation of child branches. After setting up these criteria, the results were purely analysis driven in the model. The stopping criteria prevented the overfitting of data, as all of the child branches had >0.25% of cases from the training population, as methodologically recommended (28,29). IBM SPSS statistics software, version 20, was used to conduct the statistical analysis.

RESULTS

Study population and cohort data. Data from 644 patients and 5,152 visits were analyzed. Baseline characteristics of the study population are shown in Table 1. Of the 708 patients included in the DESIR cohort, 64 patients were excluded from this analysis because they did not have a definite clinical diagnosis of axial SpA at month 60. Taking the first visit into account, the study cohort was a young (mean age 33.6 years) and sex-balanced population (53% female) with short disease duration (mean symptom duration 1.5 years) and a very

Table 4. Multivariable analyses to investigate the association between HAQ-AS (dependent variable) and BASDAI and other clinical and demographic variables (independent variables)*

Characteristics	Adjusted B (95% CI)	*P*
Female	0.154 (0.105, 0.202)	<0.001
Body mass index, kg/m^2	†	–
HLA–B27 positive	†	–
Symptom duration, years	†	–
Currently employed	†	–
Current arthritis	†	–
BASDAI	0.104 (0.096, 0.113)	<0.001
Enthesitis score	0.009 (0.006, 0.012)	<0.001
C-reactive protein, mg/liter	0.003 (0.001, 0.004)	0.001
BASMI	0.079 (0.062, 0.095)	<0.001
Presence of sacroiliitis (mNY criteria)	†	–
NSAID intake last 6 months	†	–
cDMARD intake last 6 months	†	–
TNFi treatment last 6 months	†	–

* All *P* values shown are statistically significant. 95% CI = 95% confidence interval; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASMI = Bath Ankylosing Spondylitis Metrology Index; cDMARD = conventional disease-modifying antirheumatic drug; HAQ-AS = Health Assessment Questionnaire for Ankylosing Spondylitis; mNY = modified New York; NSAID = nonsteroidal antiinflammatory drug; TNFi = tumor necrosis factor inhibitor.
† Excluded from the best-fit model.

Table 5. Overlap between the published cutoffs for ASDAS-CRP disease activity states and the patient clusters defined by the CHAID algorithm when disability was studied in early axial spondyloarthritis*

	Inactive	Low	High	Very high
ASDAS-CRP disease activity states and respective cutoffs	<1.3	≥1.3 to <2.1	≥2.1 to ≤3.5	>3.5
Current work (CHAID algorithm for disability)				
First level: ASDAS-CRP cutoff intervals with importance in the differentiation of disability clusters	≤1.0, 1.0–1.3	1.3–1.6, 1.6–1.9, 1.9–2.2	2.2–2.5, 2.5–2.7, 2.7–3.1, 3.1–3.5	>3.5
Second level: variables with discriminative power regarding disability	Sex	Enthesitis, BASMI	Enthesitis, BASMI	Enthesitis
Third level: variables with discriminative power regarding disability	BASMI	Sex, enthesitis, BASMI, TNFi	Sex, BASMI, symptom duration	NSAID, BASMI, TNFi

* ASDAS-CRP = Ankylosing Spondylitis Disease Activity Score with C-reactive protein; BASMI = Bath Ankylosing Spondylitis Metrology Index; CHAID = chi-square automatic interaction detector; NSAID = nonsteroidal antiinflammatory drug; TNFi = tumor necrosis factor inhibitor.

low level of spinal radiographic damage (mean mSASSS 0.6 units).

Univariable analysis. There was a significant univariable association (Table 2) between a higher level of disability (i.e., higher HAQ-AS scores) and female sex, higher BMI, the presence of peripheral arthritis, higher disease activity (measured by ASDAS-CRP, BASDAI, or CRP level alone), less spinal mobility (measured by BASMI), higher enthesitis score (measured by cMES), and NSAID or cDMARD treatment in the last 6 months. Conversely, HLA–B27 positivity, symptom duration, being employed, and treatment with TNFi in the last 6 months were associated with lower levels of disability. The presence of sacroiliitis according to mNY criteria was not statistically significant, but it was considered in the multivariable and hierarchical models, as the *P* value was <0.10.

Multivariable analysis. The multivariable model using ASDAS-CRP as a measure of disease activity showed a significant, independent, and positive association between HAQ-AS and ASDAS-CRP (adjusted B [adjB] 0.205 [95% CI 0.187, 0.222]), enthesitis score (adjB 0.011 [95% CI 0.008, 0.015]), BASMI score (adjB 0.087 [95% CI 0.069, 0.105]), and female sex (adjB 0.172 [95% CI 0.120, 0.225]) (Table 3).

When BASDAI was used as disease activity measure, a significant, independent, and positive association with disability was also found for BASDAI (adjB 0.104 [95% CI 0.096, 0.113]), enthesitis score (adjB 0.009 [95% CI 0.006, 0.012]), BASMI score (adjB 0.079 [95% CI 0.062, 0.095]), and female sex (adjB 0.154 [95% CI 0.105, 0.202]). Higher CRP levels (not included in the ASDAS-CRP model due to collinearity/redundancy) were also associated with disability (adjB 0.003 [95% CI 0.001, 0.004]) (Table 4).

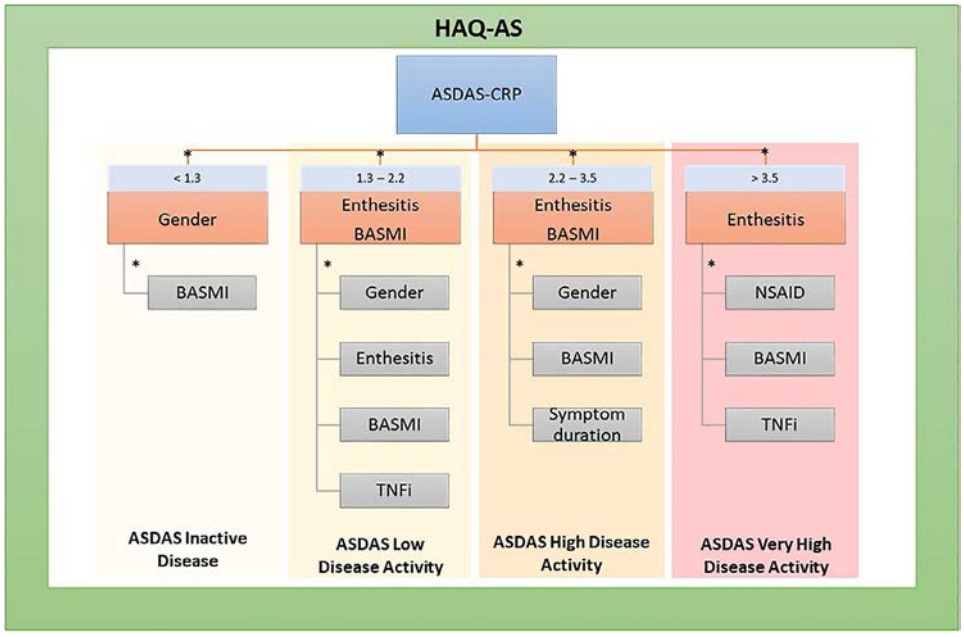

Figure 1. Decision tree showing the Ankylosing Spondylitis Disease Activity Score with C-reactive protein (ASDAS-CRP) as the first variable with discriminative power on the Health Assessment Questionnaire for Ankylosing Spondylitis (HAQ-AS). BASMI = Bath Ankylosing Spondylitis Metrology Index; NSAID = nonsteroidal antiinflammatory drug; TNFi = tumor necrosis factor inhibitor. * = $P < 0.001$.

Decision tree model. The decision tree revealed ASDAS-CRP as the first variable with the most discriminative power on HAQ-AS (Figure 1). According to this hierarchical model, after ASDAS-CRP, the variables sex, enthesitis score, and BASMI were the next most likely in explaining HAQ-AS variation, followed by TNFi treatment, symptom duration, and NSAID treatment. The decision tree model automatically created cutoffs for continuous variables that better differentiate patients' profiles. Regarding ASDAS-CRP, the cutoffs that generated a second-level variable in the tree were 1.3, 2.2, and 3.5, as shown in Table 5 and in Figure 1. The cutoffs generated for the enthesitis score as a secondary variable were 2.0, 3.0, 6.0, and 10.0; for BASMI as a secondary variable the generated cutoffs were 1.8, 2.1, 2.5, 3.1, and 3.6. The only cutoff for symptom duration was 2.7 years.

DISCUSSION

This analysis demonstrates a robust and consistent long-term association between disease activity and disability in an early axial SpA cohort and provides further evidence on the validity of the ASDAS-CRP and its cutoffs. Moreover, ASDAS-CRP disease activity categories were also able to capture distinct disability profiles in early axial SpA. Our study also suggests that there are several other factors playing an important role in disability. In addition to clinical disease activity, 3 other variables were independently associated with disability: enthesitis, spinal mobility (measured by BASMI), and sex.

Both ASDAS-CRP and BASDAI were significantly associated with HAQ-AS in the multivariable models. A higher adjB value was found for ASDAS-CRP (adjB 0.205 [95% CI 0.187, 0.222]) compared to BASDAI (adjB 0.104 [95% CI 0.096, 0.113]). These values mean that a certain increment in ASDAS-CRP leads to a higher variation in disability than the same numerical increment in BASDAI. However, the ranges of the measures are different (BASDAI ranges from 0 to 10, while ASDAS-CRP ranges from 0.6 to approximately 7). Therefore results are not comparable, and our study was not designed to directly compare these indexes.

The CHAID algorithm is a robust tool that allowed us to discriminate between subgroups of patients with different mean HAQ-AS scores, based on a certain set of variables. This analysis allowed us to conclude that ASDAS-CRP was the most important predictor of disability defined by HAQ-AS. This method allowed us to determine other important variables in the differentiation of patients' profiles regarding their level of disability, as well as their hierarchical relationship. Sex, enthesitis score, and BASMI score were the next variables explaining disability, followed by duration of symptoms, TNFi, and NSAID treatment. These results are consistent with the results found in the multivariable model for ASDAS-CRP (Table 3) and emphasize the multidimensional aspect of disability. Previous studies are consistent with our findings regarding the influence of disease activity (8–10,16), sex (16), and spinal mobility (6,7) on disability. Peripheral arthritis has previously been reported to have an association with disability in AS (30). However, in our early axial SpA cohort, the main peripheral manifestation associated with disability was enthesitis (the association with peripheral arthritis was only significant in the univariable models). The association between disability and enthesitis had previously been described in patients with AS but not in early axial SpA (31,32). Both TNFi and NSAID treatment emerged as third-level predictors of disability in the decision tree model, but not treatment with cDMARDs. These findings are in line with the expected lack of efficacy and therefore potential limited role of cDMARDs (33–35) in axial SpA, while both TNFi and NSAIDs have been shown to be effective treatments in axial SpA (35). Associations between BMI (16,17) and employment status (18) were previously found in literature, but in our study these associations were not significant.

Using an arithmetic approach, this algorithm automatically and in an unbiased manner identified ASDAS-CRP cutoffs able to differentiate between groups of patients according to their level of disability. Notably, the ASDAS-CRP cutoffs that separated different patient disability profiles largely mimicked the cutoffs previously defined for ASDAS-CRP disease activity states: inactive (<1.3), low (≥1.3 to <2.1), high (≥2.1 to ≤3.5), and very high disease activity (>3.5) (15). There was only a minor difference in the 2.1 cutoff value separating low from high disease activity, versus 2.2 as the relevant cutoff in the decision tree analysis. These findings indicate that the disease activity states defined for ASDAS-CRP also perform well in stratifying patients with different levels of disability, which adds to the validity of the ASDAS-CRP cutoffs.

We did not observe an association between disability and structural damage. This lack of association is a relevant albeit not unexpected finding, as this study is a cohort of patients with early disease and very low levels of structural damage. The association between structural damage and disability is well described in the literature for patients with longer disease duration and higher levels of structural damage (36–38). These findings suggest that structural damage contributes less to the disease burden in the early stages of the disease.

Our study has some limitations. First, the near absence of structural damage in this early axial SpA cohort (95.4% and 93.2% of patients had an mSASSS score ≤2 units at baseline and 60 months, respectively) limits the ability to study the influence of structural damage on disability. Regarding treatment, patients were not randomly assigned to the described treatments, meaning that possible confounding by indication may be present. A study designed to evaluate the impact of those treatments would be more adequate to evaluate their true impact on disability. Third, to determine the most accurate level of association between disease activity and true disability, a more robust external construct (gold standard) should be used. An attempt to use such a construct was done with data from the Program to Understand the Long-term outcomes in

Spondyloarthritis (PULSAR) registry, a nested study from the US Veterans Affairs Administration. In PULSAR, disability was defined as patients being classified as disabled or not by 1 of 2 federal agencies (the Disability Determination Services or the US Department of Veterans Affairs). That study enrolled 62 patients who completed BASDAI, ASDAS-CRP, and Bath Ankylosing Spondylitis Functional Index (BASFI) questionnaires after a routine visit. Interestingly, in that study, BASDAI was not significantly associated with disability as defined above, while BASFI and ASDAS-CRP were predictors of federally determined disability (odds ratio [OR] 1.67 [95% CI 1.12, 2.47] and OR 1.34 [95% CI 1.02, 1.76], respectively) (9). Our study would be strengthened if we had additional external criteria for disability against which we could validate our results; however, criteria such as those used by PULSAR are also subjective, have limitations, and might not necessarily represent the gold standard. Fourth, there is the possibility that some comorbidities, such as fibromyalgia, may have some influence in the association between HAQ-AS and some variables, such as sex. However, information on the presence of concomitant fibromyalgia was not available in the DESIR database.

In conclusion, we have shown that disease activity contributes longitudinally to disability and that it is hierarchically superior to any other variables or disease domains, in the context of an early axial SpA cohort with minimal structural damage. Enthesitis, sex, and spinal mobility are also key drivers of disability in early axial SpA. ASDAS-CRP, a composite index used to measure disease activity in axial SpA, also reflects the level of disability, and its cutoffs are able to discriminate between different profiles of disability in early axial SpA, reinforcing the validity of the ASDAS-CRP disease activity cutoffs.

ACKNOWLEDGMENTS

We wish to thank the different regional participating centers: Pr. Maxime Dougados, Dr. Anna Molto (Paris-Cochin), Pr. Philippe Dieudé (Paris-Bichat), Pr. Laure Gossec (Paris-La Pitié-Salpêtrière), Pr. Francis Berenbaum (Paris-Saint-Antoine), Pr. Pascal Claudepierre (Créteil), Pr. Maxime Breban (Boulogne-Billancourt), Dr. Bernadette Saint-Marcoux (Aulnay-sous-Bois), Pr. Philippe Goupille (Tours), Pr. Jean Francis Maillefert (Dijon), Dr. Emmanuelle Dernis (Le Mans), Pr. Daniel Wendling (Besançon), Pr. Bernard Combe (Montpellier), Pr. Liana Euller-Ziegler (Nice), Pr. Pascal Richette (Paris Lariboisière), Pr. Pierre Lafforgue (Marseille), Pr. Patrice Fardellone, Dr. Patrick Boumier (Amiens), Pr. Martin Soubrier (Clermont-Ferrand), Dr. Nadia Mehsen (Bordeaux), Pr. Damien Loeuille (Nancy), Pr. René-Marc Flipo (Lille), Pr. Alain Saraux (Brest), Pr. Stéphan Pavy (Le Kremlin-Bicêtre), Pr. Adeline Ruyssen-Witrand (Toulouse), Pr. Olivier Vittecoq (Rouen). We wish to thank the research nurses, the staff members of the clinical research unit of Paris Centre, the staff members of the Biological Resource Center of Bichat Hospital, the staff members of the Department of Statistics of Nîmes and all the investigators, in particular Jérôme Allain, Thierry Lequerre, Béatrice Banneville, Julien Champey, Christine Piroth, Anne Tournadre, Sophie Trijau, Salah Ferkal, Clement Prati, Marie-Agnes Timsit, and Eric Toussirot for active patient recruitment and monitoring.

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. Machado 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. Carvalho, Marreiros, Machado.

Analysis and interpretation of data. Carvalho, Ruyssen-Witrand, Marreiros, Machado.

REFERENCES

1. Rohde G, Berg KH, Prøven A, Haugeberg G. The relationship between demographic- and disease-related variables and health-related quality of life in patients with axial spondyloarthritis. BMC Musculoskelet Disord 2017;18:328.
2. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis 2009;68 Suppl 2:ii1–44.
3. Ward MM. Outcomes in ankylosing spondylitis: what makes the assessment of treatment effects in ankylosing spondylitis different? Ann Rheum Dis 2006;65 Suppl 3:iii25–8.
4. Machado P, Landewé R, Braun J, Hermann KG, Baraliakos X, Baker D, et al. A stratified model for health outcomes in ankylosing spondylitis. Ann Rheum Dis 2011;70:1758–64.
5. Almodovar R, Zarco P, Collantes E, Gonzalez C, Mulero J, Fernandez-Sueiro JL, et al. Relationship between spinal mobility and disease activity, function, quality of life and radiology: a cross-sectional Spanish registry of spondyloarthropathies (REGISPONSER). Clin Exp Rheumatol 2009;27:439–45.
6. Bostan EE, Borman P, Bodur H, Barca N. Functional disability and quality of life in patients with ankylosing spondylitis. Rheumatol Int 2003;23:121–6.
7. Ibn Yacoub Y, Amine B, Laatiris A, Abouqal R, Hajjaj-Hassouni N. Spinal mobility and its impact in Moroccan patients with ankylosing spondylitis. Clin Rheumatol 2011;30:239–43.
8. Fernández-Carballido C, Navarro-Compán V, Castillo-Gallego C, Castro-Villegas MC, Collantes-Estévez E, de Miguel E. Disease activity as a major determinant of quality of life and physical function in patients with early axial spondyloarthritis. Arthritis Care Res (Hoboken) 2017;69:150–5.
9. Quinzanos I, Luong PT, Bobba S, Steuart Richards J, Majithia V, Davis LA, et al. Validation of disease activity and functional status questionnaires in spondyloarthritis. Clin Exp Rheumatol 2015;33: 146–52.
10. Van der Heijde D, Joshi A, Pangan AL, Chen N, Betts K, Mittal M, et al. ASAS40 and ASDAS clinical responses in the ABILITY-1 clinical trial translate to meaningful improvements in physical function, health-related quality of life and work productivity in patients with non-radiographic axial spondyloarthritis. Rheumatology (Oxford) 2016; 55:80–8.
11. Machado P, Landewe R. Spondyloarthritis: is it time to replace BASDAI with ASDAS? Nat Rev Rheumatol 2013;9:388–90.
12. Machado PM, Raychaudhuri SP. Disease activity measurements and monitoring in psoriatic arthritis and axial spondyloarthritis. Best Pract Res Clin Rheumatol 2014;28:711–28.
13. Machado P, Landewé R, Lie E, Kvien TK, Braun J, Baker D, et al. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 2011;70:47–53.
14. Machado PM, Landewe RB, van der Heijde DM. Endorsement of definitions of disease activity states and improvement scores for

the Ankylosing Spondylitis Disease Activity Score: results from OMERACT 10. J Rheumatol 2011;38:1502–6.

15. Machado PM, Landewe R, van der Heijde D. Ankylosing Spondylitis Disease Activity Score (ASDAS): 2018 update of the nomenclature for disease activity states. Ann Rheum Dis 2018;77:1539–40.
16. Putrik P, Ramiro S, Molto A, Keszei AP, Norton S, Dougados M, et al. Individual-level and country-level socioeconomic determinants of disease outcomes in SpA: multinational, cross-sectional study (ASAS-COMOSPA). Ann Rheum Dis 2019;78:486–93.
17. Micheroli R, Hebeisen M, Wildi LM, Exer P, Tamborrini G, Bernhard J, et al. Impact of obesity on the response to tumor necrosis factor inhibitors in axial spondyloarthritis. Arthritis Res Ther 2017;19:164.
18. Castillo-Ortiz JD, Ramiro S, Landewé R, van der Heijde D, Dougados M, van den Bosch F, et al. Work outcome in patients with ankylosing spondylitis: results from a 12-year followup of an international study. Arthritis Care Res (Hoboken) 2016;68:544–52.
19. Calin A, Porta J, Fries JF, Schurman DJ. Clinical history as a screening test for ankylosing spondylitis. JAMA 1977;237:2613–4.
20. Rudwaleit M, Metter A, Listing J, Sieper J, Braun J. Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. Arthritis Rheum 2006;54:569–78.
21. Dougados M, Etcheto A, Molto A, Alonso S, Bouvet S, Daures JP, et al. Clinical presentation of patients suffering from recent onset chronic inflammatory back pain suggestive of spondyloarthritis: the DESIR cohort. Joint Bone Spine 2015;82:345–51.
22. Revicki D, Jin Y, Rao S, Mease PJ, Cifaldi M. Validation of the health assessment questionnaire for spondyloarthritis in patients with non-radiographic axial spondyloarthritis (poster 1345). Annual Scientific Meeting of the American College of Rheumatology, Washington, DC, 2012.
23. Lukas C, Landewe R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68: 18–24.
24. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
25. Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A, Landewe R, van ver Tempel H, Mielants H, et al. Assessment of enthesitis in ankylosing spondylitis. Ann Rheum Dis 2003;62:127–32.
26. Twisk JW. Longitudinal data analysis. A comparison between generalized estimating equations and random coefficient analysis. Eur J Epidemiol 2004;19:769–76.
27. Kass GV. An exploratory technique for investigating large quantities of categorical data. Appl Statist 1980;29:119–27.
28. Song YY, Lu Y. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry 2015;27:130–5.
29. Berry MJ, Linoff GS. Mastering data mining: the art and science of customer relationship management. New York: John Wiley; 1999.
30. Yilmaz O, Tutoglu A, Garip Y, Ozcan E, Bodur H. Health-related quality of life in Turkish patients with ankylosing spondylitis: impact of peripheral involvement on quality of life in terms of disease activity, functional status, severity of pain, and social and emotional functioning. Rheumatol Int 2013;33:1159–63.
31. Zahiroglu Y, Ulus Y, Akyol Y, Tander B, Durmus D, Bilgici A, et al. Spondyloarthritis Research Consortium of Canada (SPARCC) enthesitis index in Turkish patients with ankylosing spondylitis: relationship with disease activity and quality of life. Int J Rheum Dis 2014;17: 173–80.
32. Rezvani A, Bodur H, Ataman S, Kaya T, Bugdayci DS, Demir SE, et al. Correlations among enthesitis, clinical, radiographic and quality of life parameters in patients with ankylosing spondylitis. Mod Rheumatol 2014;24:651–6.
33. Chen J, Veras MM, Liu C, Lin J. Methotrexate for ankylosing spondylitis. Cochrane Database Syst Rev 2013;2:Cd004524.
34. Chen J, Lin S, Liu C. Sulfasalazine for ankylosing spondylitis. Cochrane Database Syst Rev 2014;11:Cd004800.
35. Van der Heijde D, Ramiro S, Landewé R, Baraliakos X, Van den Bosch F, Sepriano A, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis 2017;76:978–91.
36. Protopopov M, Sieper J, Haibel H, Listing J, Rudwaleit M, Poddubnyy D. Relevance of structural damage in the sacroiliac joints for the functional status and spinal mobility in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Res Ther 2017;19:240.
37. Poddubnyy D, Listing J, Haibel H, Knuppel S, Rudwaleit M, Sieper J. Functional relevance of radiographic spinal progression in axial spondyloarthritis: results from the GErman SPondyloarthritis Inception Cohort. Rheumatology (Oxford) 2018;57:703–11.
38. Landewe R, Dougados M, Mielants H, van der Tempel H, van der Heijde D. Physical function in ankylosing spondylitis is independently determined by both disease activity and radiographic damage of the spine. Ann Rheum Dis 2009;68:863–7.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 776–789
DOI 10.1002/acr.24519

Does Screening for Depressive Symptoms Help Optimize Duloxetine Use in Knee Osteoarthritis Patients With Moderate Pain? A Cost-Effectiveness Analysis

Nora K. Lenhard,[1] James K. Sullivan,[1] Eric L. Ross,[2] Shuang Song,[1] Robert R. Edwards,[1] David J. Hunter,[3] Tuhina Neogi,[4] Jeffrey N. Katz,[5] and Elena Losina[6]

Objective. Duloxetine is a treatment approved by the US Food and Drug Administration for both osteoarthritis (OA) pain and depression, though uptake of duloxetine in knee OA management varies. We examined the cost-effectiveness of adding duloxetine to knee OA care in the absence or presence of depression screening.

Methods. We used the Osteoarthritis Policy Model, a validated computer microsimulation of knee OA, to examine the value of duloxetine for patients with knee OA who have moderate pain by comparing 3 strategies: 1) usual care, 2) usual care plus duloxetine for patients who screen positive for depression on the Patient Health Questionnaire 9 (PHQ-9), and 3) usual care plus universal duloxetine. Outcome measures included quality-adjusted life years (QALYs), lifetime direct medical costs, and incremental cost-effectiveness ratios (ICERs), discounted at 3% annually. Model inputs, drawn from the published literature and national databases, included annual cost of duloxetine ($721–937); average pain reduction for duloxetine (17.5 points on the Western Ontario and McMaster Universities Osteoarthritis Index pain scale [0–100]), and likelihood of depression remission with duloxetine (27.4%). We considered 2 willingness-to-pay (WTP) thresholds of $50,000/QALY and $100,000/QALY. We varied parameters related to the PHQ-9 and the cost of duloxetine, efficacy, and toxicities to address uncertainty in model inputs.

Results. The screening strategy led to an additional 17 QALYs per 1,000 subjects and increased costs by $289/subject (ICER = $17,000/QALY). Universal duloxetine led to an additional 31 QALYs per 1,000 subjects and $1,205 per subject (ICER = $39,300/QALY). Under the majority of sensitivity analyses, universal duloxetine was cost-effective at the $100,000/QALY threshold.

Conclusion. The addition of duloxetine to usual care for knee OA patients with moderate pain, regardless of depressive symptoms, is cost-effective at frequently used WTP thresholds.

INTRODUCTION

Knee osteoarthritis (OA) is a prevalent, disabling, and costly condition that affects more than 14 million Americans and more than 263 million individuals worldwide (1,2). Comorbid anxiety and depression are common among individuals with OA, with up to 41% experiencing either 1 or both of these conditions (3), and place additional burdens on the health care system (4,5). Depression is associated with worse clinical outcomes, including more severe OA pain (6–8), lower quality of life (4,6,9), worse total knee replacement (TKR) outcomes (10), and increased opioid utilization (11). Despite evidence of its detrimental impact on OA outcomes,

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grants K24-AR-057827, R01-AR-074290, and P30-AR-72577 and National Institute of Mental Health grants R25-MH-094612 and K24-AR-070892).

[1]Nora K. Lenhard, BA, James K. Sullivan, BA, Shuang Song, DVM, MPH, Robert R. Edwards, PhD: Brigham and Women's Hospital, Boston, Massachusetts; [2]Eric L. Ross, MD: Massachusetts General Hospital, Boston, McLean Hospital, Belmont, and Harvard University, Boston, Massachusetts; [3]David J. Hunter, MBBS, MSc, PhD: University of Sydney and Royal North Shore Hospital, Sydney, Australia; [4]Tuhina Neogi, MD, PhD: Boston University, Boston, Massachusetts; [5]Jeffrey N. Katz, MD, MSc: Brigham and Women's Hospital and Harvard University, Boston, Massachusetts; [6]Elena Losina, PhD: Brigham and Women's Hospital, Harvard University, and Boston University, Boston, Massachusetts.

Dr. Hunter has received consulting fees from Pfizer, Merck Serono, and Eli Lilly and Company (less than $10,000 each). Dr. Neogi has received consulting fees from Pfizer, Eli Lilly and Company, EMD Serono, and Novartis (less than $10,000 each). Dr. Katz has received research support from Samumed and Flexion Therapeutics. Dr. Losina has received consulting fees from Pfizer (less than $10,000) and research support from Flexion Therapeutics and Pfizer. No other disclosures relevant to this article were reported.

Address correspondence to Elena Losina, PhD, Orthopedic and Arthritis Center for Outcomes Research, Brigham and Women's Hospital, 75 Francis Street, BTM 5016, Boston, MA 02115. Email: elosina@partners.org.

Submitted for publication June 12, 2020; accepted in revised form November 24, 2020.

SIGNIFICANCE & INNOVATIONS

- Depression is poorly managed in knee osteoarthritis (OA) populations and is associated with more severe pain and lower quality of life.
- Duloxetine is effective at reducing both OA pain and depressive symptoms, but there have been no studies evaluating the cost-effectiveness of incorporating duloxetine into knee OA care that account for the effect of duloxetine on depression.
- We used a validated microsimulation model of knee OA to evaluate the cost-effectiveness of adding duloxetine to usual knee OA care under 2 strategies: adding duloxetine only for subjects who screen positive for depressive symptoms on the Patient Health Questionnaire 9, or for all subjects, regardless of depressive symptoms.
- The depression screening strategy had an incremental cost-effectiveness ratio (ICER) of $16,961 per quality-adjusted life year (QALY), and the universal duloxetine strategy had an ICER of $39,288/QALY, indicating that incorporating duloxetine into usual knee OA care without depression assessment provides good value.

depression is not well-managed in this population; only 33% of adults with OA and depression receive adequate depression treatment (12).

Duloxetine, a serotonin-norepinephrine reuptake inhibitor that is approved by the US Food and Drug Administration to treat major depressive disorder and knee OA, is effective in treating depression and OA pain independent of depressive symptoms (13). Given the negative impact depression has on OA management, incorporating a treatment that could affect both conditions simultaneously could improve outcomes. Duloxetine may also alleviate some of the economic burden posed by medical costs associated with depression by leading to remission of this disease. However, while depression screening is recommended for all adults and may be especially important in this population (14), some studies raise questions about the feasibility and efficacy of incorporating screening into routine care (15). Additionally, rheumatologists and orthopedists may be reluctant to screen for depression due to time constraints or feeling that other providers would be better suited to depression management (16,17).

In this analysis, we aimed to examine if screening OA patients for depression and treating those who screen positive for depression with duloxetine therapy is a better value for pain management than offering duloxetine to knee OA patients regardless of depression assessment.

MATERIALS AND METHODS

Analytic overview. The Osteoarthritis Policy (OAPol) Model is a validated microsimulation model of the progressive natural history and management of knee OA (18–20). We used the OAPol Model to examine whether adding duloxetine to usual care in knee OA patients with moderate pain who no longer receive adequate pain relief from treatment with nonsteroidal anti-inflammatory drugs (NSAIDs) could be cost-effective and whether depression screening offers a way to optimize benefits and reduce adverse effects from duloxetine.

We simulated a cohort of knee OA patients with demographic and clinical characteristics similar to subjects in the Osteoarthritis Initiative (OAI). The primary outcomes were quality-adjusted life expectancy (QALE), lifetime medical costs, and incremental cost-effectiveness ratios (ICERs). We calculated ICERs as the difference in lifetime medical costs divided by the difference in QALE between 2 strategies. All costs were reported in 2018 USD, and analyses were conducted from a health care sector perspective, which includes costs paid by all payers including public, private, and individuals (21). Costs and quality-adjusted life years (QALYs) were discounted 3% annually (21). We considered 2 well-established willingness-to-pay (WTP) thresholds of $50,000/QALY and $100,000/QALY (22). Treatment strategies were considered cost-effective if they improved QALE and produced an ICER below the WTP threshold. Any strategy that increased cost and decreased QALE was described as "dominated."

To address uncertainty in our inputs, we performed sensitivity analyses by varying parameters on the Patient Health Questionnaire 9 (PHQ-9), the prevalence of depressive symptoms, and the cost, efficacy, and adverse effects of duloxetine. Model inputs and sensitivity analyses are described in Table 1.

The OAPol Model. The OAPol Model simulates cohorts of knee OA patients (model subjects) using prespecified distributions of demographic and clinical characteristics, including age, sex, race/ethnicity, body mass index (BMI), comorbidities (cardiovascular disease, cancer, diabetes, and inflammatory arthritis), and structural and symptomatic severity of knee OA. Using a Monte Carlo simulation, subjects transition through different health states that include changes to knee OA structural and symptomatic severity, BMI classes, and comorbidities. Each annual model cycle, each subject accrues QALYs, and the sum of these QALYs (measured by the time between the subject's initialization and death) averaged across all subjects is the QALE. Symptomatic severity is defined by pain ratings on the 100-point Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) scale, and ratings are split into 5 pain groups (<1, 1–15, 16–40, 41–70, and 71–100). Changes in these health states are associated with quality of life (QoL) changes, and subjects incur costs related to OA management and background (non-OA) medical costs each year. Further details about the OAPol Model have been published (18–20).

Table 1. Cohort and treatment characteristics and model input parameters with varied sensitivity analyses*

	Cohort with depressive symptoms	Cohort without depressive symptoms	WOMAC pain in obese patients	WOMAC pain in nonobese patients	Model input value	Average first year pain decrement, mean ± SD (13)	Probability of pain failure in subsequent years, % (34)	Values or range of change	Data source
Cohort characteristics†									
Age, mean ± SD years	60 ± 9	60 ± 9	–	–	–	–	–	–	Average age of those with depressive symptoms in OAI (23)
Female sex, %	74.2	60.9	–	–	–	–	–	–	OAI (23)
White Non-Hispanic, %	62.1	80.1	–	–	–	–	–	–	
WOMAC pain, no. (%)	40.1 (22.8)	24.7 (19.4)	–	–	–	–	–	–	
BMI, kg/m^2	31.4 (4.8)	29.6 (4.1)	–	–	–	–	–	–	
QoL weights by age, obesity, and WOMAC pain for cohort without depressive symptoms and with no comorbidities, 0–1 scale‡									Derivations of a model by Brazier et al (2004; 51) and data from OAI (23)
Age group, years									
45–54 years									
WOMAC 1–15	–	–	0.814	0.825	–	–	–	–	
WOMAC 16–40	–	–	0.778	0.789	–	–	–	–	
WOMAC 40–70	–	–	0.712	0.723	–	–	–	–	
WOMAC 71–100	–	–	0.654	0.664	–	–	–	–	
55-64 years									
WOMAC 1–15	–	–	0.820	0.831	–	–	–	–	
WOMAC 16–40	–	–	0.784	0.795	–	–	–	–	
WOMAC 40–70	–	–	0.718	0.729	–	–	–	–	
WOMAC 71–100	–	–	0.660	0.670	–	–	–	–	
65–74 years									
WOMAC 1–15	–	–	0.844	0.855	–	–	–	–	
WOMAC 16–40	–	–	0.808	0.819	–	–	–	–	
WOMAC 40–70	–	–	0.742	0.753	–	–	–	–	
WOMAC 71–100	–	–	0.683	0.694	–	–	–	–	
75+ years									
WOMAC 1–15	–	–	0.827	0.838	–	–	–	–	
WOMAC 16–40	–	–	0.791	0.802	–	–	–	–	
WOMAC 40–70	–	–	0.725	0.736	–	–	–	–	
WOMAC 71–100	–	–	0.666	0.677	–	–	–	–	

(Continued)

Table 1. (*Cont'd*)

	Cohort with depressive symptoms	Cohort without depressive symptoms	WOMAC pain in obese patients	WOMAC pain in nonobese patients	Model input value	Average first year pain decrement, mean ± SD (13)	Probability of pain failure in subsequent years, % (34)	Values or range of change	Data source
Background medical costs for nondepressed subjects with 0–1 comorbidity§									Pope et al (2004; 52) MCBS (27) NHANE (2015–2016) Red Book Online (26–28, 38)
Age group, years									
55–59	–	–	–	–	–	–	–	$5,199	
60–64	–	–	–	–	–	–	–	$5,741	
65–69	–	–	–	–	–	–	–	$5,587	
70–74	–	–	–	–	–	–	–	$6,244	
75–79	–	–	–	–	–	–	–	$7,138	
+80	–	–	–	–	–	–	–	$10,342	
Duloxetine treatment characteristics									
Annual cost (first year)	–	–	–	–	$721	–	–	–	Red Book Online
Annual cost (subsequent years)	–	–	–	–	$937	–	–	–	Medicare Physician Fee Schedule, Annual Cost of Dispensing Study (38–40)
Treatment-related toxicity probability	–	–	–	–	46%	–	–	–	Nelson et al (2006; 37)
Likelihood of discontinuation in first year	–	–	–	–	23%	–	–	–	Chappell et al (2011; 13)
Duloxetine pain efficacy									
WOMAC pain scale									
1–15	–	–	–	–	–	8 ± 4	24	–	
16–40	–	–	–	–	–	14 ± 5	24	–	
41–70	–	–	–	–	–	18 ± 6	50	–	
70–100	–	–	–	–	–	21 ± 6	75	–	
Duloxetine depression efficacy									
Likelihood of depressive symptoms remission	–	–	–	–	27%	–	–	–	Raskin et al (2007; 35)
Likelihood of depression relapse in subsequent years	–	–	–	–	44%	–	–	–	IMPACT trial (36)

(Continued)

Table 1. (*Cont'd*)

	Cohort with depressive symptoms	Cohort without depressive symptoms	WOMAC pain in obese patients	WOMAC pain in nonobese patients	Model input value	Average first year pain decrement, mean ± SD (13)	Probability of pain failure in subsequent years, % (34)	Values or range of change	Data source
Sensitivity analyses									
Parameter varied									
Duloxeine depression efficacy	–	–	–	–	–	–	–	21–33%; 0%	
Duloxetine pain efficacy	–	–	–	–	–	–	–	6–17 points on WOMAC, stratified by pain group; 9–23 points (CIs of data from Chappell et al [13])	
Cost of duloxetine	–	–	–	–	–	–	–	$444–1,555	
Treatment-related toxicity probability	–	–	–	–	–	–	–	56.4%	
Increased complexity of physician visit	–	–	–	–	–	–	–	$152 for first visit (HCPCS code 99214) (40) compared to $120 (base case; HCPCS code 99213)	
PHQ-9 sensitivity	–	–	–	–	–	–	–	40.7–81.3% (42)	
PHQ-9 specificity	–	–	–	–	–	–	–	42.7–85.3% (42)	
Prevalence of depressive symptoms	–	–	–	–	–	–	–	0–23% (27)	

* Values are the number (%) unless indicated otherwise. Costs were calculated in 2018 USD. Among subjects with depressive symptoms and those without, 5.95% of subjects with depressive symptoms reported strong opioid utilization; for subjects without depressive symptoms, 3.61% reported strong opioid utilization. Data derived from the Medicare Current Beneficiary Survey (MCBS), 2015–2016 (27). BMI = body mass index; CIs = confidence intervals; HCPCS = Healthcare Common Procedure Coding System; IMPACT = Improving Mood Promoting Access to Collaborative Care Treatment trial; NHANES = National Health and Nutrition Examination Survey; OAI = Osteoarthritis Initiative; PHQ-9 = Patient Health Questionnaire 9; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

† Prevalence of depressive symptoms was 23%. MCBS, 2015–2016 (27).

‡ Values are for subjects with no comorbid conditions; additional comorbid conditions carried a quality of life (QoL) decrement. QoL decrement due to depression was calculated to be 0.0625. Sullivan et al (29).

§ Additional cost for subjects with depressive symptoms was $1,081. MCBS, 2015–2016 (27).

For each treatment regimen, a pain decrement is determined from distributions reported in the published literature. Once a subject no longer receives pain relief from a treatment method, they are evaluated for the subsequent regimen (Figure 1). A subject may also move to the next regimen due to voluntary discontinuation or experiencing a toxicity that would cause their physician to discontinue that treatment. The cost of each regimen includes medication (if relevant), procedures, physician visits, and medical care for toxicities.

Cohort characteristics. The cohort characteristics were based on subjects with knee OA enrolled in the OAI (23). Using

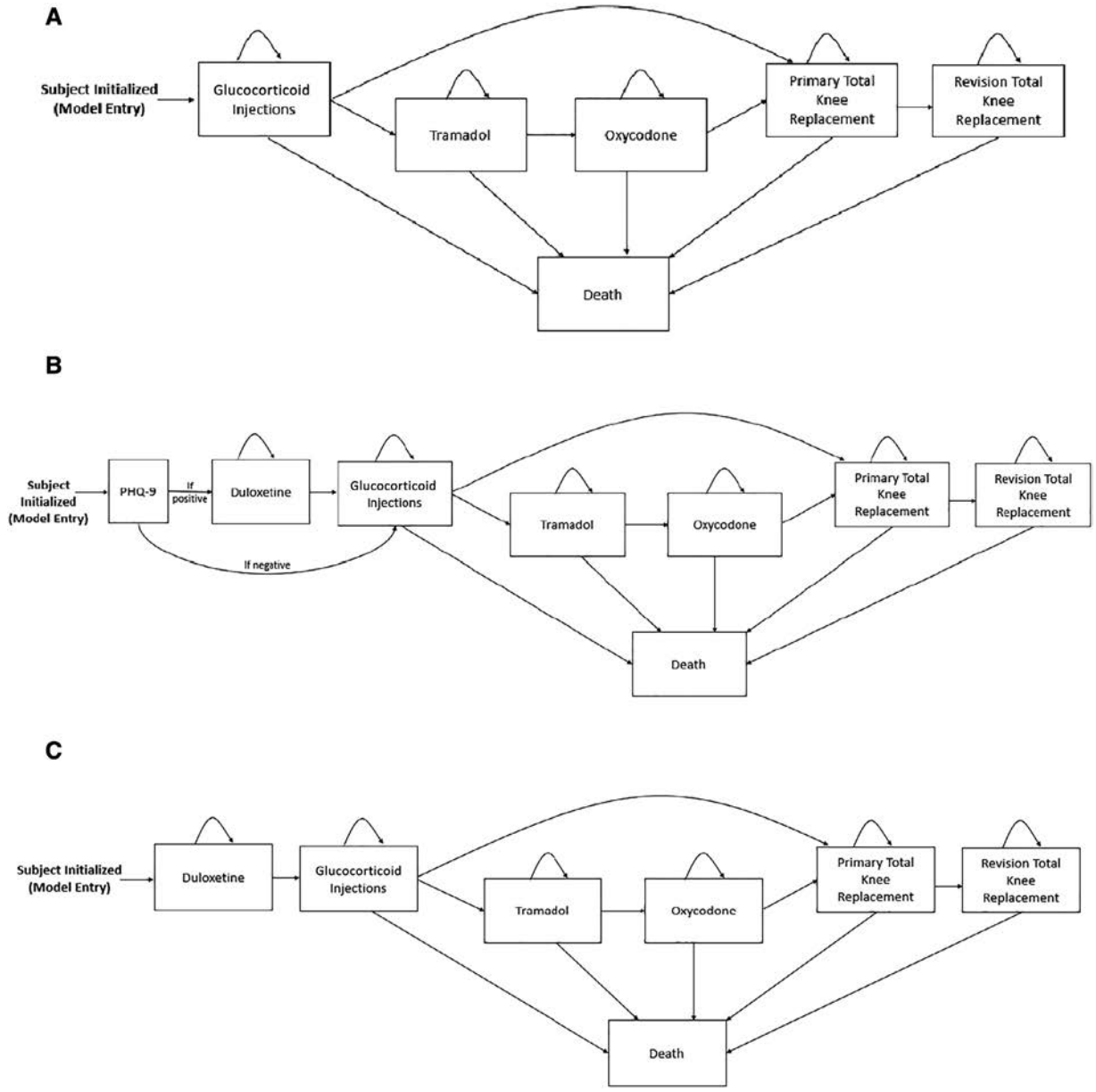

Figure 1. Treatment strategies. **A**, Usual care treatment sequence for knee osteoarthritis (OA) in the Osteoarthritis Policy (OAPol) Model. Subjects began this treatment sequence based on specific cohort characteristics and based on progress throughout the regimens outlined, which included glucocorticoids, tramadol, oxycodone (for some subjects only, as some individuals prefer to avoid opioid regimens), total knee replacement, and revision total knee replacement. **B**, Depression screening treatment sequence for knee OA in the OAPol Model. Treatment was initialized in subjects based on specific cohort characteristics, and subjects were screened for depression using the Patient Health Questionnaire 9 (PHQ-9) at treatment initialization. If a subject screened positive for depressive symptoms, they would receive duloxetine before usual care. If a subject screened negative for depressive symptoms, they would proceed directly to usual care. **C**, Universal duloxetine treatment sequence for knee OA in the OAPol Model. Treatment was initialized in subjects based on specific cohort characteristics, and subjects would receive duloxetine before progressing through the rest of the usual care treatments. In all treatment strategies outlined here, subjects continued receiving each treatment until it was no longer effective, and death could occur at any point in a given sequence.

OAI subjects' responses on the Center for Epidemiological Studies Depression Scale, we modeled the demographic characteristics of two cohorts: a cohort with depressive symptoms and one without depressive symptoms. We chose to simulate a cohort not currently receiving mental health care, as the population represented in this analysis are those for whom duloxetine would not replace ongoing depression treatments. We considered knee OA patients that had not achieved adequate pain relief from NSAIDs, physical therapy, and lifestyle modifications (24,25).

Both cohorts (one with depressive symptoms and one without depressive symptoms) had a mean ± SD starting age of 60 ± 9 years and a mean BMI of 31 kg/m^2 and 30 kg/m^2, respectively. Mean ± SD starting pain on the 100-point WOMAC scale was 40 ± 23 for those with depressive symptoms and 25 ± 19 for those without depressive symptoms (23). These OAI pain data are consistent with studies demonstrating that individuals with depression experience greater pain (6,8). The prevalence and incidence of the comorbidities were derived from the 2014–2016 National Health and Nutrition Examination Survey (NHANES) and were stratified by age, race, and sex (26). Background medical costs were derived from NHANES (26) and the Medicare Current Beneficiary Survey (MCBS) (27), and we used a risk-adjustment model developed by Pope et al to stratify costs by age and number of comorbidities (28). QoL values were derived from the OAI and were stratified by age and number of comorbidities (23).

Depression-related inputs. We used data from two depression screening questions on the MCBS to estimate the prevalence of individuals with depressive symptoms not receiving treatment for depression. Of the MCBS cohort, 23% of individuals responded in the affirmative to a question about feeling sad, "blue," or depressed or having experienced a period of 2 or more weeks where they lost interest or pleasure in valued activities in the preceding year and were not receiving treatment for depressive symptoms (27). Note that the term "depressive symptoms" is used to characterize these subjects' medical history, as the screening questionnaire did not provide enough information to determine whether they meet the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition criteria for major depressive disorder.

We applied a QoL decrement of 0.0625 to those with depressive symptoms (29). Depressed individuals utilize the health care system at higher rates (5), and those who responded in the affirmative to MCBS depression screening questions and were not receiving treatment for depression incurred an additional $1,081 in medical costs compared to those with a negative depression screening (adjusted for number of comorbidities), which we added to the background medical costs for subjects with depressive symptoms (27). Mortality rates were estimated using Centers for Disease Control and Prevention 2014 US life tables (30). We estimated a 58% increased risk of all-cause mortality in those with depressive symptoms (31).

Treatment strategies. We considered 3 treatment strategies (Figure 1). The first treatment strategy was defined as "usual care," in which subjects were treated with a sequence of glucocorticoid injections, tramadol, oxycodone, TKR, and revision TKR. A second treatment strategy, which we characterized as a "screening" approach, involved subjects being screened for depressive symptoms using the PHQ-9 and receiving therapy with duloxetine if positive on screening; if negative on screening, or if a subject's pain and depressive symptoms were inadequately controlled with duloxetine therapy, subjects received the usual care treatment sequence. The third treatment strategy was "universal duloxetine," in which all subjects received duloxetine therapy; in this treatment arm, if duloxetine failed to control symptoms, subjects were switched to the usual care treatment sequence.

We calibrated the proportion of subjects receiving strong opioids to reflect data from MCBS, stratified by depressive symptom status (Table 1) (27). The pain reduction for TKR was derived from the AViKA cohort and stratified by depressive symptom status (32). Further details on usual care regimens have been previously published (19,33).

Duloxetine characteristics. *Efficacy.* We derived the pain efficacy of duloxetine from a 13-week randomized controlled trial that reported a mean decrease on the WOMAC of 17.5 points (100-point scale) among those receiving duloxetine (13). We stratified this pain decrement by subjects' initial pain group (Table 1). Since this trial did not report the durability of pain relief after 12 months of treatment, we assumed that sustainability was similar to that reported in NSAID trials (with 24–75% of subjects experiencing pain failure while receiving NSAIDs, stratified by pain at regimen initiation) (34).

To derive the efficacy of duloxetine in the reduction of depressive symptoms ("depression efficacy"), we used a randomized trial of duloxetine for depression treatment in adults ages 65 years or older (35). Of these trial participants, 27.4% achieved remission of depressive symptoms, measured on the Hamilton Depression Scale. This trial also did not report long-term efficacy, so we used durability data from the Improving Mood Promoting Access to Collaborative Care Treatment trial, a collaborative care intervention for elderly patients with depression in which 73% of participants received therapy with antidepressants (36). Of those who achieved remission of depression at 12 months, 44% experienced a relapse in depressive symptoms at 24 months. Those who achieved remission of depressive symptoms received a QoL increase of 0.0625 (29).

Toxicity. We derived toxicities from a pooled analysis of four clinical trials (37). Only adverse effects that occurred at rates significantly different from placebo were used in our analysis, which included nausea, constipation, insomnia, dry mouth, somnolence, sweating, and fatigue. Toxicities were associated with QoL decreases, and in some cases, additional treatment costs

Table 2. Results of base case and selected scenario sensitivity analyses*

Treatment sequence	QALE	Lifetime medical costs	ICER
Base case			
Usual care	11.0458	$186,914	–
Depression screening	11.0628	$187,203	$16,961/QALY
Universal duloxetine	11.0935	$188,408	$39,288/QALY
Sensitivity analyses†			
Usual care	11.0458	$186,914	–
Depression screening	11.0632	$187,325	$23,528/QALY
Universal duloxetine	11.0940	$188,558	$40,076/QALY
Increased visit complexity for first physician visit			
Usual care	11.0458	$186,914	–
Depression screening	11.0628	$187,213	$17,515/QALY
Universal duloxetine	11.0935	$188,440	$40,007/QALY
Increased toxicity rate			
Usual care	11.0458	$186,914	–
Depression screening	11.0588	$187,202	$22,210/QALY
Universal duloxetine	11.0782	$188,373	$60,271/QALY
No pain efficacy and decreased depression efficacy of duloxetine			
Usual care	11.0458	$186,914	–
Depression screening	11.0448	$186,959	Dominated
Universal duloxetine	11.0382	$187,155	Dominated
PHQ-9 sensitivity and specificity at 50% of base‡			
Usual care	11.5714	$191,060	–
Depression screening	11.5958	$192,046	$40,330/QALY
Universal duloxetine	11.6151	$192,787	$38,430/QALY

* ICER = incremental cost-effectiveness ratio; PHQ-9 = Patient Health Questionnaire 9; QALE = quality-adjusted life expectancy; QALY = quality-adjusted life year.
† Increased background medical costs sustained for subjects with depressive symptoms in remission.
‡ Prevalence of depressive symptoms was 5.75%.

(see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24519/abstract). The likelihood of experiencing one of these toxicities was 46%. Of those who experienced a toxicity, 37.5% discontinued the duloxetine regimen (13). Overall, 23% of those receiving duloxetine stopped taking this medication, with 17.4% discontinuing duloxetine due to a toxicity and 5.6% discontinuing duloxetine for other reasons (13).

Cost. We derived the annual cost of a 60-mg daily dose of generic duloxetine from Red Book Online in October 2018 (38). The average wholesale acquisition cost for 30-unit packages was used, which was $444/year (38). In addition to the pharmaceutical cost, we included a monthly dispensing fee (as not all insurance plans cover 3-month prescriptions) (39), as well as the cost of 3 physician visits, the price of which was derived from the 2018 Medicare Physician Fee Schedule (40). We assumed that those who discontinued duloxetine did so in the first year of receiving duloxetine, and we reduced the first-year cost of the regimen accordingly. The annual cost of the treatment regimen was $721 in the first year and $937 in subsequent years (Table 1).

Depression screening characteristics. For the depression screening strategy, we used characteristics of the PHQ-9, a validated, self-administered, 9-question form designed to screen for depression during primary care and similar settings (41). The sensitivity and specificity of the PHQ-9 has been estimated at 81.3% and 85.3%, respectively (42).

Sensitivity analyses. We varied input parameters using 1-way and scenario-based deterministic sensitivity analyses and a probabilistic sensitivity analysis to examine the robustness of the results.

Deterministic sensitivity analyses. We varied parameters related to the cost, toxicities, and efficacy of duloxetine as well as the parameters of the PHQ-9 and prevalence of depressive symptoms. We conducted a "tipping point" analysis for the cost of duloxetine, in which we increased the pharmaceutical cost and determined the prices at which the 2 strategies crossed WTP thresholds. We also conducted a scenario-based sensitivity analysis in which we classified the first physician visit as more complex, and thus more expensive. Additionally, we performed a sensitivity analysis in which subjects who achieved remission of depressive symptoms continued to incur the increased background medical costs associated with depressive symptoms.

Additional studies of duloxetine have found that sexual dysfunction and falls may occur at higher rates among those taking duloxetine. Nelson et al showed that treatment-emergent sexual dysfunction occurred in 46.4% of individuals receiving duloxetine compared to 28.8% of those receiving placebo (37). Sullivan and colleagues modeled treatment for these toxicities as 2 physician

visits and a month's prescription of sildenafil for 25% of those who experienced these toxicities (43); we derived these costs from the Medicare Physician Fee Schedule and Red Book Online (38,40). Nelson et al reported that falls occurred in 17.3% of individuals receiving duloxetine therapy compared to 11.6% of those receiving placebo (44).

We also modeled the risk of fracture due to falls. Soh and colleagues found that the proportion of those who self-reported a fracture was 23% of the number that self-reported a fall (45). We applied this probability to the increased fall risk among those receiving duloxetine. With these additional toxicities, the overall toxicity probability was 56.4%, and 25.8% of those taking duloxetine discontinued the medication in the first year of treatment. Data on costs and QoL decrements for these toxicities are presented in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24519/abstract.

To examine combinations of some of the least favorable variations, we conducted scenario-based sensitivity analyses of selected reductions in the efficacy of duloxetine in controlling pain and depression. We used the lower 95% confidence intervals from the respective trials and included analyses of duloxetine having no efficacy in controlling depression or pain. We also examined scenarios in which the PHQ-9 sensitivity and specificity were 50% of base case values and the prevalence of depressive symptoms was 25% of base case values.

Probabilistic sensitivity analysis. We performed a probabilistic sensitivity analysis drawing from distributions of pain efficacy, depression efficacy, toxicity rates, PHQ-9 sensitivity, and depressive symptom prevalence. Five hundred iterations of probabilistic inputs were drawn and run through the model. A normal distribution was used for the pain decrements, and beta distributions were used for the toxicities, depression efficacy, prevalence of depressive symptoms, and PHQ-9 sensitivity. We constructed a cost-effectiveness acceptability curve to depict the proportion of time each strategy was the preferred strategy, or that which produced the greatest QALE while maintaining the ICER below a certain WTP threshold, under several WTP thresholds.

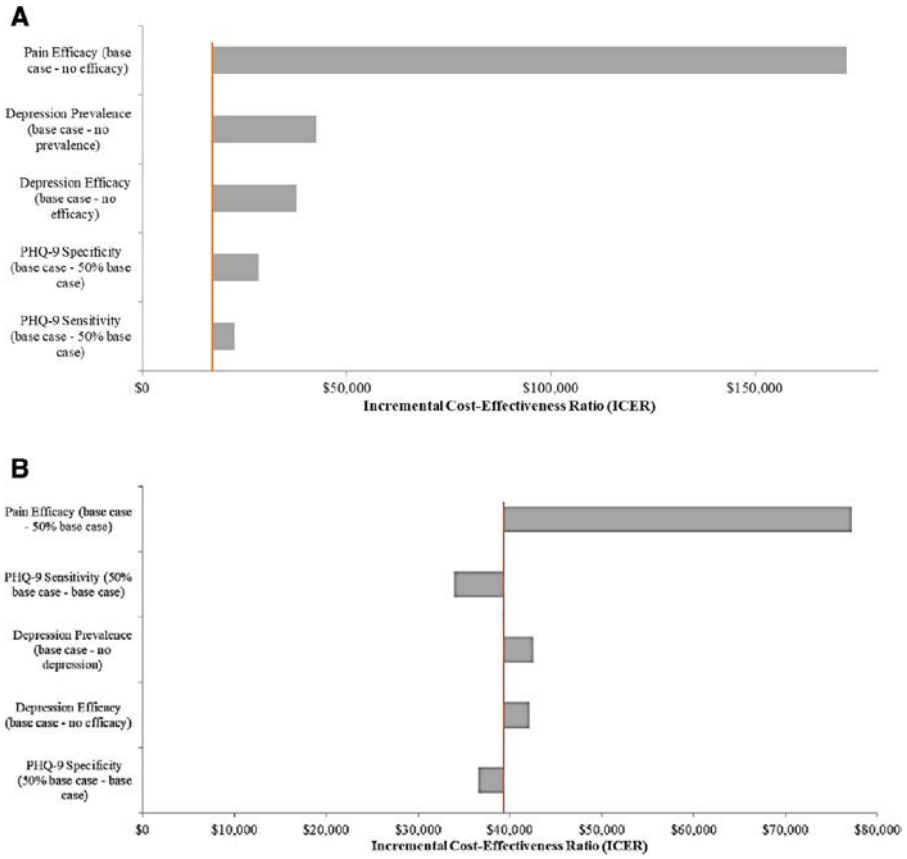

Figure 2. Results of one-way sensitivity analyses. **A** and **B**, Univariate sensitivity analysis of parameters related to duloxetine, prevalence of depressive symptoms, and Patient Health Questionnaire 9 (PHQ-9) scores used in the screening strategy (**A**) and universal duloxetine strategy (**B**). Incremental cost-effectiveness ratios for the treatment strategies are presented in relation to variations in the efficacy of duloxetine in controlling pain and depression, prevalence of depressive symptoms, and sensitivity and specificity of the PHQ-9. Orange line represents the base case. In **A**, all parameters were held at base case values except for the parameter listed on the vertical axis, which was varied according to the values listed. In **B**, parameters presented are the same as those presented for the screening strategy, with the exception that the lowest value of pain efficacy presented was 50% of base case, as the no pain efficacy scenario was dominated. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24519/abstract.

RESULTS

Base case. The average duration of duloxetine use was 2.4 years for subjects with depressive symptoms and 3.2 years for those without symptoms. The screening strategy resulted in a cost increase of $289 per subject over usual care and an increase of 17 QALYs per 1,000 subjects, which produced an ICER of $16,961/QALY (Table 2). Compared to the screening strategy, the universal duloxetine strategy led to a cost increase of $1,205 per subject and an increase of 31 QALYs per 1,000 subjects, resulting in an ICER of $39,288/QALY.

Sensitivity analyses. *One-way sensitivity analyses.* Results of the 1-way sensitivity analyses are presented in Figure 2. Pain efficacy had the greatest impact on the findings; for the screening strategy, duloxetine providing no pain efficacy resulted in an ICER of $172,455/QALY, and universal duloxetine was dominated by the screening strategy. At a 50% efficacy rate of controlling pain, the screening strategy and universal duloxetine strategy had ICERs of $23,417/QALY and $77,224/QALY, respectively. For the other parameters varied (depression efficacy, prevalence of depressive symptoms, and the sensitivity and specificity of the PHQ-9), ICERs for both the screening strategy and universal duloxetine strategy remained below $50,000/QALY for the most conservative values tested, including when the depressive symptom prevalence was zero.

Figure 3 presents the ICERs for the tipping point sensitivity analysis of the annual cost of duloxetine. We varied the cost under base case parameters and with additional toxicities. The screening strategy with base case toxicities resulted in ICERs that crossed the $50,000/QALY WTP threshold between 2.5 and 3 times the base case cost of duloxetine. The universal duloxetine strategy crossed this threshold between 1.25 and 1.5 times the base case cost. Considering the $100,000/QALY WTP threshold, the screening strategy was cost-effective at all values tested, and the universal duloxetine strategy crossed this threshold between 3 and 3.5 times the base case cost of duloxetine. Under the increased toxicity scenario, which included fracture and sexual dysfunction, the screening strategy crossed the $50,000/QALY WTP threshold between 2 and 2.5 times the base case cost. The universal duloxetine strategy crossed the $100,000/QALY threshold between 1.75 and 2 times the base case cost of duloxetine.

Scenario-based sensitivity analyses. Maintaining increased background medical costs during depressive symptom remission resulted in ICERs of $23,528/QALY and $40,076/QALY for the screening and universal duloxetine strategies, respectively. Additional toxicities increased the ICERs to $22,210/QALY and $60,271/QALY for the 2 strategies. Increasing the cost of the first physician visit, during which duloxetine would be prescribed, had minimal impact (Table 2).

We varied the efficacy of duloxetine in reducing pain and depression together using the lower 95% confidence intervals reported by previous trials (13,35). We also considered extreme scenarios in which duloxetine had no efficacy in controlling pain or depressive symptoms. Under a conservative scenario (decreased efficacy in controlling pain and no efficacy in reducing depressive symptoms), the screening strategy ICER was $51,204/QALY, and the universal duloxetine strategy ICER was $53,980/QALY. When efficacy in reducing pain was assumed to

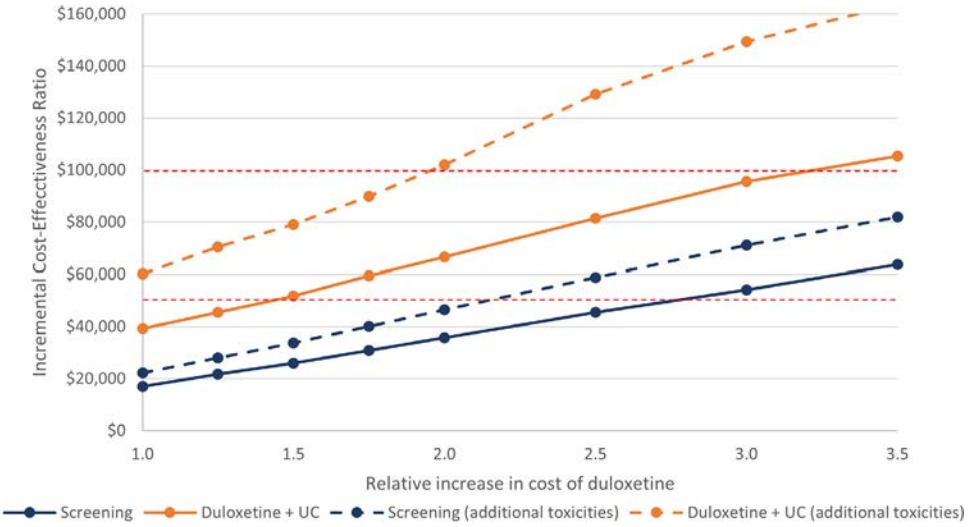

Figure 3. Tipping point sensitivity analysis of the cost of duloxetine. The incremental cost-effectiveness ratios (ICERs) for depression screening and universal duloxetine are presented in relation to relative increases in the cost of an annual prescription of 60 mg of generic duloxetine. Solid lines represent base case toxicity values; broken lines represent increased toxicity parameters, which includes an increased risk of falls and sexual dysfunction. Red lines represent the $50,000 per quality-adjusted life year (QALY) and $100,000/QALY willingness-to-pay thresholds. UC = usual care. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24519/abstract.

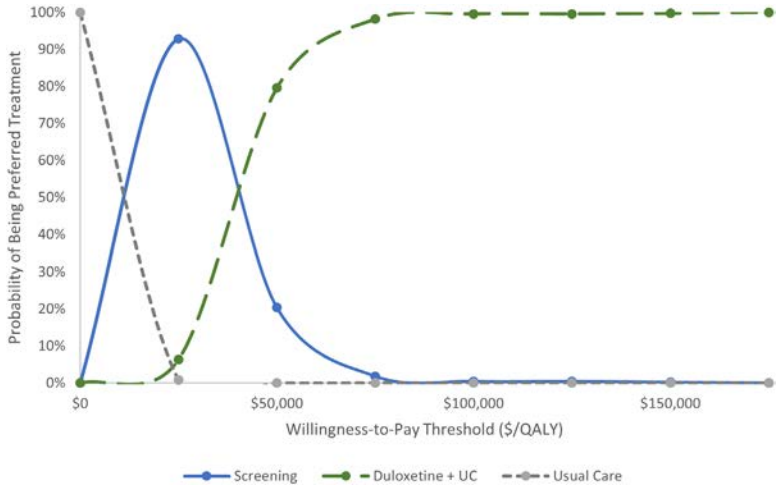

Figure 4. Probabilistic sensitivity analyses were conducted varying toxicity rates, pain and depression efficacy of duloxetine, sensitivity of the Patient Health Questionnaire 9, and prevalence of depressive symptoms. Five hundred iterations of probabilistic inputs were run. Incremental cost-effectiveness ratios (ICERs) were calculated for each iteration comparing treatment with usual care (UC) (gray broken line), the depression screening strategy (blue line), and the universal duloxetine strategy (green line). ICERs for these strategies were compared to a range of willingness-to-pay (WTP) thresholds, and the treatment strategy which produced the greatest quality-adjusted life expectancy while remaining below the WTP threshold was termed the preferred strategy. The probability of being the preferred treatment strategy is plotted against various WTP thresholds. QALY = quality-adjusted life year.

be zero and efficacy in controlling depressive symptoms was reduced, both strategies were dominated by usual care.

To examine less favorable variations of screening parameters, we reduced the sensitivity and specificity of the PHQ-9 to 50% of the base case values and the prevalence of depressive symptoms to 5.75% (25% of the base case values). Under this unfavorable scenario, the screening strategy ICER was $40,330/QALY, and the universal duloxetine ICER was $38,430/QALY (Table 2).

Probabilistic sensitivity analyses. Figure 4 shows the results of the probabilistic sensitivity analysis. The cost-effectiveness acceptability curve depicts the proportion of times each strategy was the preferred option at willingness-to-pay thresholds from $0 to $150,000/QALY. At a WTP threshold of $25,000/QALY, the screening strategy was the preferred strategy in 93% of cases, and universal duloxetine was the preferred treatment in 6% of cases. Universal duloxetine was the preferred strategy in 80% of cases at the $50,000/QALY WTP threshold. At the $100,000/QALY threshold, universal duloxetine was the preferred treatment option 100% of the time.

DISCUSSION

We used a widely published, validated computer simulation model to evaluate the cost-effectiveness of incorporating duloxetine into knee OA care for patients with moderate pain under 2 strategies: 1) the addition of duloxetine to usual care for individuals with depressive symptoms only; or 2) the incorporation of duloxetine without depression screening for individuals with knee OA. Base case results suggest that adding duloxetine to usual care without depression screening would maximize QALE while keeping ICERs under $50,000/QALY. Results from the probabilistic sensitivity analysis reinforced the value of the universal duloxetine strategy; it was the preferred strategy in 80% of iterations at the $50,000/QALY threshold.

As pharmaceutical costs vary across health systems, pharmacies, and payers, we conducted a tipping point analysis to determine at what cost the strategies crossed WTP thresholds. According to data reported on the prices that Medicare plans pay for duloxetine, 75% of plans pay less than $537/year (46), which falls between our base case ($444/year) and 1.25 times the cost of base case, indicating that the universal duloxetine strategy is likely to be cost-effective at prices paid by federal insurers. Additionally, the base case cost was greater than that reported in the Federal Supply Schedule, which is the source recommended by the Second Panel on Cost-Effectiveness for cost-effectiveness analyses (21). We chose a more conservative base case cost estimate and found universal duloxetine to be cost-effective even above this cost.

Even under a variety of conservative scenarios surrounding the efficacy and toxicities of duloxetine, the ICERs for universal duloxetine did not exceed $77,300/QALY—except for the scenario in which duloxetine had no pain efficacy. We added fracture and sexual dysfunction as additional toxicities of the drug, and while they resulted in less favorable ICERs (>$50,000/QALY), ICERS remained below the $100,000/QALY threshold. Similarly, decreasing the pain and depression efficacy of duloxetine led to somewhat less favorable results, though even with decreased pain efficacy and no depression efficacy, the ICER was $53,980/QALY. The only scenario under which universal duloxetine was

not cost-effective was when it provided no pain efficacy. The results were robust to a variety of conservative analyses, with the greatest ICER of the sensitivity analyses—except for the scenario in which duloxetine did not affect pain—being $77,224/QALY.

The universal duloxetine strategy has an advantage of being a simple one to implement. While improving depressive symptoms may increase the efficacy of OA treatments, rheumatologists and orthopedists may be hesitant to make OA-related clinical decisions based on depressive symptoms in a patient (16,17). The dual benefits of duloxetine may improve outcomes in knee OA patients, and providing this drug therapy to broader populations, rather than targeting individuals with depressive symptoms, can add value to knee OA care. The cost-effectiveness of universal duloxetine was driven more by its pain-relieving properties than its effect on depressive symptoms, as duloxetine was a cost-effective treatment even in a population without depressive symptoms (Figure 2), underscoring the benefit of not restricting its use to individuals with depressive symptoms.

To the best of our knowledge, this is the first study to examine the cost-effectiveness of duloxetine for knee OA that accounts for the efficacy of duloxetine in treating pain and depression. Wielage and colleagues found that duloxetine was a cost-effective treatment for knee OA (47), though this analysis did not account for any effect on depressive symptoms and derived QoL coefficients using a dose-response relationship between pain and QoL, which may overstate treatment effects. We addressed these 2 limitations in this analysis. Duloxetine has shown potential for being cost-effective for other chronic conditions, including fibromyalgia and chronic low back pain (48,49), though these analyses similarly did not account for depression efficacy. Our results build on Wielage et al's findings that duloxetine may provide good value in knee OA therapy in which other treatments, such as opioids, have not been shown to be cost-effective (33). Duloxetine may add to a limited range of cost-effective oral analgesics, including NSAIDs (50), that can be used before surgical intervention.

There are several limitations of this study. The trials used to derive efficacy and toxicities associated with duloxetine were relatively short-term in duration, and we had to make assumptions about the likelihood of treatment failure after 1 year. Chappell et al's trial showed WOMAC decrease in terms of the total WOMAC score, whereas we modeled the WOMAC pain subscale in our study (13). These scales are correlated, and we assumed them to be comparable, though not equivalent. Studies on diagnosed depression were also used to derive some data for our cohort with depressive symptoms. Additionally, model subjects can only be on one regimen at a time and move sequentially from one to another, though when in practice, individuals may combine or cycle through treatments. We modeled depressive symptoms as a binary state and were unable to account for partial improvement of symptoms. Sensitivity analyses were conducted to address these uncertainties, but further clinical studies could clarify these characteristics of duloxetine and provide longer-term data to model cost-effectiveness with more certainty. Last, prior studies highlight the influence of subjects' pain at regimen initiation on the cost-effectiveness of a pharmaceutical agent (19), thus these findings may have limited generalizability to populations with different characteristics, such as pain.

Incorporating the dual efficacy of duloxetine in treating pain and depressive symptoms offers a better understanding of the potential value of this therapy. Given the economic burden that depression and knee OA place on the health care system and the prevalence of inadequately treated depression in this population (12), identifying treatments that can address these concerns together is valuable. This work provides evidence that, even without screening for depressive symptoms, the introduction of duloxetine after NSAIDs are unsuccessful in providing relief to knee OA patients with moderate pain offers good value as a pain management option.

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

Study conception and design. Lenhard, Sullivan, Edwards, Losina.
Acquisition of data. Lenhard, Sullivan, Song, Losina.
Analysis and interpretation of data. Lenhard, Sullivan, Ross, Edwards, Hunter, Neogi, Katz, Losina.

REFERENCES

1. Deshpande BR, Katz JN, Solomon DH, Yelin EH, Hunter DJ, Messier SP, et al. Number of persons with symptomatic knee osteoarthritis in the US: Impact of race and ethnicity, age, sex, and obesity. Arthritis Care Res (Hoboken) 2016;68:1743–50.
2. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;398:1789–858.
3. Axford J, Butt A, Heron C, Hammond J, Morgan J, Alavi A, et al. Prevalence of anxiety and depression in osteoarthritis: use of the Hospital Anxiety and Depression Scale as a screening tool. Clin Rheumatol 2010;29:1277–83.
4. Rosemann T, Gensichen J, Sauer N, Laux G, Szecsenyi J. The impact of concomitant depression on quality of life and health service utilisation in patients with osteoarthritis. Rheumatol Int 2007;27:859–63.
5. Simon GE, VonKorff M, Barlow W. Health care costs of primary care patients with recognized depression. Arch Gen Psychiatry 1995;52: 850–6.
6. Sharma A, Kudesia P, Shi Q, Gandhi R. Anxiety and depression in patients with osteoarthritis: impact and management challenges. Open Access Rheumatol 2016;8:103–13.
7. Kim KW, Han JW, Cho HJ, Chang CB, Park JH, Lee JJ, et al. Association between comorbid depression and osteoarthritis symptom severity in patients with knee osteoarthritis. J Bone Joint Surg Am 2011;93:556–63.
8. Rosemann T, Laux G, Szecsenyi J, Wensing M, Grol R. Pain and osteoarthritis in primary care: factors associated with pain perception in a sample of 1,021 patients. Pain Med 2008;9:903–10.

9. Sambamoorthi U, Shah D, Zhao X. Healthcare burden of depression in adults with arthritis. Expert Rev Pharmacoecono Outcomes Res 2017;17:53–65.
10. Lopez-Olivo MA, Landon GC, Siff SJ, Edelstein D, Pak C, Kallen MA, et al. Psychosocial determinants of outcomes in knee replacement. Ann Rheum Dis 2011;70:1775–81.
11. Wright EA, Katz JN, Abrams S, Solomon DH, Losina E. Trends in prescription of opioids from 2003-2009 in persons with knee osteoarthritis. Arthritis Care Res (Hoboken) 2014;66:1489–95.
12. Harman JS, Edlund MJ, Fortney JC, Kallas H. The influence of comorbid chronic medical conditions on the adequacy of depression care for older Americans. J Am Geriatr Soc 2005;53:2178–83.
13. Chappell AS, Desaiah D, Liu-Seifert H, Zhang S, Skljarevski V, Belenkov Y, et al. A double-blind, randomized, placebo-controlled study of the efficacy and safety of duloxetine for the treatment of chronic pain due to osteoarthritis of the knee. Pain Pract 2011;11: 33–41.
14. Guglielmo D, Hootman JM, Boring MA, Murphy LB, Theis KA, Croft JB, et al. Symptoms of anxiety and depression among adults with arthritis: United States, 2015–2017. MMWR Morb Mortal Wkly Rep 2018;67:1081–7.
15. Kigozi J, Jowett S, Nicholl BI, Lewis M, Bartlam B, Green D, et al. Cost-utility analysis of routine anxiety and depression screening in patients consulting for osteoarthritis: results from a clinical, randomized controlled trial. Arthritis Care Res (Hoboken) 2018;70: 1787–94.
16. Heiman E, Kravitz RL, Wise BL. Rheumatologists' approaches to diagnosis and treatment of depression. J Clin Rheumatol 2016;22: 307–11.
17. Sheehy C, Murphy E, Barry M. Depression in rheumatoid arthritis: underscoring the problem. Rheumatology (Oxford) 2006;45:1325–7.
18. Losina E, Silva GS, Smith KC, Collins JE, Hunter DJ, Shrestha S, et al. Quality-adjusted life-years lost due to physical inactivity in the United States osteoarthritis population. Arthritis Care Res (Hoboken) 2020;72:1349–57.
19. Losina E, Usiskin IM, Smith SR, Sullivan JK, Smith KC, Hunter DJ, et al. Cost-effectiveness of generic celecoxib in knee osteoarthritis for average-risk patients: a model-based evaluation. Osteoarthritis Cartilage 2018;26:641–50.
20. Losina E, Walensky RP, Reichmann WM, Holt HL, Gerlovin H, Solomon DH, et al. Impact of obesity and knee osteoarthritis on morbidity and mortality in older Americans. Ann Intern Med 2011;154: 217–26.
21. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. JAMA 2016;316:1093–103.
22. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness: the curious resilience of the $50,000-per-QALY threshold. N Engl J Med 2014;371:796–7.
23. National Institutes of Health. The Osteoarthritis Initiative. 2013. URL: https://nda.nih.gov/oai/.
24. Bannuru RR, Osani MC, Vaysbrot EE, Arden NK, Bennell K, Bierma-Zeinstra SMA, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthritis Cartilage 2019;27:1578–89.
25. Kolasinski SL, Neogi T, Hochberg MC, Oatis C, Guyatt G, Block J, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis Rheumatol 2020;72:220–33.
26. Centers for Disease Control and Prevention. National Health and Nutrition Examination Survery (NHANES), 2014–2016. 2016. URL: https://www.cdc.gov/nchs/nhanes/index.htm.
27. Medicare Current Beneficiary Survey (MCBS). 2015–2016. URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS.
28. Pope G. KJ, Ingber M, Freeman S, Sekar R, Newhart C. Evaluation of the CMS-HCC Risk Adjustment Model RTI International. March 2011. Centers for Medicare and Medicaid Services.
29. Sullivan PW, Ghushchyan V. Preference-based EQ-5D index scores for chronic conditions in the United States. Med Decis Making 2006; 26:410–20.
30. Arias E, Heron M, Xu J. United States life tables, 2014. Natl Vital Stat Rep 2017;66:1–64.
31. Cuijpers P, Vogelzangs N, Twisk J, Kleiboer A, Li J, Penninx BW. Comprehensive meta-analysis of excess mortality in depression in the general community versus patients with specific illnesses. Am J Psychiatry 2014;171:453–62.
32. Losina E, Collins JE, Wright J, Daigle ME, Donnell-Fink LA, Strnad D, et al. Postoperative care navigation for total knee arthroplasty patients: a randomized controlled trial. Arthritis Care Res (Hoboken) 2016;68:1252–9.
33. Smith SR, Katz JN, Collins JE, Solomon DH, Jordan JM, Suter LG, et al. Cost-effectiveness of tramadol and oxycodone in the treatment of knee osteoarthritis. Arthritis Care Res (Hoboken) 2017;69:234–42.
34. Scott DL, Berry H, Capell H, Coppock J, Daymond T, Doyle DV, et al. The long-term effects of non-steroidal anti-inflammatory drugs in osteoarthritis of the knee: a randomized placebo-controlled trial. Rheumatology (Oxford) 2000;39:1095–101.
35. Raskin J, Wiltse CG, Siegal A, Sheikh J, Xu J, Dinkel JJ, et al. Efficacy of duloxetine on cognition, depression, and pain in elderly patients with major depressive disorder: an 8-week, double-blind, placebo-controlled trial. Am J Psychiatry 2007;164:900–9.
36. Hunkeler EM, Katon W, Tang L, Williams JW Jr, Kroenke K, Lin EH, et al. Long term outcomes from the IMPACT randomised trial for depressed elderly patients in primary care. BMJ 2006;332:259–63.
37. Nelson JC, Lu Pritchett Y, Martynov O, Yu JY, Mallinckrodt CH, Detke MJ. The safety and tolerability of duloxetine compared with paroxetine and placebo: a pooled analysis of 4 clinical trials. Prim Care Companion J Clin Psychiatry 2006;8:212–9.
38. Micromedex. Red Book online. 2018. URL: https://www.micromedexsolutions.com.
39. National Community Pharmacists Association, Coalition for Community Pharmacy Action. Cost of dispensing study: an independent comparative analysis of US prescription dispensing costs. 2015.
40. Centers for Medicare and Medicaid Services. Medicare Fee Schedules 2018. URL: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/FeeScheduleGenInfo.
41. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16:606–13.
42. Mitchell AJ, Yadegarfar M, Gill J, Stubbs B. Case finding and screening clinical utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis of 40 studies. BJPsych Open 2016;2:127–38.
43. Sullivan PW, Valuck R, Saseen J, MacFall HM. A comparison of the direct costs and cost effectiveness of serotonin reuptake inhibitors and associated adverse drug reactions. CNS Drugs 2004;18:911–32.
44. Nelson JC, Oakes TM, Liu P, Ahl J, Bangs ME, Raskin J, et al. Assessment of falls in older patients treated with duloxetine: a secondary analysis of a 24-week randomized, placebo-controlled trial. Prim Care Companion CNS Disord 2013;15:PCC.12m01419.
45. Soh SE, Barker AL, Morello RT, Ackerman IN. Applying the International Classification of Functioning, Disability and Health framework to determine the predictors of falls and fractures in people with osteoarthritis or at high risk of developing osteoarthritis: data from the Osteoarthritis Initiative. BMC Musculoskelet Disord 2020;21:138.

46. 46brooklyn. New Medicare Part D "Ski Slope" shows seniors' wild drug pricing ride. 2019. URL: https://www.46brooklyn.com/research/2019/7/14/new-viz-shows-part-d-pricing-distortions.
47. Wielage RC, Bansal M, Andrews JS, Klein RW, Happich M. Cost-utility analysis of duloxetine in osteoarthritis: a US private payer perspective. Appl Health Econ Health Policy 2013;11:219–36.
48. Beard SM, Roskell N, Le TK, Zhao Y, Coleman A, Ang D, et al. Cost effectiveness of duloxetine in the treatment of fibromyalgia in the United States. J Med Econ 2011;14:463–76.
49. Wielage RC, Bansal M, Andrews JS, Wohlreich MM, Klein RW, Happich M. The cost-effectiveness of duloxetine in chronic low back pain: a US private payer perspective. Value Health 2013;16:334–44.
50. Katz JN, Smith SR, Collins JE, Solomon DH, Jordan JM, Hunter DJ, et al. Cost-effectiveness of nonsteroidal anti-inflammatory drugs and opioids in the treatment of knee osteoarthritis in older patients with multiple comorbidities. Osteoarthritis Cartilage 2016;24:409–18.
51. Brazier HE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care 2004;42:851–9.
52. Pope GC, Kautter J, Ellis RP, Ash AS, Ayanian JZ, Lezzoni Li, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. Health Care Financ Rev 2004;25:119–41.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 790–798
DOI 10.1002/acr.24516

American College of Rheumatology
Empowering Rheumatology Professionals

Screening to Identify Postoperative Pain and Cross-Sectional Associations Between Factors Identified in This Process With Pain and Function, Three Months After Total Knee Replacement

Vikki Wylde,[1] Emily Sanderson,[2] Tim J. Peters,[2] Wendy Bertram,[3] Nicholas Howells,[4] Julie Bruce,[5] Christopher Eccleston,[6] and Rachael Gooberman-Hill[1]

Objective. To describe the screening and recruitment process of a randomized trial and evaluate associations with knee pain and function 3 months after total knee replacement (TKR).

Methods. In order to screen for a multicenter trial, a total of 5,036 patients were sent the Oxford Knee Score (OKS) questionnaire 10 weeks post-TKR. Patients who reported pain in their replaced knee (score of ≤14 on the OKS pain component) completed a second OKS questionnaire 12 weeks post-TKR. Those patients who were still experiencing pain 12 weeks post-TKR completed a detailed questionnaire 13 weeks post-TKR. These data were used to characterize pain in a cross-sectional analysis. Multivariable regression was performed in order to identify factors associated with pain and function at 13 weeks post-TKR.

Results. We received OKS questionnaires from 3,058 of 5,063 TKR patients (60%), and 907 of the 3,058 (30%) reported pain in their replaced knee 10 weeks post-TKR. By 12 weeks, 179 of 553 patients (32%) reported improved pain (score of >14 on the OKS pain component). At 13 weeks, 192 of 363 patients (53%) who completed a detailed questionnaire reported neuropathic pain, 94 of 362 (26%) reported depression symptoms, and 95 of 363 (26%) anxiety symptoms. More severe pain at 13 weeks postoperatively was associated with poorer general health, poorer physical health, more pain worry, and lower satisfaction with surgery outcome. More severe functional limitation was associated with higher levels of depression, more pain worry, lower satisfaction with surgery outcome, and higher pain acceptance.

Conclusion. Screening after TKR identified individuals with pain. We identified several potential targets (physical and mental health outcomes, acceptance of pain, and quality of life) for tailored intervention to improve outcomes for patients. Future trials of multidisciplinary interventions warranted.

INTRODUCTION

Primary total knee replacement (TKR) is a common operation, with over 100,000 operations performed in the UK's National Health Service (NHS) in 2019 (1,2). The main indications for TKR are chronic pain and functional limitations, which are predominately related to osteoarthritis. Although the operation is successful for many, 10–34% of patients experience ongoing pain in the months and years after surgery (3). Despite its prevalence, knowledge about the onset and postoperative trajectory of chronic pain after TKR is not well understood (4). The evidence base for treatment and management is sparse (5,6), and

The views expressed herein are those of the authors and do not necessarily represent those of the NIHR or the Department of Health and Social Care.

Supported by the NIHR (program grant for applied research RP-PG-0613-20001) and the NIHR Biomedical Research Centre at the University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. Dr. Bruce's work was supported by University Hospitals Coventry and Warwickshire (NIHR research capability funding).

[1]Vikki Wylde, PhD, Rachael Gooberman-Hill, PhD: Bristol Medical School, University of Bristol and NIHR Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK; [2]Emily Sanderson, MSc, Tim J. Peters, PhD: Bristol Medical School, University of Bristol, Bristol, UK; [3]Wendy Bertram, MSc: Bristol Medical School, University of Bristol and North Bristol NHS Trust, Bristol, UK; [4]Nicholas Howells, MD: North Bristol NHS Trust, Bristol, UK; [5]Julie Bruce, PhD: University of Warwick, Warwick, UK; [6]Christopher Eccleston, PhD: The University of Bath, Bath, UK.

Dr. Wylde and Ms. Sanderson contributed equally to this work.

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

Address correspondence to Vikki Wylde, PhD, Musculoskeletal Research Unit, Bristol Medical School, University of Bristol, Learning and Research Building, Southmead Hospital, Bristol, BS10 5NB. Email: V.Wylde@bristol.ac.uk.

Submitted for publication April 22, 2020; accepted in revised form November 12, 2020.

SIGNIFICANCE & INNOVATIONS

- Our study demonstrated good uptake to early postoperative screening of pain and function after total knee replacement (TKR).
- Half of the patients with pain at 3 months after TKR reported neuropathic pain symptoms.
- One-fourth of the patients with pain at 3 months after TKR report depression and/or anxiety.
- Multiple factors, such as quality of life, physical and mental health outcomes, and acceptance of pain, are associated with more severe pain and functional limitations after TKR, which highlights the need for multidisciplinary interventions.

referrals for assessment and care are inconsistent (7,8). There is no preoperative model as yet that can accurately predict who will have chronic pain after surgery (5,9). People with chronic pain after TKR can feel abandoned by health care services and struggle to understand ongoing pain (10).

The improvement trajectory following TKR is variable; however, most pain relief occurs within the first 3 months postoperatively (11). Persistent pain at 3 months could be due to slower recovery or an early indication of long-term chronic pain. Chronic pain is difficult to treat once established (12) and the identification and characterization of pain early in the recovery trajectory could facilitate the delivery of targeted interventions to support recovery and improve long-term pain outcomes. Hence, there is potential for early identification of these patients to explore whether intervention is warranted.

Previous studies have described pain after TKR (3,13–17), but these studies have methodologic shortcomings that have contributed to the poor quantification and characterization of pain after TKR. These shortcomings include the use of surgeon-administered tools to assess pain, limited assessment of the multidimensional nature of pain, variable definitions of pain resulting in different prevalence estimates, and single-center studies, which all limit generalizability (3,16,18). A robust method of identifying patients with pain after TKR using the Oxford Knee Score (OKS) pain component has been developed (19). Using data from a national population-based cohort across England, patients with a postoperative score of ≤14 on the OKS pain component were identified as having pain likely to negatively impact health-related quality of life (19). Applying this method for identification of patients with pain in the first 3 months postoperatively allows the early investigation of pain characteristics. The aim here is to describe our screening procedures to identify people with postoperative pain and to identify associations with pain and function among patients with pain in the first 3 months after primary TKR.

PATIENTS AND METHODS

Design. The data analyzed in the present study are from the Support and Treatment After Joint Replacement (STAR) trial, a multicenter randomized trial evaluating the effectiveness of a care pathway for patients with chronic pain at 3 months after TKR (20). Screening data included in the analyses of the present study were collected before randomization and were analyzed as observational data. Study methods relevant to these analyses were described and reported following the Strengthening and Reporting of Observational Studies in Epidemiology guidance (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24516).

Patient and public involvement. This research was conducted in collaboration with the Patient Experience Partnership in Research STAR group, which is a specialized group comprised of 5 patients who have experienced chronic pain after TKR. Through regular group meetings, patient representatives contributed to project design and management.

Participants. Between September 2016 and May 2019, eligible patients were recruited into the STAR trial from 8 NHS orthopedic centers in Bristol, Cardiff, Exeter, Mansfield, Oswestry, Wrightington, Leicester, and Birmingham. Inclusion criteria included adults who received a primary TKR for osteoarthritis and reported pain in their replaced knee 12 weeks postoperatively. Exclusion criteria included lack of capacity to provide informed consent, previous study participation for the contralateral knee, or participation in another project that interfered with the STAR trial. The STAR trial complied with the Declaration of Helsinki and was approved by the Southwest–Central Bristol Research Ethics Committee (16/SW/0154) and the Health Research Authority. All participants provided written informed consent in 2 stages: 1) for the screening study only, comprising OKS measurements at 10 and 12 weeks after TKR and 2) for the main STAR trial, comprising a detailed baseline questionnaire at 13 weeks after TKR. Identification of patients with pain after TKR began 10 weeks postoperatively to ensure timely identification of those with pain that persisted 3 months postoperatively. We reported our findings of screening procedures and the cross-sectional analysis of associations with pain and function 13 weeks after TKR surgery.

Initial postal screening to identify patients with pain 10 weeks after TKR. Patients who received a primary TKR due to osteoarthritis 8 weeks previously were sent a study information leaflet, consent form, and short initial screening questionnaire, including the OKS (21) and sociodemographic questions. Nonresponders received a single reminder. The OKS is a joint-specific measure of pain and function consisting of 12 items with 5 ordinal response options for each item (21). There is evidence of validity and reliability, with the OKS being reported as the best performing site-specific patient-reported outcome measure in a psychometric review of 32 measures used in hip and knee replacement surgery (22). It has an overall score ranging from 0–48 (worst to best). Two subscales can be calculated,

including a 5-item OKS function component (raw score of 0–20) and a 7-item OKS pain component (raw score of 0–28). Patients with a score of 0–14 on the raw OKS pain component were considered to have pain that was likely to negatively impact health-related quality of life (19). It is recommended that the component scores are standardized to a 0–100 scale (worst to best) for analysis (23).

Second telephone screening to confirm ongoing pain 12 weeks after TKR. All responding patients who reported an OKS pain score of ≤14 at 10 weeks were contacted by telephone at 12 weeks and invited to complete a second screening questionnaire that repeated the OKS questionnaire, in order to confirm their pain status. Those still reporting clinically meaningful pain (defined as an OKS pain score of ≤14) at 12 weeks were eligible for invitation to enter the trial.

Detailed study questionnaire at 13 weeks after TKR for patients with pain. Participants who gave their consent to the trial completed a third OKS questionnaire as part of a more detailed study questionnaire prior to randomization. If questionnaires were not returned within 1 week, the participant was offered support on the telephone with a researcher.

The outcomes assessed in the questionnaire administered 13 weeks postoperatively reflected the 8 domains of the core outcome set for chronic pain after TKR (24). Pain severity and pain interference were assessed using the Brief Pain Inventory (BPI) (subscale scores range 0–10 [best to worst]) (25). Knee pain and function were measured using the OKS questionnaire.

Pain with neuropathic features was assessed using 2 questionnaires. First, we used the PainDETECT (26), which can be analyzed as a continuous score (range –1 to 38, with a higher score indicating greater likelihood of neuropathic pain) or categorized into nociceptive pain (range –1 to 12), possible neuropathic pain component (range 13–18), or probable neuropathic pain component (range 19–38). Second, the Dolour Neuropathic scale (DN4) (27), with scores ranging from 0–7 (best to worst) and a score of ≥3 indicating neuropathic pain characteristics, was used. Single questions evaluated the frequency of pain in the past 24 hours and 4 weeks and how this pain compared to preoperative pain.

General health was measured using the Short Form 12 (SF-12) health survey (28), comprising a physical component score and a mental component score (range 0–100 [worst to best]). Health-related quality of life was assessed by the 5-level version of the EuroQol 5-domain instrument (EQ-5D-5L [29]) (range –0.594 to 1, where 1 indicates "perfect health" and 0 indicates "dead"). Capability was assessed by the Icepop Capability Measure for Adults (ICECAP-A [30]) (range –0.001 to 1 [worst to best]).

Depression and anxiety were assessed using the Hospital Anxiety and Depression Scale (HADS) (31), with subscale scores (HADS anxiety and HADS depression) ranging from 0 to 21 (best to worst) and categorized into unlikely symptoms of depression/anxiety (range 0–7), possible depression/anxiety (range 8–10), and probable depression/anxiety (range 11–21). Worry about pain was assessed with the Pain Catastrophizing Scale (PCS) (range 0–52 [best to worst]) (32), which consists of 3 subscales labeled rumination (scored 0–16), magnification (scored 0–12), and helplessness (scored 0–24). The Possible Solutions to Pain Questionnaire (33) was also completed, and the 4 subscales were analyzed, including solving pain (scored 0–24 [worst to best]), meaningfulness of life despite pain (scored 0–30), acceptance of the insolubility of pain (scored 0–18), and belief in solutions (scored 0–12).

Patient satisfaction with the outcome of surgery was measured by the Self-Administered Patient Satisfaction Scale (34), a 4-item arthroplasty-specific score (range 25–100 [worst to best]). Painful body regions were indicated on a body diagram, and widespread pain was defined as pain in at least 2 sections of each 2 contralateral arms or legs and in the axial skeleton (35). Sociodemographic questions included age, sex, marital and living status, ethnicity, and education level.

Statistical analysis. *Screening questionnaire.* In addition to response rates, distributions of screening OKS scores were assessed at each phase using histograms and summary statistics, such as the mean ± SD. Regression analyses were performed on the OKS pain and function subscores as the outcome variables to explore the associations with age and sex. Results are shown as regression coefficients, 95% confidence intervals, and *P* values. The relationship between the OKS subscales were assessed using scatter plots, replicated stratified by age group (age ranges of <60, 60–70, 71–80, and >80 years) and sex.

Study questionnaire analyses. Summary statistics for sociodemographic data and patient-reported outcomes were presented as mean ± SD, median (interquartile range), and number (%). Distributions of the OKS scales were presented as histograms, in order to assess normality. Correlation coefficients between pain outcomes were evaluated. Linear regression was used to evaluate the factors independently associated with the OKS pain and function scores. Staged regression was then used to select variables systematically for the linear regression model (36). Associations were explored between OKS pain and groups of factors, including sociodemographic variables, general health, and mental health measures. Each group of variables was first explored separately in multivariable regression models, with (iterative) exclusion of variables without strong associations with OKS pain when adjusted for other variables in the model. The process was then extended to consider all groups together, resulting in a final regression model containing only variables that were strongly associated with OKS pain, adjusted for other variables. This process was repeated for OKS function exploring associations with sociodemographic variables and mental health outcomes. In all

analyses, the standardized OKS pain and function scores (range 0–100) were used (23).

Data completeness is reported in the Tables and Figures. For the OKS component scores, the mean of other items on the subscale was used to impute a missing item, if only 1 item was missing. If more than 1 item was missing, a score was not calculated (37). The approach to missing data for other validated questionnaires followed guidance recommended by the questionnaire developers; further details are in the STAR trial statistical analysis plan (38).

Sample size. The sample size for the STAR trial was based on detecting a minimal clinically important difference between trial arms in the BPI subscales 12 months after randomization (20). We did not undertake a separate power calculation for the analyses presented herein, as our intention was to investigate characteristics of the study population collected prior to randomization; rather, the levels of achieved precision are indicated through the relevant confidence intervals.

Data availability. The data sets generated during the current study will be available in the University of Bristol Research Data Repository (https://data.bris.ac.uk/data/). Data will be available following publication of the trial results. Access to the data will be restricted to ensure that data is only made available to bona fide researchers for ethically approved research projects, on the understanding that confidentiality will be maintained and after a data access agreement has been signed by an institutional signatory.

RESULTS

Recruitment, screening, and participant flow. An overview of participant flow through the study is provided in Figure 1. Screening questionnaires to identify patients with pain after TKR were sent to 5,036 patients who had a TKR at 1 of 8 orthopedic centers. Completed screening questionnaires were returned postoperatively by 3,058 patients (61%) at a mean ± SD of 10 ± 2 weeks. Of these patients, 907 (30%) reported pain in their replaced knee at 10 weeks, of whom 553 (61%) completed a second telephone OKS to confirm pain status at a mean ± SD of 12 ± 2 weeks. The mean ± SD age of the 553 patients who completed a telephone questionnaire was 67.7 ± 8.6 years, with 56% of the patients being female. Those who did not complete a telephone questionnaire at 12 weeks (n = 354) were slightly older, with a mean ± SD age of 69.4 ± 10.4 years, and 62% were female. Patients who completed the 12-week telephone-administered OKS questionnaire had a slightly higher mean ± SD OKS score (18.2 ± 5.4) than those who did not complete a telephone OKS questionnaire

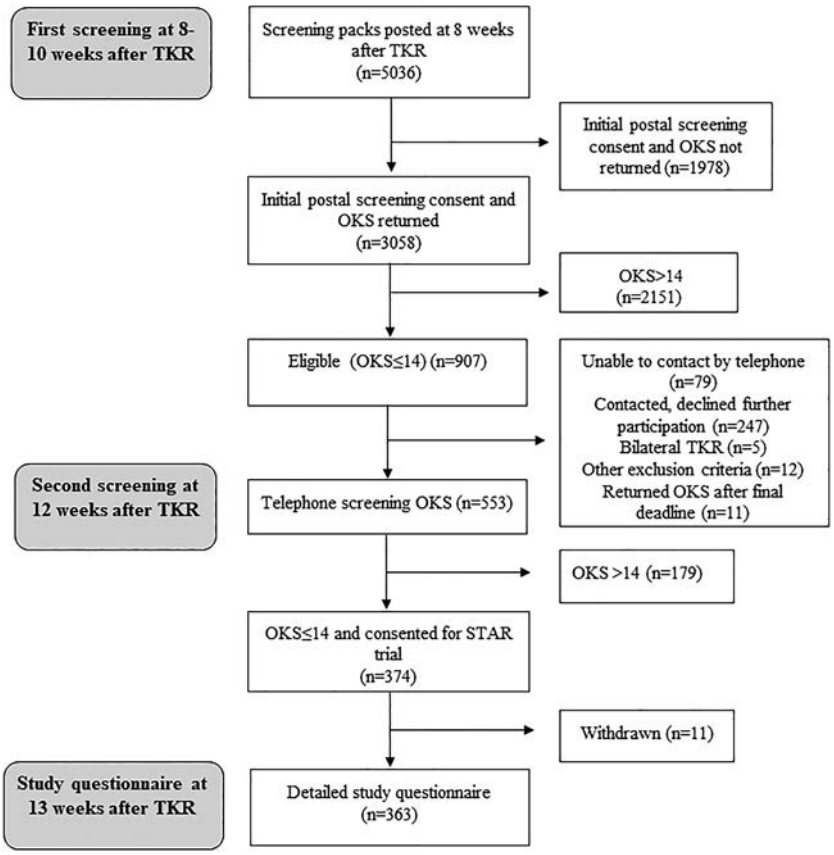

Figure 1. Participant flow chart. TKR = total knee replacement; OKS = Oxford Knee Score; STAR = Support and Treatment After Joint Replacement.

Table 1. Characteristics of responders and nonresponders to screening questionnaire at 10 and 12 weeks post-TKR*

	Responders	Nonresponders
Screening questionnaire 10 weeks post-TKR		
Total, no. (%)	3,058 (61)	1,977 (39)
Age, years	69.7 ± 8.8	69.9 ± 9.8
Female sex, %	54.5	62.2
Total OKS (0–48; worst to best)	29.3 ± 9.6	–
OKS pain component (0–100; worst to best)	62.3 ± 21.2	–
OKS function component (0–100; worst to best)	59.4 ± 20.8	–
Telephone-administered screening questionnaire at 12 weeks post-TKR†		
Total, no. (%)	553 (61)	354 (39)
Age, years	67.7 ± 8.6	69.4 ± 10.4
Female sex, %	56.2	62.0
Total OKS 10 weeks post-TKR	18.2 ± 5.4	17.2 ± 5.7
OKS pain component 10 weeks post-TKR	36.6 ± 11.3	35.6 ± 12.2
OKS function component 10 weeks post-TKR	40.0 ± 14.4	36.2 ± 15.0

* Values are the mean ± SD unless indicated otherwise. OKS = Oxford Knee Score; TKR = total knee replacement.
† Eligible at 10 weeks (n = 907).

(17.2 ± 5.7), indicating less pain and better function in responders compared with nonresponders at 12 weeks. A total of 363 of 553 patients (66%) completed a detailed questionnaire at a mean ± SD 13 ± 2 weeks.

Sociodemographic characteristics of responders and nonresponders to the screening questionnaire at 10 weeks are provided in Table 1. The mean age was comparable (70 years), although female patients were slightly less likely to respond than male patients (55% versus 62%). The OKS overall and component scores are in Table 1 and Supplementary Figures 1 and 2 (available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24516). Overall, 907 of 3,058 patients (30%) reported clinically meaningful pain in their replaced knee at 10 weeks (OKS pain component score ≤14).

Scatter plots of the OKS component scores demonstrated a linear relationship between pain and function, with similar patterns when stratified by sex and age (see Supplementary Figures 3–5, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24516). Younger age and female sex were associated with worse knee pain severity and functional limitations at 10 weeks (Table 2).

Of the 533 of 907 patients to complete the OKS by telephone, 179 (32%) reported an improvement in pain (OKS score of >14), but 374 (68%) remained in pain. A summary of the statistics of age, sex, and week-10 OKS scores for those who did and did not respond at week 12 are presented in Table 1. Responders were slightly younger than nonresponders at 12 weeks, with a lower proportion of female responders. Responders at 12 weeks had slightly higher week-10 OKS scores compared with those who did not respond at 12 weeks.

Characterization of people reporting pain 13 weeks post-TKR. Sociodemographic characteristics and patient-reported outcomes for the 363 of 374 (97%) participants who completed a detailed questionnaire at 13 weeks are shown in Table 3. The mean ± SD age of these participants was 67 ± 9 years, and 60% were female. Neuropathic pain characteristics were common, with half of the participants (53%) having a PainDETECT score that indicated likely neuropathic pain (score >19) and three-fourths (74%) of patients having neuropathic pain characteristics according to the DN4 (score of ≥3). A total of 47% of patients likely had both neuropathic pain according to PainDETECT and neuropathic pain according to the DN4. Poor mental health was also common, with patients having HADS scores indicative of either probable depression (26%) or anxiety (26%); of these patients, 60 of 362 (17%) reported symptoms of both depression and anxiety. Over the previous 4 weeks, 96% of patients had experienced pain frequently, defined as pain being present "often," "most of the time," or "all the time." Almost half of patients (44%) reported their pain as "a bit worse" or "much worse" than their preoperative pain. Despite still being in pain at 12 weeks, most patients (74%) were satisfied with their overall outcome from TKR, and 55% were satisfied with their pain relief, although satisfaction rates with ability to do activities of daily living and leisure activities were lower (39% and 38%, respectively).

Regression analysis. Results of the linear regression model with the OKS pain component as the outcome are shown in Table 4. In this cross-sectional analysis, having more severe knee pain at 13 weeks was associated with lower general health measured by the EQ-5D-5L utility score, lower physical health measured by the SF-12 health survey, higher pain worry (PCS), and lower satisfaction with the outcome of surgery.

Table 2. Univariable associations between age, sex, pain, and function 10 weeks after TKR*

		OKS pain component		OKS function component	
	No.	Coefficient (95% CI)	*P*	Coefficient (95% CI)	*P*
Age	2,915	0.29 (0.20, 0.37)	<0.001	0.09 (0.009, 0.18)	0.030
Sex (ref. = male)	3,042	−3.09 (−4.60, −1.58)	<0.001	−7.18 (−8.64, −5.71)	<0.001

* 95% CI = 95% confidence interval; OKS = Oxford Knee Score; ref. = reference; TKR = total knee replacement.

Table 3. Characteristics of participants with pain 3 months after TKR (n = 363)*

Characteristic	
Age, years	
Mean ± SD	67.2 ± 8.7
Median (IQR)	67 (61–73)
Range	40–88
Sex	
Female	217 (60)
Male	146 (40)
Marital status (n = 356)	
Single	25 (7)
Married/partner	251 (71)
Divorced/separated	35 (10)
Widowed	45 (13)
Living arrangement (n = 356)	
Live alone	78 (22)
With spouse/partner	253 (71)
With someone else	22 (6)
Other	3 (1)
Ethnicity (n = 356)	
White	335 (94)
Asian	13 (4)
Black	5 (1)
Mixed	1 (<1)
Other	2 (<1)
Education level (n = 318)	
School left <16 years	22 (7)
School left 16 years	194 (61)
College	63 (20)
University	15 (5)
Other postgraduate	24 (8)
BPI scores, mean ± SD	
Severity	5.2 ± 1.7
Interference	6.28 ± 1.92
OKS scores, mean ± SD	
Total	18.23 ± 5.83
Pain	36.75 ± 12.70
Function	39.70 ± 14.28
Pain Catastrophizing Scale, median (IQR) (n = 360)	
Total	18 (9.25–30.5)
Rumination	8 (4–12)
Magnification	2 (1–5)
Helplessness	8 (4–14)
Pain solution (PaSol), median (IQR)	
Solving pain (n = 362)	18 (14–22)
Meaningful life (n = 362)	22 (18–26)
Acceptance of the insolubility of pain (n = 358)	8 (5–11)
Belief in solution (n = 359)	9 (6–12)
Patient Satisfaction, mean ± SD (n = 360)	62.88 ± 18.99
ICECAP-A, median (IQR) (n = 362)	0.78 (0.55–0.89)
SF-12, mean ± SD	
Physical score	33.44 ± 6.51
Mental score	42.19 ± 11.12
EQ-5D-5L, median (IQR) (n = 358)	0.53 (0.30–0.62)
DN4 score, mean ± SD (n = 359)	3.79 ± 1.71
Neuropathic pain characteristics according to DN4?	
Yes	267 (74)
No	92 (26)
PainDETECT score, mean ± SD	18.19 ± 6.77
Neuropathic pain characteristics according to PainDETECT?	
Unlikely	76 (21)
Ambiguous	96 (26)
Likely	191 (53)
HADS: Anxiety	
Normal	197 (54)
Borderline anxiety	71 (20)
Clinical anxiety	95 (26)
HADS: Depression (n = 362)	
Normal	177 (49)
Borderline depression	91 (25)
Clinical depression	94 (26)
Pain frequency in past 24 hours (n = 361)	
Rarely	1 (<1)
Sometimes	40 (11)
Often	98 (27)
Most of the time	164 (45)
All of the time	58 (16)
Pain frequency in past 4 weeks (n = 362)	
Rarely	0 (0)
Sometimes	14 (4)
Often	102 (28)
Most of the time	156 (43)
All of the time	90 (25)
Satisfaction with overall results of TKR (n = 359)	
Very dissatisfied	21 (6)
Somewhat dissatisfied	72 (20)
Somewhat satisfied	154 (43)
Very satisfied	112 (31)
Satisfaction with improving pain (n = 359)	
Very dissatisfied	47 (13)
Somewhat dissatisfied	117 (33)
Somewhat satisfied	139 (39)
Very satisfied	56 (16)
Satisfaction with improving ability to do housework or gardening (n = 358)	
Very dissatisfied	65 (18)
Somewhat dissatisfied	152 (42)
Somewhat satisfied	111 (31)
Very satisfied	30 (8)
Satisfaction with improving ability to do leisure activities (n = 359)	
Very dissatisfied	86 (24)
Somewhat dissatisfied	140 (39)
Somewhat satisfied	106 (30)
Very satisfied	27 (8)
Comparison of pain to preoperative pain (n = 362)	
Much better	79 (22)
A bit better	70 (19)
The same	54 (15)
A bit worse	77 (21)
Much worse	82 (23)
Presence of chronic widespread pain (Manchester definition)	
Yes	16 (4)
No	347 (96)

* Values are the number (%) unless indicated otherwise. BPI = Brief Pain Inventory; DN4 = Dolour Neuropathic scale; EQ-5D-5L = 5-level version of the EuroQol 5-domain instrument; HADS = Hospital Anxiety and Depression Scale; ICECAP-A = Icepop Capability Measure for Adults; IQR = interquartile range; OKS = Oxford Knee Score; PaSOL = Pain Solutions Questionnaire; SF-12 = Short Form 12 health survey; TKR = total knee replacement.

Table 4. Final model from the linear regression for associations with pain 3 months after TKR (n = 352)*

Variable	Coefficient (95% CI)	*P*
EQ-5D-5L	19.9 (14.1, 25.8)	<0.001
SF-12 (physical)	0.25 (0.09, 0.42)	0.003
Pain Catastrophizing Scale	–0.27 (–0.36, –0.17)	<0.001
Patient Satisfaction Scale	0.11 (0.05, 0.16)	<0.001

* 95% CI = 95% confidence interval; EQ-5D-5L = 5-level version of the EuroQol 5-domain instrument; SF-12 = Short Form 12 health survey.

From the linear regression model with the OKS function component as the outcome (Table 5), in patients with pain at 13 weeks postoperatively, more severe functional limitation was associated with higher levels of depression, higher pain catastrophizing, lower satisfaction with the outcome of surgery, and higher levels of acceptance of the insolubility of pain.

DISCUSSION

The present study examined characteristics of people reporting pain 10 to 13 weeks after TKR. We used the validated OKS questionnaire pain component threshold to identify patients with pain in the first 3 months after TKR. Using this standardized pain definition, 30% of patients reported pain in their replaced knee 10 weeks after surgery. Of the 553 patients who completed a second OKS questionnaire by telephone (12 weeks after TKR), 30% reported an improvement in their OKS pain score from the 10-week measurement. However, for the majority (70%), the pain was still present at 3 months. Applying the OKS pain threshold allowed an in-depth evaluation of the characteristics of patients with pain 3 months after TKR. We found that more than half of the patients reported pain with neuropathic characteristics, one-fourth of the patients reported probable depression or anxiety, and 17% reporting both depression and anxiety. Despite still having problems with pain, three-fourths of these patients were satisfied with their TKR outcome. Patients with more severe knee pain at 3 months were likely to have poorer general health, poorer physical health, higher pain worry (measured as pain catastrophizing), and lower satisfaction with the outcome of surgery. Patients with greater functional limitations were more likely to have higher levels of depression, higher pain worry, lower satisfaction with the outcome of surgery, and higher levels of acceptance of the pain's insolubility.

Previously, the lack of a robust approach to screening has been a barrier to the implementation of new services to improve care for patients with pain after TKR (7). Our study demonstrated that early screening using the OKS questionnaire definition of chronic pain as a standardized approach to identify patients with pain is achievable. In our large multicenter trial, one-third of patients met our definition of pain at 10 weeks; this is not unexpected, as TKR has a long recovery period and individual patients' recovery trajectories vary (11). One-third of patients with pain at 10 weeks had improved by 12 weeks, demonstrating that patients can experience rapid recovery during this early postoperative period. However, 70% of responding patients with pain at 10 weeks still had pain at 12 weeks, and for some, this pain is likely to persist for the long term. Early screening to identify patients with pain at 3 months could facilitate targeted care delivery (e.g., through transitional pain clinics) to prevent the transition of acute pain to chronic pain (39).

The prevalence of neuropathic pain after TKR and other types of surgery differs in the literature, likely due to variation in definition and measurement (40). This warrants further research and suggests a potential role for routine screening and treatment of neuropathic pain after TKR. A systematic review by Finnerup et al has identified inadequate response to pharmacotherapy for neuropathic pain, which relates to modest efficacy, high placebo rates, and poor phenotyping (41). Further work could examine the development of targeted interventions including nonpharmacologic treatments. For example, the National Institute for Health and Care Excellence currently recommends trials comparing the effectiveness of combination therapy versus monotherapy for neuropathic pain (42). Another potential target for intervention is depression and anxiety, reported by one-fourth of participants in our study. Given the known association between mental health and chronic pain (43), concurrent treatment of both conditions may improve outcomes for patients.

An interesting finding in the present study was that, despite ongoing pain, satisfaction with treatment was high. This may have been in part influenced by the relatively early time point of assessment postsurgery and an acceptance that initial postoperative pain is part of the recovery trajectory. Satisfaction is a complex construct that can be influenced by a wide array of interrelated factors (44). The degree of dissatisfaction experienced by patients with chronic pain after TKR has been associated with various factors, including instability in the coronal plane, stiffness, and negative social support (17). Further research would help to further understand the factors that influence patients' satisfaction with their outcome.

Our analysis also identified factors that were associated with more severe pain and functional limitations at 3 months. These associations are consistent with previous studies of pain conditions (44–46) and present potential areas for intervention to improve patient outcomes. Any such intervention should be

Table 5. Final model from the linear regression for associations with function 3 months after TKR (n = 353)*

Variable	Coefficient (95% CI)	*P*
HADS depression	–1.18 (–1.55, –0.80)	<0.001
PaSOL (acceptance of pain)	–0.52 (–0.78, –0.25)	<0.001
Pain Catastrophizing Scale	–0.24 (–0.36, –0.12)	<0.001
Patient Satisfaction Scale	0.09 (0.01, 0.16)	0.019

* 95% CI = 95% confidence interval; HADS = Hospital Anxiety and Depression Scale; PaSOL = Pain Solutions Questionnaire; TKR = total knee replacement.

multidisciplinary to address the varied nature of factors associated with pain. The association of more severe functional limitations with higher levels of pain acceptance of the insolubility of pain that is beyond the general measures of mental health is highly unusual and deserves further attention. This association might be artifactual (floor effect), as 27% of the sample recorded "not applicable" to the item "I can accept that there is no solution for my pain." Many patients found the idea of accepting the lack of a solution as simply not relevant to their early postoperative phase. The association could be explained by some patients entertaining the idea of accepting the insolubility of pain because of severity of symptoms. Speculatively, it could also demonstrate a fatalistic coping strategy in which one expects pain after surgery. This coping style could be negative, acting as a barrier to engaging with treatment-seeking for pain, or could be positive, acting as a means to disengage from unachievable goals (47).

There are several factors limiting the interpretation of the results from this study. First, the response rate of 61% to the initial postal screening questionnaire at 10 weeks, although comparable to other surveys of orthopedic populations (17,48), may have introduced a responder bias (49). Of note, women were slightly less likely to respond to the screening questionnaire, and female sex was associated with more severe pain at 10 weeks; this may underestimate pain prevalence. Second, our screening of patients with pain after TKR began 10 weeks postoperatively, sooner than the internationally accepted 3-month definition of chronic postsurgical pain (50). This approach was necessary to ensure the timely identification of patients with postoperative pain at 3 months. Treatment of pain becomes more difficult once pain is established and becomes chronic. Our study demonstrates that identification of patients with pain early in the recovery trajectory is feasible to undertake (12). Third, the data were cross-sectional so the direction of effects could not be determined. Fourth, our study sample was limited to those with pain 3 months after TKR, which limits the generalizability of our results. When interpreting the baseline factors associated with pain and function, we cannot know if these associations are unique to those with pain 3 months after TKR. Such interpretation is further limited by the lack of preoperative data on our patient cohort, which was not feasible to obtain but would allow further examination of those at higher risk of post-TKR pain. Finally, the measurement tools limit interpretation; although the PainDETECT and DN4 are widely used self-report screening tools for pain with neuropathic characteristics, a detailed clinical examination is recommended to confirm diagnosis (51).

In conclusion, large-scale early screening after TKR identified ongoing pain in a relatively high proportion of people who may benefit from tailored intervention to prevent chronicity. Our study found a high prevalence of pain with neuropathic characteristics and identified several potential intervention targets to improve outcomes for patients with pain at 3 months post-TKR. Research is needed to build on our findings and evaluate multidisciplinary and targeted interventions to improve outcomes for people with pain after TKR.

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. Wylde 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. Wylde, Peters, Howells, Bruce, Eccleston, Gooberman-Hill.

Acquisition of data. Bertram.

Analysis and interpretation of data. Sanderson, Peters.

REFERENCES

1. National Joint Registry. 16th annual report for England, Wales, Northern Ireland and the Isle of Man. Hemel Hempstead: NJR centre; 2019.
2. Scottish Arthroplasty Project. Biennial report. NHS National Services Scotland; 2019.
3. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. BMJ Open 2012;2:e000435.
4. Wylde V, Beswick A, Bruce J, Blom A, Howells N, Gooberman-Hill R. Chronic pain after total knee arthroplasty. EFORT Open Rev 2018;3:461–70.
5. Beswick AD, Wylde V, Gooberman-Hill R. Interventions for the prediction and management of chronic postsurgical pain after total knee replacement: systematic review of randomised controlled trials. BMJ Open 2015;5:e007387.
6. Wylde V, Dennis J, Beswick AD, Bruce J, Eccleston C, Howells N, et al. Systematic review of management of chronic pain after surgery. Br J Surg 2017;104:1293–306.
7. Wylde V, Mackichan, Dixon S, Gooberman-Hill R. Service provision for patients with chronic post-surgical pain after total knee replacement: an evaluation of current practice. J Pain Manag 2014;7:147–54.
8. MacKichan F, Wylde V, Gooberman-Hill R. Pathways through care for long-term pain after knee replacement: a qualitative study with healthcare professionals. Musculoskeletal Care 2015;13:127–38.
9. Judge A, Arden NK, Cooper C, Kassim Javaid M, Carr AJ, Field RE, et al. Predictors of outcomes of total knee replacement surgery. Rheumatology (Oxford) 2012;51:1804–13.
10. Jeffery AE, Wylde V, Blom AW, Horwood JP. "It's there and I'm stuck with it": patients' experiences of chronic pain following total knee replacement surgery. Arthritis Care Res (Hoboken) 2011;63:286–92.
11. Wylde V, Penfold C, Rose A, Blom AW. Variability in long-term pain and function trajectories after total knee replacement: A cohort study. Orthop Traumatol Surg Res 2019;105:1345–50.
12. Turk DC, Wilson HD, Cahana A. Treatment of chronic non-cancer pain. Lancet 2011;377:2226–35.
13. Petersen KK, Simonsen O, Laursen MB, Nielsen TA, Rasmussen S, Arendt-Nielsen L. Chronic postoperative pain after primary and revision total knee arthroplasty. Clin J Pain 2015;31:1–6.
14. Phillips JR, Hopwood B, Stroud R, Dieppe PA, Toms AD. The characterisation of unexplained pain after knee replacement. Br J Pain 2017;11:203–9.
15. Rice DA, Kluger MT, McNair PJ, Lewis GN, Somogyi AA, Borotkanics R, et al. Persistent postoperative pain after total knee arthroplasty: a prospective cohort study of potential risk factors. Br J Anaesth 2018;121:804–12.

16. Wylde V, Bruce J, Beswick A, Elvers K, Gooberman-Hill R. The assessment of chronic post-surgical pain after knee replacement: a systematic review. Arthritis Care Res (Hoboken) 2013;65:1795–803.
17. Howells N, Murray J, Wylde V, Dieppe P, Blom A. Persistent pain after knee replacement: do factors associated with pain vary with degree of patient dissatisfaction? Osteoarthritis Cartilage 2016;24:2061–8.
18. Bafeta A, Dechartres A, Trinquart L, Yavchitz A, Boutron I, Ravaud P. Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. BMJ 2012;344:e813.
19. Pinedo-Villanueva R, Khalid S, Wylde V, Gooberman-Hill R, Soni A, Judge A. Identifying individuals with chronic pain after knee replacement: a population-cohort, cluster-analysis of Oxford knee scores in 128,145 patients from the English National Health Service. BMC Musculoskelet Disord 2018;19:354.
20. Wylde V, Bertram W, Beswick AD, Blom AW, Bruce J, Burston A, et al. Clinical- and cost-effectiveness of the STAR care pathway compared to usual care for patients with chronic pain after total knee replacement: study protocol for a UK randomised controlled trial. Trials 2018;19:132.
21. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg Br 1998;80:63–9.
22. Harris K, Dawson J, Gibbons E, Lim CR, Beard DJ, Fitzpatrick R, et al. Systematic review of measurement properties of patient-reported outcome measures used in patients undergoing hip and knee arthroplasty. Patient Relat Outcome Meas 2016;7:101–8.
23. Harris K, Dawson J, Doll H, Field RE, Murray DW, Fitzpatrick R, et al. Can pain and function be distinguished in the Oxford Knee Score in a meaningful way? An exploratory and confirmatory factor analysis. Qual Life Res 2013;22:2561–8.
24. Wylde V, Mackichan F, Bruce J, Gooberman-Hill R. Assessment of chronic post-surgical pain after knee replacement: development of a core outcome set. Eur J Pain 2015;19:611–20.
25. Cleeland C. The Brief Pain Inventory: user guide. Houston, Texas; 2009.
26. Freynhagen R, Baron R, Gockel U, Tolle TR. PainDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Curr Med Res Opin 2006;22:1911–20.
27. Bouhassira D, Attal N, Alchaar H, Boureau F, Brochet B, Bruxelle J, et al. Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4). Pain 2005;114:29–36.
28. Ware J Jr, Kosinski M, Keller SD. A 12-item Short-Form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220–33.
29. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20:1727–36.
30. Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Qual Life Res 2012; 21:167–76.
31. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–70.
32. Sullivan MJ, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychol Assess 1995;7:524–32.
33. De Vlieger P, Bussche EV, Eccleston C, Crombez G. Finding a solution to the problem of pain: conceptual formulation and the development of the Pain Solutions Questionnaire (PaSol). Pain 2006;123: 285–93.
34. Mahomed N, Sledge CB, Daltroy L, Fossel A, Katz J. Self-administered satisfaction scale for joint replacement arthroplasty. J Bone Joint Surg 1997;80B Suppl:9.
35. Hunt IM, Silman AJ, Benjamin S, McBeth J, Macfarlane GJ. The prevalence and associated features of chronic widespread pain in the community using the 'Manchester' definition of chronic widespread pain. Rheumatology (Oxford) 1999;38:275–9.
36. Hamilton W, Lancashire R, Sharp D, Peters TJ, Cheng K, Marshall T. The risk of colorectal cancer with symptoms at different ages and between the sexes: a case-control study. BMC Med 2009;7:17.
37. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al. The use of the Oxford hip and knee scores. J Bone Joint Surg Br 2007;89:1010–4.
38. Sanderson E, Peters T, Gooberman-Hill R. STAR: Support and treatment after joint replacement statistical analysis plan. University of Bristol Publications Repository. 2017.
39. Glare P, Aubrey KR, Myles PS. Transition from acute to chronic pain after surgery. Lancet 2019;393:1537–46.
40. Haroutiunian S, Nikolajsen L, Finnerup NB, Jensen TS. The neuropathic component in persistent postsurgical pain: a systematic literature review. Pain 2013;154:95–102.
41. Finnerup NB, Attal N, Haroutounian S, McNicol E, Baron R, Dworkin RH, et al. Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. Lancet Neurol 2015;14:162–73.
42. National Institute for Health and Care Excellence. Neuropathic pain in adults: pharmacological management in non-specialist settings. 2013. URL: https://www.nice.org.uk/guidance/cg173.
43. Stubbs B, Aluko Y, Myint PK, Smith TO. Prevalence of depressive symptoms and anxiety in osteoarthritis: a systematic review and meta-analysis. Age Ageing 2016;45:228–35.
44. Gunaratne R, Pratt DN, Banda J, Fick DP, Khan RJ, Robertson BW. Patient dissatisfaction following total knee arthroplasty: a systematic review of the literature. J Arthroplasty 2017;32:3854–60.
45. Leadley RM, Armstrong N, Reid KJ, Allen A, Misso KV, Kleijnen J. Healthy aging in relation to chronic pain and quality of life in Europe. Pain Pract 2014;14:547–58.
46. Fletcher D, Stamer UM, Pogatzki-Zahn E, Zaslansky R, Tanase NV, Perruchoud C, et al. Chronic postsurgical pain in Europe: an observational study. Eur J Anaesthesiol 2015;32:725–34.
47. Eccleston C, Tabor A, Edwards RT, Keogh E. Psychological approaches to coping with pain in later life. Clin Geriatr Med 2016; 32:763–71.
48. Wylde V, Hewlett S, Learmonth ID, Dieppe P. Persistent pain after joint replacement: prevalence, sensory qualities, and postoperative determinants. Pain 2011;152:566–72.
49. Kim J, Lonner JH, Nelson CL, Lotke PA. Response bias: effect on outcomes evaluation by mail surveys after total knee arthroplasty. J Bone Joint Surg Am 2004;86:15–21.
50. Werner MU, Kongsgaard UE. I. Defining persistent post-surgical pain: is an update required? Br J Anaesth 2014;113:1–4.
51. Attal N, Bouhassira D, Baron R. Diagnosis and assessment of neuropathic pain through questionnaires. Lancet Neurol 2018;17:456–66.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 799–808
DOI 10.1002/acr.24512

American College of Rheumatology
Empowering Rheumatology Professionals

Novel Framework for Measuring Whole Knee Osteoarthritis Progression Using Magnetic Resonance Imaging

Jeffrey B. Driban,[1] Lori Lyn Price,[2] Michael P. LaValley,[3] Grace H. Lo,[4] Ming Zhang,[1] Matthew S. Harkey,[1] Amanda Canavatchel,[1] and Timothy E. McAlindon[1]

Objective. We developed and validated a set of composite scores that combine quantitative magnetic resonance imaging (MRI)–based measurements of hyaline cartilage damage, bone marrow lesions (BMLs), and effusion-synovitis into composite scores.

Methods. We selected 300 participants (n = 100 for development cohort; n = 200 for validation cohort) from the Osteoarthritis Initiative with complete clinical, radiographic, and MRI data at baseline and 24 months. We used semiautomated programs to quantify tibiofemoral and patellar cartilage damage, BML volume, and whole-knee effusion-synovitis volume. The candidate composite scores were formed by summing changes from baseline to 24 months based on prespecified methods. We evaluated the candidate composite scores for 1) the ability to differentiate groups with and without knee osteoarthritis progression (17 radiographic and patient-reported definitions), 2) sensitivity to change (standardized response means), and 3) relative performance relating to legacy outcome measures of knee osteoarthritis progression.

Results. Three of 13 developed composite scores qualified for testing in the validation cohort (ranked by sensitivity to change): whole-knee cumulative cartilage damage, unweighted total knee score, and BML plus effusion-synovitis volume. Change in cumulative cartilage damage associated with radiographic progression (Kellgren/Lawrence grade: odds ratio [OR] 1.84; joint space width progression: OR 2.11). Changes in the unweighted total knee score (OR 1.97) and BML plus effusion-synovitis score (OR 1.92) associated with Western Ontario and McMaster Universities Osteoarthritis Index knee pain progression.

Conclusion. Two composite scores emerged, reflecting discrete domains of knee osteoarthritis progression. First, cumulative damage, which is measured by a whole-knee cartilage damage score, reflects the damage accrued over time. Second, dynamic disease activity, which is measured by a BML plus effusion-synovitis score, relates to changes in a patient's state of disease and symptoms.

INTRODUCTION

Osteoarthritis (OA) of the knee is a highly prevalent age-related disorder and a leading cause of pain and functional impairment in the population (1). Despite this, there are few effective interventions for knee OA and none accepted to reduce its structural progression (2). One of the most significant obstacles to testing and developing disease-modifying therapies for knee

The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or the NIH.

This article was prepared using an Osteoarthritis Initiative (OAI) public-use data set, and its contents do not necessarily reflect the opinions or views of the OAI Study Investigators, the NIH, or the private funding partners of the OAI. The OAI is a public–private partnership between the NIH (contracts N01-AR-2-2258, N01-AR-2-2259, N01-AR-2-2260, N01-AR-2-2261, and N01-AR-2-2262) and private funding partners (Merck Research Laboratories, Novartis Pharmaceuticals, GlaxoSmithKline, and Pfizer, Inc.) and is conducted by the OAI Study Investigators. Private sector funding for the OAI is managed by the Foundation for the NIH. The funding sources had no role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; nor in the decision to submit the article for publication.

Supported by a grant from the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS; grant U01-AR-067168), from the NIH/National Center for Advancing Translational Sciences (grant UL1-TR-002544), and from resources at the Center for Innovations in Quality, Effectiveness and Safety (CIN 13-413; Veterans Affairs Health Services Research & Development Service). Dr. Lo's work was supported by the NIH/NIAMS (grant K23-AR-062127) with an NIH/NIAMS funded mentored award, which provided support for design and conduct of the study, analysis, and interpretation of the data. Dr. Harkey's work was supported by the NIH (grant 5-TL1-TR-454-3).

[1]Jeffrey B. Driban, PhD, ATC, CSCS, Ming Zhang, PhD, Matthew S. Harkey, PhD, ATC, Amanda Canavatchel, BS, Timothy E. McAlindon, MD, MPH: Tufts Medical Center, Boston, Massachusetts; [2]Lori Lyn Price, MAS, MLA: Tufts University and Tufts Medical Center, Boston, Massachusetts; [3]Michael P. LaValley, PhD: Boston University School of Public Health, Boston, Massachusetts; [4]Grace H. Lo, MD, MSc: Baylor College of Medicine, Houston, Texas.

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

Address correspondence to Jeffrey Driban, PhD, ATC, CSCS, Division of Rheumatology, Allergy, and Immunology, Tufts Medical Center, 800 Washington Street, Box 406, Boston, MA 02111. Email: Jeffrey.driban@tufts.edu.

Submitted for publication April 27, 2020; accepted in revised form November 10, 2020.

SIGNIFICANCE & INNOVATIONS

- Clinical trials for knee osteoarthritis (OA) urgently need quantitative composite whole-knee scores that address 1) the multifactorial and complex etiopathogenesis of knee osteoarthritis, 2) the discordance between structural changes and signs/symptoms/function, and 3) the lack of consensus for the best definitions of disease progression.
- We offer 2 composite scores that reflect discrete domains of knee OA progression.
- Cumulative damage is defined as the joint damage accrued over time, reflected by a whole-knee cartilage damage score.
- Disease activity is defined as a patient's state of disease and symptoms that may ebb and flow, reflected by a bone marrow lesion plus effusion-synovitis score.

OA is the absence of measures of disease progression that meaningfully reflect a change in patient status (3). For regulatory agencies to accept an imaging biomarker of knee OA progression the biomarker must: 1) reflect the complex pathology throughout the joint, 2) address the apparent discordance between structural changes and signs/symptoms/function, 3) offer a standardized definition of disease progression, and 4) reliably predict meaningful changes in symptoms and function (3).

Semiquantitative scales are comprehensive, but the approach is inherently insensitive to change (4) and limited by the absence of a validated composite measurement reflecting severity in the whole knee (5,6). Quantitative measurements in knees have also been limited by a focus on single structures (e.g., hyaline cartilage) (7–11).

Over the past decade, there has been a paradigm shift from conceptualizing knee OA as a single-structure disorder, based in hyaline cartilage, to a multi-tissue "whole-organ" failure of diarthrodial joints (2). Magnetic resonance imaging (MRI) has revealed structural features, such as bone marrow lesions (BMLs) and effusion-synovitis, which are clinically relevant (12). BMLs reflect altered periarticular bone morphology and density (13,14) and associate with hyaline cartilage damage and pain (12,15,16). BMLs appear to be responsive to clinical intervention (17–19) and have been proposed as a therapeutic target for knee OA disease modification (18,20). Effusion-synovitis is also common in knee OA and is strongly associated with pain (12,21,22) and hyaline cartilage loss (23,24).

Combining these relevant features into a single quantitative composite knee score, rather than measuring each feature separately, could advance the field forward by addressing 1) the multifactorial and complex etiopathogenesis of knee OA, 2) the discordance between structural changes and signs/symptoms/function, and 3) the lack of consensus for the best definitions of disease progression. However, while this is an appealing concept, there are methodologic issues that need to be addressed to determine if a composite score is practical, feasible, and valid. The first question is whether measures of each feature can be combined in a way that is internally (heuristically) consistent. For example, it is unclear if a change in BMLs, which relates to a

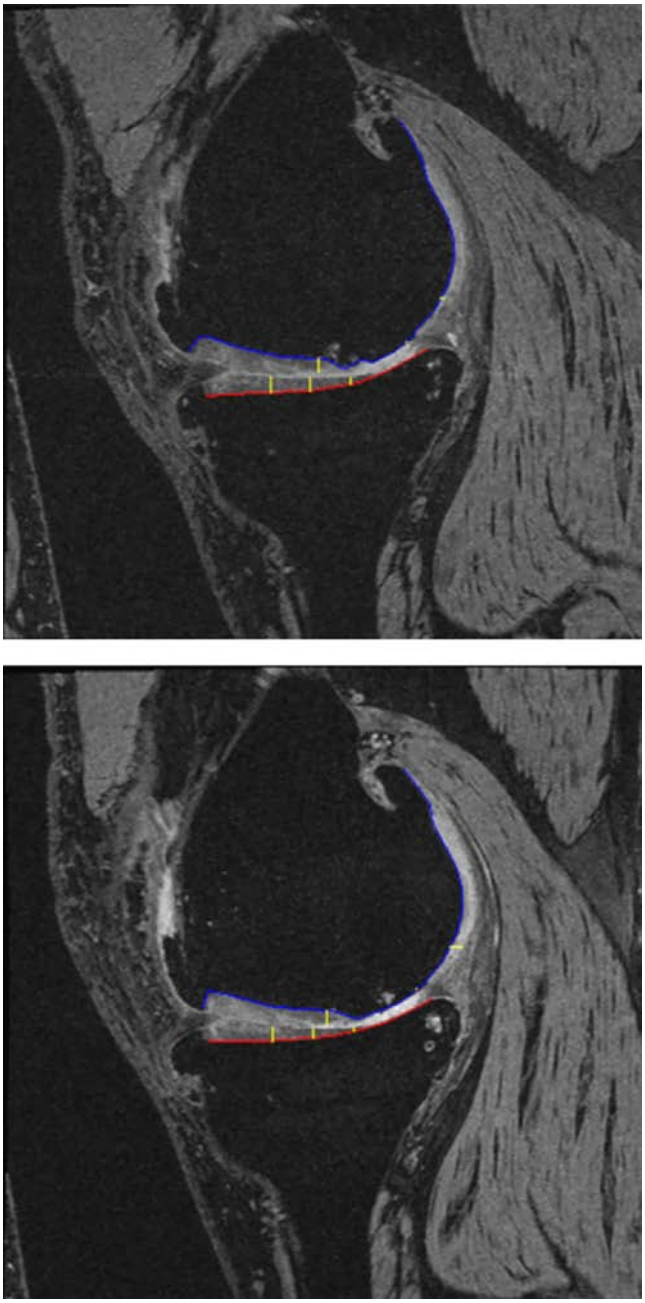

Figure 1. Measurement of medial tibiofemoral cartilage damage index (CDI) on paired baseline (top) and 24-month (bottom) follow-up magnetic resonance images. Images show the medial tibiofemoral CDI with cartilage thickness being assessed at 3 out of 9 informative locations (yellow lines) on a tibia (red line) and femur (blue line). These measurements contribute to the whole-knee cumulative cartilage damage.

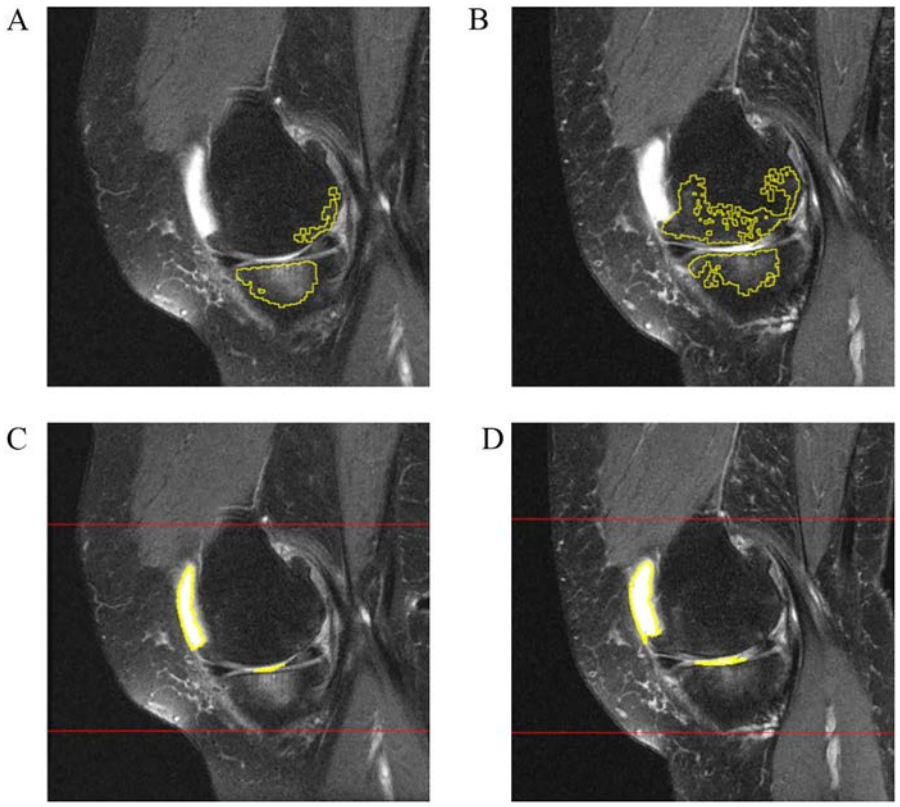

Figure 2. Measurement of medial tibiofemoral bone marrow lesion (BML) volume (mm^3) and whole-knee effusion-synovitis volume (mm^3) on paired baseline (**A** and **C**) and 24-month follow-up (**B** and **D**) magnetic resonance images. Images show a segmented BML (**A** and **B**; yellow) and effusion-synovitis segmentation (**C** and **D**; yellow). These measurements are components of the score based on BML plus effusion-synovitis volume.

change in pain, can or should be combined with a change in hyaline cartilage, which relates to joint space narrowing (JSN). A second challenge is to determine how to mathematically combine the individual features into a single composite score that retains construct validity and optimizes the ability to detect a treatment difference (discriminative validity). Therefore, the goals of this study were to develop and validate candidate composite scores that combine quantitative measurements of hyaline cartilage damage, BMLs, and effusion/synovitis (Figures 1 and 2) into composite scores that are heuristically consistent, and then determine sensitivity to change, and identify the best performing composite scores in relation to established legacy measures of knee OA progression (i.e., joint space width [JSW], Kellgren/Lawrence [K/L] grade, and Western Ontario and McMaster Universities Osteoarthritis Index [WOMAC] pain score).

MATERIALS AND METHODS

Overview. To develop and validate composite, quantitative, MRI-based outcome scores, we selected 2 subcohorts from the Osteoarthritis Initiative (OAI), a multicenter cohort study of individuals with or at risk for symptomatic knee OA in the US. We used a development cohort (n = 100) to develop and perform preliminary analyses to assess construct validity of candidate composite scores based on measures of radiographic progression as well as changes in quality of life and knee pain (Table 1). Then we used a validation cohort (n = 200) to confirm the construct validity of the best-performing composite scores by determining which composite scores exhibit the best discriminative ability and sensitivity to change.

Selection of subcohorts. The OAI included 4 clinical sites (Memorial Hospital of Rhode Island, Ohio State University, University of Maryland and Johns Hopkins University, and the University of Pittsburgh) and recruited participants between February 2004 and May 2006. Each clinical site and the coordinating center obtained institutional review board approval for the OAI. Informed consent was obtained from every participant.

In the course of prior analyses, the OAI investigators formed a Core Image Assessment sample that reflected people typically enrolled in clinical trials. This convenience sample consisted of one knee from 600 participants with symptomatic knee OA at baseline (frequent knee pain and K/L grade 2 or 3) and complete baseline and 24-month data for clinical, radiographic, and MRI outcomes (25). The current project used newer central radiographic readings, which explained why a few people had K/L

grades of 0, 1, or 4. We used the Core Image Assessment sample to assemble our 2 subcohorts for these analyses. We first selected participants for the development cohort. Within each K/L grade, we randomly sampled participants with and without radiographic progression (any increase in K/L grade over 24 months). The process was then repeated for the validation cohort after excluding the participants in the development cohort.

MRI. Each OAI site acquired knee MRI images at baseline and 24-month follow-up visits using identical 3T Trio MRI systems and knee coils (Siemans). The OAI used study-certified, licensed MRI technologists and completed comprehensive quality control measurements to promote quality, provide consistency across sites, and ensure longitudinal uniformity (26,27).

Acquisitions included a 3-dimensional dual echo steady state (3D DESS) sequence (field of view 140 mm, slice thickness 0.7 mm, skip 0 mm, flip angle 25°, echo time 4.7 msec, recovery time 16.3 msec, matric 307 × 384 pixels, x-resolution 0.365 mm, and y-resolution 0.456 mm) and an intermediate-weighted fat suppressed (IWFS) sequence (field of view 160 mm, slice thickness 3 mm, skip 0 mm, flip angle 180°, echo time 30 msec, recovery time 3,200 msec, 313 × 448 matrix, x-resolution 0.357 mm, and y-resolution 0.511 mm).

Radiographic outcomes. OAI participants had annual, bilateral weight-bearing, fixed-flexion posteroanterior knee radiographs taken. Central readers provided K/L grades, medial and lateral JSN grades, and medial JSW measurements. The intrarater agreement for K/L and JSN grades using the weighted kappa ranged from 0.70 to 0.88, and the intraclass correlation coefficient (ICC) for JSW was >0.9 (27). We used the baseline and 24-month data to define K/L grade, JSN, and JSW progression. The operational definitions of progression in these constructs are in Table 1. Data are publicly available at https://data-archive.nimh.nih.gov/oai (kxr_sq_bu00 [version 0.8], kxr_sq_bu00 [version 3.7], kxr_qjsw_duryea00 [version 0.8], and kxr_qjsw_duryea03 [version 3.7]).

Patient-centered outcomes. The WOMAC questionnaire measuring knee-specific pain (0 to 20 scale) was obtained at each OAI visit. We used baseline and 24-month WOMAC pain scores to calculate pain progression. We used the 4 Knee Injury and Osteoarthritis Outcome Score (KOOS) quality of life questions because we were interested in the unique concepts they reflect, as follows: 1) awareness of knee problem, 2) modifying activities, 3) lack of confidence, and 4) difficulty with a knee. The operational definitions of progression in these constructs are in Table 1. Data are publicly available at https://data-archive.nimh.nih.gov/oai (allclinical00 [version 0.2.2] and allclinical03 [version 3.2.1]).

Quantitative measurement of change in cartilage damage. We used a validated approach (28–30) to quantify cartilage damage on the 3D DESS MR images from the baseline and 24-month visits (Figure 1). The Cartilage Damage Index (CDI) generates a single measure from measurements of cartilage thickness obtained from predetermined regions of interest

Table 1. Binary outcomes used for construct validity steps in the development and validation cohorts*

Outcome	Details of change from baseline to 24-month visit
Radiographic	
K/L grade progression	Any change in K/L grade
Lateral JSN progression	Any change in lateral JSN
Medial JSN progression	Any change in medial JSN
JSN progression	Any change in medial or lateral JSN
Medial JSW progression	Change ≥ median change in the development dataset
Quality of life	
KQOL 1 improvement	KQOL item 1 improves from 2 worst responses to 3 best responses (how often aware of problems with knee[s])
KQOL 1 worsens	KQOL item 1 worsens from 3 best responses to 2 worst responses (how often aware of problems with knee[s])
KQOL 2 improvement	KQOL item 2 improves from 2 worst responses to 3 best responses (modified lifestyle to avoid potentially damaging activities to knee[s])
KQOL 2 worsens	KQOL item 2 worsens from 3 best responses to 2 worst responses (modified lifestyle to avoid potentially damaging activities to knee[s])
KQOL 3 improvement	KQOL item 3 improves from 2 worst responses to 3 best responses (how much troubled with lack of confidence in knee[s])
KQOL 3 worsens	KQOL item 3 worsens from 3 best responses to 2 worst responses (how much troubled with lack of confidence in knee[s])
KQOL 4 improvement	KQOL item 4 improves from 2 worst responses to 3 best responses (in general, how much difficulty have with knee[s])
KQOL 4 worsens	KQOL item 4 worsens from 3 best responses to 2 worst responses (in general, how much difficulty have with knee[s])
Pain	
WOMAC pain progression	Pain increased by at least 2 points on 0–20 scale
WOMAC pain progression	Pain increased by at least 3 points on 0–20 scale
WOMAC pain improvement	Pain decreased by at least 2 points on 0–20 scale
WOMAC pain improvement	Pain decreased by at least 3 points on 0–20 scale

* JSN = joint space narrowing; K/L = Kellgren/Lawrence; KQOL = Knee Injury and Osteoarthritis Outcome Score quality of life question; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

distributed across 6 regions in the knee joint: medial and lateral tibia, femur, and patella. We deployed a semiautomated interface to delineate the bone-cartilage boundary on the automatically selected MR images. We also used the interface to automatically identify the predefined informative locations to measure cartilage thickness in each of the 6 regions (9 locations in each tibiofemoral region, 12 locations in patellar regions). The CDI summary measure is calculated by summing the products of cartilage thickness, cartilage length, and voxel size from each of the informative locations. The ICC for the baseline measures ≥0.86 with ≥72 hours between read and re-read (28–31).

Quantitative measurement of change in BMLs. We measured BML volume in the medial and lateral tibia, femur, and patella regions using the IWFS MR sequence and a validated semiautomated software (Figure 2) (32). We defined BMLs as regions of high-signal intensity within a bone that appeared on ≥2 slices and were within 10 mm of the articular surface (15,33,34). The software uses thresholding to delineate suspected BML regions. The reader (AC) manually adjusted the threshold and removed artifacts. All BML segmentation results were reviewed for quality by a second reader (JBD) and adjusted when necessary. Our reader had good intrareader reliability among 20 knees measured twice with ≥48 hours between readings: ICC 0.98 for baseline BML volume.

Quantitative measurement of change in effusion-synovitis volume. We measured whole-knee effusion-synovitis volume using the IWFS MR sequence and a validated semiautomated software, using an approach that has been described previously (Figure 2) (35–37). Effusion-synovitis segmentation results were read by one reader and reviewed for quality by a second reader (MZ), with changes made when necessary. Our reader had good intrareader reliability among 15 knees measured twice with ≥48 hours between readings: ICC 0.87 for baseline effusion-synovitis volume.

Analytic approach. *Standardizing MRI-based measures of change to the same scale.* We generated 13 MRI-based structural measurements: articular cartilage damage measured by CDI (6 measurements: medial and lateral femur, tibia, and patella), BML volume (6 measurements: medial and lateral femur, tibia, and patella), and effusion-synovitis volume (single whole-knee measurement). We adjusted each structural measure for knee size based on femur width (medial to lateral femoral epicondyle). Next, we subtracted the baseline mean and divided by the baseline SD ([measure – baseline mean]/baseline SD) to transform all measurements to the same scale (mean 0 and SD 1). We calculated change in each structural measurement by subtracting the baseline from the 24-month value. We multiplied the CDI change measurements by (–1) to ensure greater change represented worsening disease burden among all the structural measurements.

Development of candidate composite knee scores. In the development cohort, we implemented 4 approaches to select features to combine into 13 candidate composite knee scores.

For the first approach, we created whole-knee composite scores, which included all 13 measures of change in cartilage, BML, and effusion-synovitis. We used 3 mathematical methods to calculate total MRI-based scores: 1) unweighted summation, 2) inverse variance weighting, and 3) weighting according to the eigenvectors from a principal component analysis (PCA).

For the second approach, we used expert judgment based on the existing literature to select MRI-based features that might reasonably be combined: 1) a cumulative cartilage damage score comprising all 6 measures of cartilage change and 2) a BML plus effusion-synovitis composite score using all 7 measures of change in BML and effusion-synovitis volumes. We used an unweighted sum (simple addition) to calculate these 2 candidate composite scores.

For the third approach, we used a data-driven process based on a PCA to identify principal components (i.e., structural measures that cluster together based on correlated measures of change in MRI-based features). The PCA enabled us to determine if structural measures group together according to anatomic location (e.g., medial femoral BML volume and cartilage change) or within a given tissue (e.g., all BML volumes). Each principal component with an eigenvalue of >1 was considered a candidate composite knee score. For each component with an eigenvalue of >1, we calculated a candidate composite knee score by multiplying standardized measures of each MRI-based feature that had a factor loading of >0.4 (absolute value) in a principal component by its eigenvector divided by the sum of the eigenvectors of the contributory components. For the fourth approach, we calculated 2 variables using an unweighted sum of BML or effusion-synovitis volumes in the whole knee.

Preliminary evaluation of the candidate composite score in the development cohort. To prune the set of candidate composite scores before advancing these to the validation sample, we performed preliminary analyses in the development cohort. We evaluated the performance of each candidate composite score to differentiate knee OA progressors from nonprogressors based on a 2-year change in 17 dichotomous measures of pain, quality of life, and radiography (Table 1) using Wilcoxon's rank sum test.

Our decision rule to identify the best-performing composite scores that warranted further validation was that a candidate composite score needed to differentiate ($P < 0.20$ by Wilcoxon's rank sum test [38]) between groups with and without knee OA progression for at least 1 measure in each of the 3 constructs: 5 measures of radiographic progression, 8 measures of quality of life, and 4 measures of knee pain (Table 1).

Confirmation of validity of the best-performing composite scores. In the validation cohort, we examined the best performing

Table 2. Participant characteristics*

	Development cohort (n = 99)	Validation cohort (n = 197)
Female	59 (59.6)	107 (54.3)
Age, mean ± SD years	61.5 ± 8.6	61.2 ± 8.8
BMI, mean ± SD kg/m^2	30.2 ± 4.6	30.1 ± 5.1
K/L grade		
0	1 (1.0)	1 (0.5)
1	9 (9.1)	9 (4.6)
2	40 (40.4)	83 (42.1)
3	47 (47.5)	101 (51.3)
4	2 (2.0)	3 (1.5)
Progression, K/L grade†	29 (29.9)	56 (28.9)
Progression, medial JSW‡	47 (47.5)	92 (46.7)
WOMAC pain, mean ± SD (baseline)	4.4 ± 3.6	5.0 ± 3.6
WOMAC pain, worsen at least 3 points	20 (20.2)	31 (15.7)

* Values are the number (%) unless otherwise indicated. BMI = body mass index; JSW = joint space width; K/L = Kellgren/Lawrence; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† K/L grade progression is any worsening in the K/L score.
‡ Medial JSW progression is more change than the median change.

candidate composite scores from the development cohort. Specifically, we re-assessed their performance based on 3 criteria: 1) an ability to differentiate groups with and without knee OA progression (radiographic and patient-reported), 2) sensitivity to change, and 3) associations with several commonly used outcome measures of knee OA progression. First, to confirm whether a candidate quantitative composite knee score differentiated groups, we evaluated the same 17 dichotomized measures tested in the development cohort (requiring a *P* value of less than 0.05 by Wilcoxon's rank sum test). Second, we calculated sensitivity to change for each candidate composite knee score using a standardized response mean (SRM) and interquartile range (IQR). Finally, we compared the performance of each candidate composite score using the area under the receiver operating characteristic curve (AUC) and odds ratios (ORs) in relation to commonly used outcome measures of knee OA progression: change in K/L grade, JSW, and pain. To help compare our results with a prior biomarker study, we adjusted for the same variables they selected: sex, race, age, body mass index (BMI), K/L grade, WOMAC pain measure, use of pain medications, and minimum JSW (39). The ORs and 95% confidence intervals were calculated using 1 SD as the unit of analysis.

The assumption of linearity of the candidate composite scores with the outcomes was assessed. We used model diagnostics to identify potential outliers and influential points in the multivariable models. *P* values were not adjusted to account for multiple comparisons. We performed all analyses using SAS software, version 9.4.

RESULTS

We removed 1 participant from the development cohort and 3 participants from the validation cohort due to concerns with MRI quality (e.g., metal artifact, pathologic changes unrelated to knee OA). The 99 participants in the development cohort had a mean age of 62 years, mean BMI of 30 kg/m^2, and 60% were female (Table 2). Similarly, the 197 participants in the validation cohort had a mean age of 61 years, mean BMI of 30 kg/m^2, and 54% were female (Table 2).

Development and preliminary evaluation of the candidate composite scores in the development cohort. The development process produced 13 candidate composite scores of structural changes (Table 3). Four of the 13 candidate composite scores met our decision rule based on discriminatory ability to advance as a best-performing composite score (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24512/abstract). These included 1) the unweighted total knee score, 2)

Table 3. Candidate quantitative composite knee scores in the development cohort*

Score	Calculation
Principal component 1†	Sum of CDI in the lateral and medial patella and BML in the medial patella
Principal component 2‡	Sum of BML in lateral and medial tibia
Principal component 3‡	Sum of CDI in the lateral patella, effusion-synovitis volume, and BML in the lateral patella
Principal component 4‡	Sum of CDI in the medial femur and tibia and BML in the medial femur
Principal component 5‡	Sum of BML in lateral femur, medial femur, and lateral patella
Principal component 6‡	Sum of CDI in the lateral femur and tibia
PCA-weighted total	Sum of all 13 measurements, weighted by the principal component loading
Inverse variance weighting total	Sum of all 13 measures, Inverse variance weighted
Unweighted total†	Sum of all 13 measurements (simple addition)
Cumulative cartilage damage†	Sum of all 6 CDI measures
BML plus effusion-synovitis†	Sum of all 6 BML and effusion-synovitis measures
Whole knee BML volume	Sum of all 6 BML measures
Whole knee effusion-synovitis volume	Sum of all effusion-synovitis volumes

* BML = bone marrow lesion; CDI = cartilage damage index; PCA = principal component analysis.
† Met criteria for further testing in the validation cohort.
‡ All knee composite scores based on the principal components analysis are multiplied by the eigenvector for the component and divided by the sum of the eigenvectors of the 6 components.

Table 4. Odds ratios (ORs) and area under the receiver operating characteristic curve (AUC) for composite scores in the validation cohort*

Knee measure	Unadjusted		Adjusted†	
	OR (95% CI)	AUC	OR (95% CI)	AUC
K/L grade progression				
Unweighted total score	1.29 (0.94–1.77)	0.60	1.38 (0.97–1.96)	0.71
Cumulative cartilage damage	1.54 (1.12–2.11)	0.64	1.84 (1.27–2.69)	0.72
BML plus effusion-synovitis	1.17 (0.85–1.61)	0.59	1.21 (0.86–1.70)	0.70
Medial JSW progression				
Unweighted total score	1.20 (0.90–1.61)	0.56	1.21 (0.89–1.64)	0.63
Cumulative cartilage damage	1.93 (1.38–2.72)	0.67	2.11 (1.44–3.07)	0.70
BML plus effusion-synovitis	1.05 (0.79–1.39)	0.52	1.05 (0.78–1.41)	0.63
WOMAC pain progression				
Unweighted total score	1.72 (1.18–2.52)	0.63	1.97 (1.26–3.08)	0.76
Cumulative cartilage damage	1.28 (0.89–1.84)	0.55	1.31 (0.86–1.98)	0.72
BML plus effusion-synovitis	1.69 (1.15–2.48)	0.65	1.92 (1.23–3.01)	0.76

* 95% CI = 95% confidence interval; BML = bone marrow lesion; JSW = joint space width; K/L = Kellgren/Lawrence; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.
† Adjusted for age, sex, body mass index, race, WOMAC pain, K/L grade, JSW, and use of pain medication.

cumulative cartilage damage, 3) BML plus effusion-synovitis volume, and 4) change in the first principal component (patellar damage component).

Confirmation of validity of the best-performing composite scores in the validation cohort. Change in the BML plus effusion-synovitis score was the only candidate composite knee score to differentiate progressors and nonprogressors for at least 1 measure within each of the 3 constructs (see Supplementary Table 1, at http://onlinelibrary.wiley.com/doi/10.1002/acr.24512/abstract). Specifically, change in the BML plus effusion-synovitis score differentiated progressors and nonprogressors for 1 out of 5 radiographic measures, 1 out of 6 quality of life measures, and 3 out of 4 knee pain measures. Change in the cumulative cartilage damage differentiated groups for 4 radiographic measures, 1 quality of life measure, but none of the knee pain measures. Change in the unweighted total knee score differentiated groups for 3 radiographic measures, none of the quality-of-life measures, and 3 knee pain measures. Finally, change in the first principal component only differentiated progressors and nonprogressors for one measure of radiographic progression. Hereafter, we omitted the first principal component from the analyses.

Sensitivity to change in the validation cohort. The cumulative cartilage damage had the greatest sensitivity to change (SRM 1.19; mean ± SD 0.88 ± 0.74), followed by the unweighted total knee score (SRM 0.48; mean ± SD 1.52 ± 3.16) and BML plus effusion-synovitis (SRM 0.22; mean ± SD 0.65 ± 2.96). Alternatively, the unweighted total knee score had the largest IQR of change (IQR 2.49; 25th percentile 0.12, 75th percentile 2.61) followed by BML plus effusion-synovitis (IQR 1.90; 25th percentile –0.43, 75th percentile 1.47) and cumulative cartilage damage (IQR 0.89; 25th percentile 0.40, 75th percentile 1.29).

Associations with commonly used knee OA progression outcomes. Table 4 summarizes the associations between the candidate composite scores and progression in K/L grade, medial JSW, and WOMAC pain. In adjusted models, change in cumulative cartilage damage was the only candidate associated with K/L grade progression (OR 1.84) or JSW progression (OR 2.11). Changes in the unweighted total knee score (OR 1.97) and BML plus effusion-synovitis score (OR 1.92) was associated with WOMAC knee pain progression.

DISCUSSION

We showed that quantitative composite knee scores using quantitative measurements of cartilage damage, BML, and effusion-synovitis have construct validity with radiographic structural progression (cartilage damage score) and worsening of knee pain (unweighted total score and BML plus effusion-synovitis score). We also considered a data-driven approach to identify novel candidate composite scores (e.g., tibial BMLs, patellar pathology); however, these failed to demonstrate adequate performance in the development data set. Furthermore, despite our presumption that one definition of knee OA structural progression may be ideal, we found that 2 composite scores of disease progression may provide a more pragmatic approach to standardize the definitions of knee OA progression. By adopting 2 definitions of disease progression—one for cumulative disease burden and another for more dynamic processes related to patient-centered outcomes—we can overcome some significant obstacles to testing disease-modifying therapies for knee OA. Specifically, these definitions, and their respective scores, address the multifactorial

and complex nature of knee OA, represent structural changes in the whole knee, the discordance between structural changes and signs/symptoms/function, and offer a new framework to define knee OA progression (3).

This work builds on the potential brought by MRI imaging as a powerful outcome tool to measure structures throughout a joint. Thus far, no assessment strategy has met the demand to define structural severity in the whole joint or to reflect changes in important patient-relevant outcomes. Semiquantitative scales have great utility but are insensitive to change (4) and lack a validated composite score to reflect severity in the whole knee (5,6). To overcome these challenges, we deployed our parsimonious validated methods to quantify cartilage loss, BMLs, and effusion-synovitis, which are clinically relevant features of knee OA. Furthermore, the cartilage damage score (representing tibial, femoral, and patellar regions) and BML plus effusion-synovitis score (representing tibial, femoral, and patellar regions and effusion-synovitis throughout the knee) reflect pathologic changes throughout the knee. These new composite scores rely on quantitative measurements, which address concerns about combining semiquantitative scores to create a whole-knee score (5,6). Furthermore, these new whole-knee scores offer an alternative to the typical focus on a single region or compartment of the knee (e.g., medial tibiofemoral) with burdensome quantitative methods.

Based on our findings, the cumulative cartilage damage and BML plus effusion-synovitis composite scores reflect different constructs of knee OA progression. Cumulative cartilage damage represents hyaline cartilage damage throughout the knee, relates to radiographic severity, and reflects the damage attributable to knee OA over the course of the disease. The relationship between cartilage damage with radiographic severity complements prior evidence that regional and compartment-specific semiquantitative and quantitative MRI-based assessments of cartilage damage progression discriminate and predict structural progression defined by radiographs better than progression in patient-reported outcomes (12,25,39). The novelty of the cumulative cartilage damage score is that it is a fully quantitative composite score representing key informative locations throughout the knee, which is optimized to discriminate change. Future studies may clarify whether short-term changes in the cumulative cartilage damage score predict long-term symptoms—a hypothesis supported by regional measures of cartilage loss modestly predicting future total knee replacements (40,41). In contrast to cumulative cartilage damage, the BML plus effusion-synovitis composite score may reflect a more dynamic disease activity process that relates to knee pain and reflects a patient's current state of disease and symptoms. These findings are consistent with prior reports that effusion-synovitis and BMLs relate to changes in patient-reported outcomes (12,32,42). The novelty of the BML plus effusion-synovitis composite score is that it is a fully quantitative composite score that combines BML volumes and effusion-synovitis volumes throughout the knee to maximize the potential to discriminate change. Overall, these findings are consistent with the heuristic used in other rheumatologic diseases that separately measure damage and disease activity (43).

Our study does have some limitations. Meniscal damage, an important but complex component of knee OA, is not included in our composite scores. There is merit in determining the optimal measures of meniscal change among people with knee OA and then working on including the meniscus into future composite scores. However, the exclusion of a meniscal measure does not undermine our primary finding that knee OA progression may be conceptualized as 2 constructs. Future research will clarify if meniscal measurements contribute to the cumulative damage score or if meniscal changes represent another unique construct of knee OA progression. Osteophytes were also not included in our composite scores because their progression is slow and unlikely to change meaningfully in 24 months.

We also observed modest AUCs, possibly reflecting the imprecision in using construct validity in the absence of a gold standard for knee OA progression. For example, we were limited to using radiographic outcomes focused on the tibiofemoral joint despite our focus on MRI-based changes throughout the joint. We also did not account for multiple testing in our significance tests. While we used a sequential testing strategy with development and validation subcohorts to reduce the number of falsely recommended quantitative composite scores, these composite scores need further testing on other data for full validation, including evaluating the internal consistency of the scores. Finally, we demonstrated that the new composite scores outperform whole-knee BML or effusion-synovitis volumes and relate to radiographic outcomes (e.g., JSW change, K/L grade change). However, future studies are needed to compare these new scores to traditional measures of structural progression (e.g., JSW change, mean cartilage thickness change) in regard to predicting patient-centered outcomes and discriminating group differences in clinical trials.

Our study has several strengths. First, we used a robust, longitudinal cohort with high-quality MR images. Second, we used validated semiautomated quantitative measures of cartilage damage (29,30), BML volume (32), and effusion-synovitis volume (35–37) that are less burdensome than other measurement methods. Furthermore, these software programs can be disseminated for use by technicians with a modest training requirement. Finally, we used separate development and validation subcohorts for more rigorous validation of the composite knee scores.

In conclusion, we showed that quantitative composite knee scores based on MRI measures of cartilage damage, BML, and effusion-synovitis are associated with knee OA progression. Our analyses offer a new perspective on knee OA progression reflected in the 2 different domains. First, cumulative damage, which is measured by a cumulative cartilage damage score, is associated with radiographic progression and reflects the

damage attributable to knee OA over the course of the disease. Second, dynamic disease activity, which is measured by a BML plus effusion-synovitis score, relates to changes in a patient's state of disease and symptoms. These domains of knee OA progression, and their respective scores, address the multifactorial and complex nature of knee OA, the discordance between structural changes and signs/symptoms/function, offer a new framework to define knee OA progression, and have promise as MRI-based biomarkers.

ACKNOWLEDGMENT

We want to thank Rachel Sachs for her work in analyzing the effusion-synovitis data.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Driban 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. Driban, Price, LaValley, Lo, Zhang, Harkey, Canavatchel, McAlindon.

Acquisition of data. Driban, Zhang, Canavatchel, McAlindon.

Analysis and interpretation of data. Driban, Price, LaValley, Lo, Zhang, Canavatchel, McAlindon.

REFERENCES

1. Cisternas MG, Murphy L, Sacks JJ, Solomon DH, Pasta DJ, Helmick CG. Alternative methods for defining osteoarthritis and the impact on estimating prevalence in a US population-based survey. Arthritis Care Res (Hoboken) 2016;68:574–80.
2. Lane NE, Brandt K, Hawker G, Peeva E, Schreyer E, Tsuji W, et al. OARSI-FDA initiative: defining the disease state of osteoarthritis. Osteoarthritis Cartilage 2011;19:478–82.
3. Guidance for industry: osteoarthritis. Structural endpoints for the development of drugs, devices, and biological products for treatment. Rockville (MD): US Food and Drug Administration; 2018.
4. Eckstein F, Nevitt M, Gimona A, Picha K, Lee JH, Davies RY, et al. Rates of change and sensitivity to change in cartilage morphology in healthy knees and in knees with mild, moderate, and end-stage radiographic osteoarthritis: results from 831 participants from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2011;63:311–9.
5. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage;19:990–1002.
6. Conaghan PG, Tennant A, Peterfy CG, Woodworth T, Stevens R, Guermazi A, et al. Examining a whole-organ magnetic resonance imaging scoring system for osteoarthritis of the knee using Rasch analysis. Osteoarthritis Cartilage 2006;14 Suppl A:A116–21.
7. Hunter DJ, Conaghan PG, Peterfy CG, Bloch D, Guermazi A, Woodworth T, et al. Responsiveness, effect size, and smallest detectable difference of magnetic resonance imaging in knee osteoarthritis. Osteoarthritis Cartilage 2006;14 Suppl A:A112–5.
8. Duryea J, Neumann G, Niu J, Totterman S, Tamez J, Dabrowski C, et al. Comparison of radiographic joint space width with magnetic resonance imaging cartilage morphometry: analysis of longitudinal data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2010; 62:932–7.
9. Hunter DJ, Niu J, Zhang Y, Totterman S, Tamez J, Dabrowski C, et al. Change in cartilage morphometry: a sample of the progression cohort of the Osteoarthritis Initiative. Ann Rheum Dis 2009;68:349–56.
10. Wirth W, Nevitt M, Hellio Le Graverand MP, Benichou O, Dreher D, Davies RY, et al. Sensitivity to change of cartilage morphometry using coronal FLASH, sagittal DESS, and coronal MPR DESS protocols: comparative data from the Osteoarthritis Initiative (OAI). Osteoarthritis Cartilage 2010;18:547–54.
11. Eckstein F, Cotofana S, Wirth W, Nevitt M, John MR, Dreher D, et al. Painful knees have greater rates of cartilage loss than painless knees after adjusting for radiographic disease stage: data from the OA initiative. Arthritis Rheum 2011;63:2257–67.
12. Hunter DJ, Zhang W, Conaghan PG, Hirko K, Menashe L, Li L, et al. Systematic review of the concurrent and predictive validity of MRI biomarkers in OA. Osteoarthritis Cartilage 2011;19:557–88.
13. Kazakia GJ, Kuo D, Schooler J, Siddiqui S, Shanbhag S, Bernstein G, et al. Bone and cartilage demonstrate changes localized to bone marrow edema-like lesions within osteoarthritic knees. Osteoarthritis Cartilage 2013;21:94–101.
14. Zanetti M, Bruder E, Romero J, Hodler J. Bone marrow edema pattern in osteoarthritis knees: correlation between MR imaging and histologic findings. Radiology 2000;215:835–40.
15. Driban JB, Lo GH, Lee JY, Ward RJ, Miller E, Pang J, et al. Quantitative bone marrow lesion size in osteoarthritic knees correlates with cartilage damage and predicts longitudinal cartilage loss. BMC Musculoskelet Disord 2011;12:217–10.
16. Felson DT, Chaisson CE, Hill CL, Totterman SM, Gale ME, Skinner KM, et al. The association of bone marrow lesions with pain in knee osteoarthritis. Ann Intern Med 2001;134:541–9.
17. Callaghan MJ, Parkes MJ, Hutchinson CE, Gait AD, Forsythe LM, Marjanovic EJ, et al. A randomised trial of a brace for patellofemoral osteoarthritis targeting knee pain and bone marrow lesions. Ann Rheum Dis 2015;74:1164–70.
18. Laslett LL, Dore DA, Quinn SJ, Boon P, Ryan E, Winzenberg TM, et al. Zoledronic acid reduces knee pain and bone marrow lesions over 1 year: a randomised controlled trial. Ann Rheum Dis 2012;71: 1322–8.
19. Pelletier JP, Roubille C, Raynauld JP, Abram F, Dorais M, Delorme P, et al. Disease-modifying effect of strontium ranelate in a subset of patients from the Phase III knee osteoarthritis study SEKOIA using quantitative MRI: reduction in bone marrow lesions protects against cartilage loss. Ann Rheum Dis 2015;74:422–9.
20. Carbone LD, Nevitt MC, Wildy K, Barrow KD, Harris F, Felson D, et al. The relationship of antiresorptive drug use to structural findings and symptoms of knee osteoarthritis. Arthritis Rheum 2004;50:3516–25.
21. Torres L, Dunlop DD, Peterfy C, Guermazi A, Prasad P, Hayes KW, et al. The relationship between specific tissue lesions and pain severity in persons with knee osteoarthritis. Osteoarthritis Cartilage 2006;14: 1033–40.
22. Yusuf E, Kortekaas MC, Watt I, Huizinga TW, Kloppenburg M. Do knee abnormalities visualised on MRI explain knee pain in knee osteoarthritis? A systematic review. Ann Rheum Dis 2011;70:60–7.
23. Roemer FW, Guermazi A, Felson DT, Niu J, Nevitt MC, Crema MD, et al. Presence of MRI-detected joint effusion and synovitis increases the risk of cartilage loss in knees without osteoarthritis at 30-month follow-up: the MOST study. Ann Rheum Dis 2011;70:1804–9.
24. Roemer FW, Kwoh CK, Hannon MJ, Green SM, Jakicic JM, Boudreau R, et al. Risk factors for magnetic resonance imaging-detected patellofemoral and tibiofemoral cartilage loss during a six-month period: the joints on glucosamine study. Arthritis Rheum 2012;64:1888–98.

25. Eckstein F, Collins JE, Nevitt MC, Lynch JA, Kraus VB, Katz JN, et al. Brief report: cartilage thickness change as an imaging biomarker of knee osteoarthritis progression. Data from the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol 2015;67:3184–9.
26. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis Cartilage 2008;16:1433–41.
27. The Osteoarthritis Initiative. 2019. URL: https://nda.nih.gov/oai/.
28. Zhang M, Canavatchel AR, Driban JB, Price L, Moreno-koehler A, Lo GH, et al. Patellar cartilage change is not associated with difficulty with stairs and crepitus: a two year longitudinal study of data from the osteoarthritis initiative. Osteoarthritis Cartilage 2016;24:S356–7.
29. Zhang M, Driban JB, Price LL, Harper D, Lo GH, Miller E, et al. Development of a rapid knee cartilage damage quantification method using magnetic resonance images. BMC Musculoskelet Disord 2014; 15:264.
30. Zhang M, Driban JB, Price LL, Lo GH, Miller E, McAlindon TE. Development of a rapid cartilage damage quantification method for the lateral tibiofemoral compartment using magnetic resonance images: data from the Osteoarthritis Initiative. BioMed Res Int 2015;2015: 634275.
31. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420–8.
32. Zhang M, Driban JB, Price LL, Lo GH, McAlindon TE. Magnetic resonance image sequence influences the relationship between bone marrow lesions volume and pain: data from the Osteoarthritis Initiative. BioMed research international 2015;2015:731903.
33. Roemer FW, Frobell R, Hunter DJ, Crema MD, Fischer W, Bohndorf K, et al. MRI-detected subchondral bone marrow signal alterations of the knee joint: terminology, imaging appearance, relevance and radiological differential diagnosis. Osteoarthritis Cartilage 2009;17:1115–31.
34. Roemer FW, Khrad H, Hayashi D, Jara H, Ozonoff A, Fotinos-Hoyer AK, et al. Volumetric and semiquantitative assessment of MRI-detected subchondral bone marrow lesions in knee osteoarthritis: a comparison of contrast-enhanced and non-enhanced imaging. Osteoarthritis Cartilage 2010;18:1062–6.
35. Davis JE, Ward RJ, MacKay JW, Lu B, Price LL, McAlindon TE, et al. Effusion-synovitis and infrapatellar fat pad signal intensity alteration differentiate accelerated knee osteoarthritis. Rheumatology (Oxford) 2019;58:418–26.
36. Harkey MS, Price LL, McAlindon TE, Davis JE, Stout AC, Lu B, et al. Association between declining walking speed and increasing bone marrow lesion and effusion volume in individuals with accelerated knee osteoarthritis. Arthritis Care Res (Hoboken) 2019;71:259–70.
37. Stout AC, Barbe MF, Eaton CB, Amin M, Al-Eid F, Price LL, et al. Inflammation and glucose homeostasis are associated with specific structural features among adults without knee osteoarthritis: a cross-sectional study from the osteoarthritis initiative. BMC Musculoskelet Disord 2018;19:1.
38. Heinze G, Dunkler D. Five myths about variable selection. Transpl Int 2017;30:6–10.
39. Collins JE, Losina E, Nevitt MC, Roemer FW, Guermazi A, Lynch JA, et al. Semiquantitative imaging biomarkers of knee osteoarthritis progression: data from the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. Arthritis Rheumatol 2016;68: 2422–31.
40. Kwoh CK, Guehring H, Aydemir A, Hannon MJ, Eckstein F, Hochberg MC. Predicting knee replacement in participants eligible for disease-modifying osteoarthritis drug treatment with structural endpoints. Osteoarthritis Cartilage 2020;28:782–91.
41. Hitzl W, Wirth W, Maschek S, Cotofana S, Nevitt M, John MR, et al. Greater lateral femorotibial cartilage loss in osteoarthritis initiative participants with incident total knee arthroplasty: a prospective cohort study. Arthritis Care Res (Hoboken) 2015;67:1481–6.
42. Driban JB, Price L, Lo GH, Pang J, Hunter DJ, Miller E, et al. Evaluation of bone marrow lesion volume as a knee osteoarthritis biomarker: longitudinal relationships with pain and structural changes: data from the Osteoarthritis Initiative. Arthritis Res Ther 2013;15:R112.
43. Fransen J, van Riel PL. Outcome measures in inflammatory rheumatic diseases. Arthritis Res Ther 2009;11:244.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 809–817
DOI 10.1002/acr.24530

American College of Rheumatology
Empowering Rheumatology Professionals

Osteitis in Systemic Sclerosis: A Nationwide Case–Control Retrospective Study

Cyril Cosse,[1] Solen Kernéis,[2] Alain Lescoat,[3] Gregory Pugnet,[4] Marie-Elise Truchetet,[5] Pascal Priollet,[6] Elisabeth Diot,[7] Mickael Martin,[8] François Maurier,[9] Jean François Viallard,[10] Christian Agard,[11] Brigitte Granel,[12] Sabine Berthier,[13] Dorothée Fagedet,[14] Bénédicte Watelet,[15] Ségolène Toquet,[16] David Luque Paz,[17] Chloé Giret,[7] Olivier Cerles,[1] Jérémie Dion,[1] Christelle Nguyen,[18] Loïc Raffray,[19] Julien Bertolino,[12] Wendy Jourde,[10] Claire Le Jeunne,[1] Luc Mouthon,[1] and Benjamin Chaigne[1]

Objective. Systemic sclerosis (SSc) is an autoimmune connective tissue disorder characterized by skin fibrosis, vasculopathy, and dysimmunity. Data regarding osteitis in SSc are scarce.

Methods. We performed a nationwide multicenter, retrospective, case–control study including patients with SSc, according to the 2013 American College of Rheumatology/European Alliance of Associations for Rheumatology classification, with a diagnosis of osteitis. The objectives of the study were to describe, to characterize, and to identify associated factors for osteitis in patients with SSc.

Results. Forty-eight patients were included. Twenty-six patients (54.1%) had osteitis beneath digital tip ulcers. Physical symptoms included pain (36 of 48, 75%), erythema (35 of 48, 73%), and local warmth (35 of 48, 73%). Thirty-one (65%) patients had median (interquartile range) C-reactive protein levels >2 mg/liter of 8 (2.7–44.3) mg/liter. On radiography, computed tomography, or magnetic resonance imaging, osteitis was characterized by swelling or abscess of soft tissues, with acro-osteolysis or lysis in 28 patients (58%). Microbiological sampling was performed in 45 (94%) patients. Most pathogens were *Staphylococcus aureus* (43.8%), anaerobes and Enterobacteriaceae (29.1%), and *Pseudomonas aeruginosa* (10.4%). Management comprised antibiotics in 37 (77.1%) patients and/or surgery in 26 (54.2%). Fluoroquinolones were used in 22 (45.8%) patients, and amoxicillin plus β-lactamase inhibitor in 7 (14.6%). Six (12.6%) patients relapsed, 6 (12.6%) patients had osteitis recurrence, 15 (32%) sequelae, and 2 patients had septic shock and died.

Conclusion. This study confirmed digital tip ulcers as an associated factor for osteitis and revealed a high rate of functional sequelae. Antimicrobial therapy with oral fluoroquinolone or intravenous amoxicillin and β-lactamase inhibitor are used as first-line antibiotic therapy in SSc patients with osteitis.

INTRODUCTION

Systemic sclerosis (SSc) is an autoimmune connective tissue disorder characterized by skin sclerosis due to fibroblast activation and extracellular matrix synthesis, vasculopathy including vascular hyperactivity and remodeling, and dysimmunity (1–3). Because of microangiopathy, patients with SSc have a compromised vascular supply and present with Raynaud's

[1]Cyril Cosse, PhD, Olivier Cerles, PhD, Jérémie Dion, MD, Claire Le Jeunne, MD, PhD, Luc Mouthon, MD, PhD, Benjamin Chaigne, MD, PhD: Service de Médecine Interne, Centre de Référence Maladies Systémiques Auto-Immunes Rares d'Ile de France, Hôpital Cochin, Université Paris Descartes, Assistance Publique-Hôpitaux de Paris, Paris, France; [2]Solen Kernéis, MD, PhD: Assistance Publique-Hôpitaux de Paris, Cochin Hospital, University of Paris, Paris, France; [3]Alain Lescoat, MD, PhD: Université Rennes, CHU Rennes, Inserm, EHESP, Institut de Recherche en Santé, Environnement et Travail, UMR S 1085, and Department of Internal Medicine and Clinical Immunology, Rennes University Hospital, Rennes, France; [4]Gregory Pugnet, MD, PhD: Service de Médecine Interne, CHU Toulouse, and Centre d'Investigation Clinique 1436, CHU Toulouse, Toulouse, France; [5]Marie-Elise Truchetet, MD, PhD: Rheumatology Department and National Center of Reference for Rare Autoimmune Diseases, Bordeaux University Hospital, Bordeaux, France; [6]Pascal Priollet, MD: Service de Médecine Vasculaire, Hôpital Saint Joseph, Paris, France; [7]Elisabeth Diot, MD, PhD, Chloé Giret, MD: Service de Médecine Interne, Hôpital Bretonneau, Tours, France; [8]Mickael Martin, MD, PhD: Service de Médecine Interne, Maladies Infectieuses et Tropicales, CHU de Poitiers, Poitiers, France; [9]François Maurier, MD: Centre de Compétence des Maladies Rares, Hôpitaux privés de Metz, Metz, France; [10]Jean François Viallard, MD, PhD, Wendy Jourde, MD: Département de Médecine Interne et Maladies Infectieuses, Centre Hospitalier Universitaire Haut Lévêque, Université de Bordeaux, Pessac, France; [11]Christian Agard, MD, PhD: Service de Médecine Interne, Hôtel-Dieu, CHU Nantes, Hôpital, Université de Nantes, Nantes, France; [12]Brigitte Granel, MD, PhD, Julien Bertolino, MD: Service de Médecine Interne, Hôpital Nord, Aix-Marseille Université, Assistance Publique-Hôpitaux de Marseille, Marseille, France; [13]Sabine Berthier, MD: Service de Médecine Interne et Immunologie Clinique, CHU F. Mitterrand, Dijon, France; [14]Dorothée Fagedet, MD: Service de Médecine Interne, Centre Hospitalier Intercommunal des Alpes du Sud, Hôpital de Gap, Gap, France; [15]Bénédicte Watelet, MD: Service de Médecine Interne, CHU Caen, Caen, France; [16]Ségolène Toquet, MD: Pôle Médecine, CHU Robert Debré, Reims, France; [17]David Luque Paz, MD, PhD: Service des Maladies Infectieuses et Réanimation Médicale, CHU Rennes, Rennes, France; [18]Christelle Nguyen, MD, PhD: Service de Rééducation et de Réadaptation de l'Appareil Locomoteur et des Pathologies du Rachis, Hôpitaux Universitaires Paris

SIGNIFICANCE & INNOVATIONS

- 54.1% of osteitis in patients with systemic sclerosis (SSc) occurred beneath digital ulcers. The most frequent pathogens in SSc osteitis were *Staphylococcus aureus*, anaerobes and Enterobacteriaceae, and *Pseudomonas aeruginosa*.
- SSc osteitis is associated with 25.2% of relapse or recurrence, 32% of sequelae, and 4.2% of deaths due to septic shock.
- Antibiotic biotherapy with either oral fluoroquinolone or intravenous amoxicillin and β-lactamase inhibitor was mainly used as a first-line antibiotic therapy in SSc patients with osteitis.

phenomenon and digital ulcers, which put them at risk of osteitis (4–7).

Osteitis is an infection resulting from the contamination of the bone by ≥1 pathogens. Such infection results either from direct inoculation or from secondary hematogenous spread. It causes an inflammatory reaction with the secretion of proinflammatory cytokines such as tumor necrosis factor (8). Inflammatory cytokines then stimulate the activation of osteoclasts and osteoblasts, which leads to bone resorption and periosteal reaction, respectively. In case of prolonged osteitis or in patients with diabetic foot infection, ischemic lesions enhance the occurrence of an infection and favor bone necrosis and osteolysis (9). Patients are subsequently at risk of gangrene and amputation.

Gangrene has been reported in SSc and occurs mainly in smoking, anti–Scl-70 positive patients with digital ulcers, and those with a history of previous infection, including osteitis (10). Oppositely, osteitis has been described less in SSc. Even though it is a severe infection, which inevitably impacts patients' quality of life and is responsible for high morbidity, very few studies have focused on its description and management in patients with SSc. Herein, we report on the nationwide SCLEROS study, which aimed at characterizing osteitis in patients with SSc.

PATIENTS AND METHODS

Study design. SCLEROS is a retrospective, multicenter, nationwide case–control study performed in France. Between January and June 2019, physicians belonging to the following scientific societies were asked to recall and report all cases with SSc and osteitis managed between 2000 and 2019: Groupe Francophone de Recherche sur la Sclérodermie (GFRS); Société Nationale Française de Médecine Interne (SNFMI); Société de Pathologie Infectieuse de Langue Française (SPILF); and Club Rhumatismes et Inflammations (CRI). The study was approved by the local ethics committee (CER Paris Descartes; number 2018-93).

Objectives. The primary objective of the study was to describe and characterize osteitis in SSc patients. The secondary objective was to identify associated factors for osteitis in patients with SSc.

Study population. Inclusion criteria were as follows: male or female patients >18 years of age; diagnosis of SSc according to the 2013 American College of Rheumatology (ACR)/European Alliance of Associations for Rheumatology classification (EULAR) (11); diagnosis of osteitis according to the physician in charge of the patient, suspected on the basis of clinical and laboratory findings and confirmed by radiologic findings or surgical bacteriologic sampling, namely, the presence of clinical signs or symptoms (local pain, swelling, erythema around the skin ulcer, warmth, purulent discharge, fever), blood chemistry alterations (C-reactive protein [CRP] level and/or increased white blood cell count), positive microbiological surgical sample, or radiologic findings. The latter included the following: soft-tissue swelling; periosteal reaction or elevation; loss of cortex with bone erosion; focal loss of trabecular pattern; new bone formation, bone sequestrum, involucrum, and/or cloacae (12,13). For each case, 4 control patients with SSc according to the 2013 ACR/EULAR classification without osteitis were randomly extracted from the SSc database of the national referral center for autoimmune and systemic disease (14,15).

Data source. Data comprised a subset classification of SSc (limited cutaneous SSc or diffuse cutaneous SSc) according to LeRoy and Medsger classification (16), disease duration, clinical profiles (modified Rodnan skin thickness score [MRSS], digital vasculopathy, and other visceral impairment), antibody profile, and patients' comorbidities. Collected data related to osteitis were the following: location, clinical and biological presentation of osteitis, and microbiological data and management. Superficial samples were defined as thin-needle aspiration, while deep samples were defined as bone biopsies performed during surgery. A new episode of osteitis was defined as osteitis at a different location within 6 months.

Centre, Groupe Hospitalier Cochin, Assistance Publique-Hôpitaux de Paris, Université de Paris, Faculté de Santé, UFR Médecine Paris Descartes, Sorbonne Paris Cité, INSERM UMRS 1124, Toxicité Environnementale, Cibles Thérapeutiques, Signalisation Cellulaire et Biomarqueurs, UFR Sciences Fondamentales et Biomédicales, Centre Universitaire des Saints-Pères, Paris, France; [19]Loïc Raffray, MD, PhD: Service Médecine Interne et Dermatologie, CHU Réunion, Hôpital Félix Guyon, Saint Denis, France.

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

Address correspondence to Benjamin Chaigne, MD, PhD, Service de Médecine Interne, Centre de Référence Maladies Systémiques Auto-Immunes Rares d'Ile de France, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris, 75014 Paris, France. Email: benjamin.chaigne@aphp.fr.

Submitted for publication June 18, 2020; accepted in revised form December 1, 2020.

Statistical analysis. Quantitative variables were expressed as the mean ± SD or (for a non-Gaussian distribution) the median (interquartile range [IQR]) and compared using the Mann-Whitney test. Qualitative variables were expressed as the number and percentage and compared using the chi-square test.

Associated factors for osteitis occurrence and osteitis healing were identified according to 2-step logistic regression. First variables with a *P* value <0.15 in the univariate analysis were included in the multivariate analysis. These variables were considered as independent associated factors if the *P* value after the multivariate analysis was ≤0.05.

Statistical analysis was performed using Stata software, version 15.1. Results are reported according to the STROBE guidelines for reporting data of observational studies (17).

RESULTS

Forty-eight patients with SSc and osteitis were included in the SCLEROS study. No patient was removed from the analysis because of lacking data. The baseline characteristics of the populations are shown in Table 1. SCLEROS cases comprised typical SSc patients, with 77.1% being female patients (median age 58.5 years [IQR 48.5–68.5 years]). Eighty-five percent had puffy fingers, 79.4% sclerodactyly, 95.8% Raynaud's phenomenon, and 54.2% abnormal nailfold capillaries on capillaroscopy. Fifty-two percent of patients had positivity for anticentromere autoantibodies, and 6.3% had positivity for anti–RNA polymerase III autoantibodies. Only 25% patients had been treated with immunosuppressant drugs at the time of osteitis. Five (10.4%)

Table 1. Patients' baseline characteristics*

Characteristic	Cases (n = 48)	Controls (n = 192)	*P*
Age, median (IQR) years	58.5 (48.5–68.5)	56 (42.5–68)	0.09
Female sex	37 (77.1)	162 (84.4)	0.23
Ethnicity			0.34
White	43 (89.6)	139 (90.9)	
Black	3 (6.3)	9 (5.9)	
Asian	1 (2.1)	5 (3.3)	
Hispanic	1 (2.1)	0 (0)	
BMI, median (IQR) kg/m^2	21.1 (19.7–23.7)	23.0 (20.1–26.4)	0.94
Active smokers	17 (35.4)	49 (32.5)	0.70
Typical features of SSc			0.03
dSSc	17 (35.4)	40 (20.8)	
lSSc	31 (64.6)	152 (79.2)	
Duration of SSc, median (IQR) months	110 (30–204)	44 (17–109)	0.004
Puffy fingers	41 (85.4)	151 (80.8)	0.46
Sclerodactyly	38 (79.2)	137 (71.7)	0.54
Mouth opening, median (IQR) mm	35 (34.5–40)	40 (35–42)	0.95
MRSS, median (IQR)	7 (4–18)	4 (2–8)	0.008
dSSc, median (IQR)	18 (15.5–21.5)	15 (8–22.5)	
lSSc, median (IQR)	4 (3–8)	3 (0–6)	
Raynaud's phenomenon	46 (95.8)	188 (97.9)	0.41
Abnormal nailfold capillaries	26 (54.2)	88 (45.8)	0.30
Digital tip ulcers			0.87
Active	20 (41.7)	27 (19.7)	
Healed	10 (20.8)	27 (19.7)	
Telangiectasia	23 (47.9)	89 (53.6)	0.49
Calcinosis	17 (35.4)	38 (23.5)	0.09
Gastroesophageal reflux disease	32 (66.7)	112 (59.3)	0.35
Interstitial lung disease	17 (35.4)	78 (53.4)	0.03
DLco (% of predicted), median (IQR)	63.5 (54–72)	64 (43–76)	0.15
FVC (% of predicted), median (IQR)	90.5 (82–102)	91 (74–112)	0.66
Pulmonary arterial hypertension	4 (8.3)	22 (15.3)	0.22
Systolic PAP, median (IQR) mm Hg	33 (30–39)	31.5 (27–37)	0.30
Musculoskeletal disorders	19 (39.6)	60 (32.6)	0.36
Scleroderma renal crisis	4 (8.3)	22 (15.3)	0.22
Autoantibodies†			
Anticentromere	25 (52.1)	73 (38.0)	0.08
Anti–topoisomerase I	11 (22.9)	46 (23.9)	0.88
Anti–RNA polymerase III	3 (6.3)	6 (3.1)	0.31
Immunosuppressants	12 (25.0)	54 (28.1)	0.67
Glucocorticoids	18 (37.5)	68 (47.6)	0.23

* Values are the number (%) unless indicated otherwise. BMI = body mass index; DLco = diffusing capacity for carbon monoxide; dSSc = diffuse cutaneous systemic sclerosis; FVC = forced vital capacity; IQR = interquartile range; lSSc = limited cutaneous SSc; MRSS = modified Rodnan skin thickness score; PAP = pulmonary artery pressure.
† No. (% positive in patients tested).

patients were treated with bosentan (3 with digital tip ulcers, and 2 without).

When compared to SSc controls (n = 192), SCLEROS cases had a higher MRSS (median 7 [IQR 4–18] versus median 4 [IQR 2–8], P = 0.008). The occurrence of interstitial lung disease was less frequent (35.4% versus 53.4%; P = 0.03). Osteitis occurred mostly in patients with a long disease duration of SSc (median 110 months [IQR 30–204 months] versus median 44 months [IQR 17–109 months], P = 0.004) and predominantly in patients with diffuse cutaneous SSc (35.4% versus 20.8%; P = 0.03).

Table 2. Osteitis characteristics in patients with systemic sclerosis (SSc)*

Characteristic	SCLEROS cases (n = 48)	Patients with osteitis overlying digital tip ulcers (n = 26)	Patients with osteitis without digital tip ulcers (n = 22)
Clinical manifestations			
Time from SSc diagnosis, median (IQR) months	113 (27–201)	90 (6–120)	144 (42–240)
dSSc	48 (24–120)	96 (26–131)	45 (28.5–48)
lSSc	121 (63–210)	108 (48–120)	144 (122–240)
Osteitis location			
Toe (R/L)	19 (39.6)	7 (26.9)	12 (54.6)
Hand (R/L)	25 (52.1)	18 (69.2)	7 (31.8)
Elbow (R/L)	2 (4.2)	1 (3.9)	1 (4.6)
Other location			
Femur	1 (2.1)	0 (0)	1 (2.1)
Sternum	1 (2.1)	0 (0)	1 (2.1)
Overlying calcinosis	13 (27.1)	8 (30.8)	5 (22.7)
Duration of osteitis, median (IQR) months	3 (1–6)	3 (1–8)	2.5 (1–4)
Fever (>37.8°C)	4 (8.3)	2 (7.7)	2 (9.1)
Local pain	36 (75)	22 (84.6)	14 (63.6)
Functional disability	29 (60.4)	17 (65.4)	12 (54.6)
Purulent discharge	27 (56.3)	15 (57.7)	12 (54.6)
Local swelling	26 (54.2)	19 (73.1)	7 (31.8)
Erythema	35 (72.9)	22 (84.6)	13 (59.1)
Local warmth	35 (72.9)	20 (76.9)	15 (68.2)
Biological manifestations			
CRP, median (IQR) mg/liter	8 (2.7–44.3)	7.2 (2.7–37.5)	12 (2.7–60.0)
Increased CRP value	31 (65.9)	15 (60.0)	16 (72.7)
Hyperleukocytosis, >8,000/mm^3	13 (27.1)	8 (30.8)	5 (22.7)
Leukocytes, median (IQR) mm^3	7,300 (5,000–8,890)	7,400 (5,435–9,820)	6,385 (4,390–7,900)
Hemoglobin, median (IQR) gm/dl	11.8 (10.6–12.8)	12.3 (11.1–12.9)	11.4 (10.6–12.5)
Platelets, median (IQR) 10^3/liter	285 (218–341)	284.5 (215–341)	288 (219–310)
Imaging manifestations			
Localized collection	3 (6.3)	0 (0)	3 (13.6)
Bone hypermetabolism	3 (6.3)	0 (0)	3 (13.6)
Acro-osteolysis or lysis	23 (47.9)	10 (38.5)	13 (59.1)
Swelling or abscess of soft tissues on MRI or standard radiograph			
With acro-osteolysis	28 (58.3)	14 (53.8)	14 (63.6)
Without acro-osteolysis	3 (6.3)	2 (7.8)	1 (4.5)
Sampling			
Intraoperative contributive superficial or needle-guided sampling†	14/18 (77.8)	8/10 (80.0)	6/8 (75.0)
Contributive deep or surgical sampling†	24/27 (88.9)	14/15 (93.3)	10/12 (83.3)
Microbiological documentation			
Staphylococcus aureus	21 (43.8)	15 (57.7)	6 (27.3)
Coagulase-negative *Staphylococcus*	4 (8.3)	3 (11.5)	1 (4.5)
Enterobacteriaceae	14 (29.1)	8 (30.8)	6 (27.3)
Anaerobes	14 (29.1)	7 (26.9)	7 (31.8)
Pseudomonas aeruginosa	5 (10.4)	1 (3.8)	4 (18.2)
Streptococcus species	4 (8.3)	4 (15.4)	0 (0)
Negative culture‡	3 (6.5)	1 (3.8)	2 (9.1)

* Values are the number (%) unless indicated otherwise. Contributive examinations correspond to examinations in which imaging aspects of osteitis were reported. CRP = C-reactive protein; dSSc = diffuse cutaneous SSc; IQR = interquartile range; lSSc = limited cutaneous SSc; MRI = magnetic resonance imaging.
† No./total no. (% of performed examination).
‡ Sterile samples were reported in the entire SCLEROS population (n = 48).

Description of osteitis episodes. *Clinical presentation.* Osteitis characteristics are shown in Table 2. Osteitis in the SCLEROS population occurred predominantly in distal portions of the limbs (91.7% of patients), mainly hands (52.1%). Five (10.4%) patients had osteitis over the proximal phalangeal joints. Other locations included toes (n = 19), elbows (n = 2), femur (n = 1), and sternum (n = 1). Digital ulcers and calcinosis were reported overlying the osteitis location in 26 (54.1%) and 13 (27.1%) of cases, respectively. Physical symptoms included pain (75%), erythema (72.9%), and local warmth (72.9%). In most cases (44 of 48, 91.7 %), fever was absent. Purulent discharge was reported in 56.3% of patients.

The patient with sternum osteitis (a 56-year-old man) presented with a 6-month diagnostic history of diffuse cutaneous SSc and positivity for anti–topoisomerase I antibodies. He had history of calcinosis, interstitial lung disease, and gastroesophageal reflux disease. Osteitis was suspected on local pain associated with local warmth, functional disability, and erythema and confirmed on radiography. CRP level was 60 mg/liter. His condition was successfully managed with antibiotics, with no sequelae and no recurrence. The patient with femur osteitis (a 72-year-old woman) presented with a 4-year diagnostic history of limited cutaneous SSc and positivity for anticentromere antibodies. She had history of calcinosis and gastroesophageal reflux disease. Osteitis was suspected due to local pain on the amputation site on her femur associated with local warmth, functional disability, purulent discharge, and erythema and confirmed on magnetic resonance imaging (MRI). Her CRP level was 66 mg/liter. Her condition was successfully managed with antibiotics and surgical joint cleansing, with no sequelae. She presented with a new episode of osteitis at the same location 6 months after the first episode, and her condition was successfully managed with antibiotics.

When comparing osteitis beneath digital ulcers to other types of osteitis (i.e., those without underlying digital ulcer), osteitis beneath digital tip ulcers mostly resembled other types of osteitis (Tables 2 and 3), but osteitis beneath ulcers was more frequently associated with local swelling and erythema (73.1%; $P = 0.004$ and 84.6%; $P = 0.04$, respectively).

Laboratory investigations. Thirty-one (64.6%) patients had an increased CRP level with a median (IQR) CRP value of 8 (2.7–44.3) mg/liter (n < 5 mg/liter). Hyperleukocytosis (>8,000 leukocytes/mm^3) was present in 27.1% of patients. Other blood count parameters were within normal range. No difference was reported for biological documentation for osteitis beneath digital tip ulcers when compared to osteitis without digital tip ulcers.

Imaging manifestations. In the SCLEROS population, only 2 (4.2%) patients did not have any imaging performed. Standard radiography (n = 36 patients) showed osteolysis or acro-osteolysis in 17 patients, subcutaneous calcifications in 6 patients, and thickening of soft tissues in 5 patients. MRI (n = 17 patients) reported localized collection in 3 patients, soft-tissue swelling in 12 patients, thickening of soft tissues in 10 patients, loss of surrounding tissue in 2 patients, and acro-osteolysis in 5 patients. Computed tomography (CT) scan (n = 8 patients) showed cortical osteolysis with loss of surrounding tissue in 6 patients. Fifteen patients (31.3%) had results from either radiographs and MRI, or

Table 3. Management and prognosis of osteitis in patients with systemic sclerosis (SSc)*

Characteristic	SCLEROS cases (n = 48)	Patients with osteitis overlying digital tip ulcers (n = 26)	Patients with osteitis without digital tip ulcers (n = 22)
Management of osteitis			
Multidisciplinary management	24 (50.0)	11 (42.3)	13 (59.1)
Surgical treatment	26 (54.2)	16 (61.5)	10 (45.5)
Antibiotics	37 (77.1)	21 (80.8)	16 (72.7)
Local dressing	35 (72.9)	21 (80.8)	14 (63.6)
Surgical joints cleansing	8 (16.7)	4 (15.4)	4 (18.2)
Opiates	10 (20.8)	7 (26.9)	3 (13.6)
Amputation	11 (22.9)	8 (30.8)	3 (13.6)
Synovectomy	2 (4.2)	1 (3.9)	1 (4.6)
Re-evaluation			
Follow-up, median (IQR) months	5 (2–18)	3.8 (2–12)	8 (2.5–24.0)
Clinical cure	35 (72.9)	19 (73.1)	16 (72.7)
Relapse	6 (12.6)	4 (15.4)	2 (9.1)
No. of relapses per patient, median (IQR)	1 (0–1)	1 (0–1)	1 (0–1)
No. of homolateral relapse	4	3	1
Time between 2 episodes, median (IQR) months	2 (1–4)	2 (1–4)	3 (2.5–6.0)
New episode of osteitis	8 (16.6)	7 (26.9)	1 (4.5)
Time to second episode, median (IQR) months	6 (5–9)	5 (3–10)	6 (5–7)
Sequelae	15 (31.3)	10 (38.5)	5 (22.7)
Death related to osteitis	2 (4.2)	2 (7.7)	0 (0)

* Values are the number (%) unless indicated otherwise. IQR = interquartile range.

Table 4. Antibiotic therapy management in systemic sclerosis osteitis*

Characteristic	SCLEROS cases (n = 48)
Duration of treatment, days	41 (5.5–55)
Bitherapy, no. (%)	23 (47.9)
Duration of bitherapy, days	41 (30–49)
Duration of intravenous administration, days	9 (0–56)
Duration of oral administration, days	29 (0–51.5)
Duration of switch from intravenous to oral, days	1 (0–40)

* Values are the median (interquartile range) unless indicated otherwise.

radiographs and CT scan, or MRI and CT scan. On radiography, CT, or MRI, osteitis was characterized by swelling or abscess of soft tissues, with acro-osteolysis or lysis in 28 patients (58.3%). Additional description of osteitis included localized collection and swelling of soft tissues without acro-osteolysis, which were present in only 6.3% of radiologic reports.

Microbiological documentation. Of 48 patients, 40 (83.3%) patients had undergone microbiological sampling. Sampling was performed either superficially or deeply. Seventy-seven percent of superficial samples and 88.9% of deep samples allowed the identification of at least 1 pathogen species. Of 40 patients with samples, 3 had sterile samples (7.5%), including 2 surgical samples and 1 needle-guided sample. Of 37 samples with identified microorganisms, 18 samples (48.6%) had >1 identified microorganism. Overall, the most frequent pathogens were *Staphylococcus aureus* (43.8%), anaerobes and Enterobacteriaceae (29.1%), and *Pseudomonas aeruginosa* (10.4%). No difference was found in terms of pathogen occurrence considering the presence of digital tip ulcers.

Osteitis management. Management of osteitis involved several specialists (internists, infectious diseases specialists, and orthopedic surgeons) in 50% of cases. Management consisted of antibiotics in 77.1% of patients and/or surgery in 54.2% of patients (synovectomy [4.2%] or amputation [22.9%]) (Table 3).

The main antibiotic regimens are summarized in Tables 4 and 5. Fluoroquinolones were used in 45.8% of patients, macrolides or lincosamides in 18.8%, and amoxicillin and β-lactamase inhibitor in 14.6%. First-line antibiotics were fluoroquinolone, orally as a bitherapy for a median of 34 days (IQR 10–50 days), or amoxicillin ± β-lactamase inhibitor orally as a bitherapy for a median of 30 days (IQR 12–42 days). Second-line antibiotics were fluoroquinolone, orally for a median of 32 days (IQR 27–42 days), or rifampin, orally for a median of 42 days (IQR 35–67 days) as a bitherapy. Detailed regimens are presented in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24530/abstract.

After a median 5 months (IQR 2–18 months) of follow-up, healing of osteitis was reported in 72.9% of patients. Relapse was observed in 6 patients (12.6%; 1 relapse per patient) within 2 months of the diagnosis of the first episode. Six (12.6%) patients experienced a new episode of osteitis. Notably, the location of such osteitis recurrence was contralateral to the first episode in 5 of 6 patients. The time before recurrence was a median 4 months (IQR 2–6 months). For 1 patient, the same antibiotic was continued for 3 months (amoxicillin plus clavulanic acid). For another, the antibiotic was switched to piperacillin/tazobactam, and then cefazolin for 3 months. For the other 4, an amputation of the affected limb was required.

Thirty-two percent of patients had functional sequelae after management, including those related to amputation. Two patients died of septic shock leading to multiple organ failure secondary to osteitis. When comparing osteitis beneath digital ulcers to the locations of other types of osteitis, management and prognosis were similar.

Osteitis management in patients with sterile samples. Eleven patients were treated without any microbiological documentation. Three patients had sterile samples. The first patient's condition was managed with an oral bitherapy (levofloxacin plus clindamycin) for 6 weeks, with a clinical healing of osteitis and no recurrence at 2 months. The second patient's condition was managed with an oral monotherapy (cloxacillin) for 20 days, and then an oral monotherapy (clindamycin) for 6 weeks, with a clinical healing of osteitis and no recurrence at 2 months after the beginning of treatment. The condition of the third patient was managed

Table 5. Main antibiotics regimens in systemic sclerosis osteitis*

	Antibiotics	Count (n = 48)	Duration, median (IQR) days	Oral/IV, no.	Monotherapy/ bitherapy, no.
First-line					
First	Fluoroquinolone	19	34 (10–50)	15/4	4/15
Second	Amoxicillin ± β-lactamase inhibitor	15	30 (12–42)	14/1	4/11
Third	Macrolides	7	35 (15–42)	7/0	1/6
Second-line					
First	Fluoroquinolone	10	32 (27–42)	8/2	2/8
Second	Rifampin	4	42 (35–67)	3/1	0/4
Third	Penicillin ± inhibitor	3	5 (3–17)	1/2	3/0

* IQR = interquartile range; IV = intravenous.

Table 6. Factors associated with osteitis in patients with systemic sclerosis (SSc)*

	Univariate analysis		Multivariable analysis	
	OR (95% CI)	*P*	OR (95% CI)	*P*
Age	1.01 (0.99–1.03)	0.19	–	–
Female sex	0.62 (0.29–1.36)	0.23	–	–
Body mass index	0.94 (0.88–1.02)	0.12	0.96 (0.87–1.07)	0.51
Active smoking	1.14 (0.58–2.26)	0.70	–	–
lSSc	0.48 (0.24–0.95)	0.04	0.76 (0.17–3.35)	0.72
Duration of SSc	1.01 (1.01–1.02)	0.02	1.01 (0.99–1.01)	0.44
Puffy fingers	1.39 (0.58–3.37)	0.46	–	–
Sclerodactyly	1.53 (0.71–3.27)	0.28	–	–
MRSS	1.05 (1.01–1.09)	0.02	1.02 (0.94–1.10)	0.65
Raynaud's phenomenon	0.49 (0.09–2.75)	0.42	–	–
Abnormal nailfold capillaries	0.79 (0.09–3.19)	0.74	–	–
Digital tip ulcers	2.29 (1.18–4.45)	0.01	4.26 (1.43–12.66)	0.009
Telangiectasia	0.79 (0.41–1.51)	0.49	–	–
Calcinosis	1.49 (0.89–3.59)	0.10	3.13 (1.16–8.47)	0.03
Gastroesophageal reflux disease	1.38 (0.71–2.68)	0.35	–	–
Interstitial lung disease	0.48 (0.24–0.94)	0.03	0.62 (0.22–1.73)	0.36
Pulmonary arterial hypertension	0.50 (0.16–1.54)	0.23	–	–
Musculoskeletal disorders	1.35 (0.70–2.61)	0.37	–	–
Scleroderma renal crisis	0.55 (0.18–1.69)	0.29	–	–
Anticentromere autoantibodies	1.26 (0.98–2.17)	0.08	1.17 (0.85–3.01)	0.41
Anti–topoisomerase I autoantibodies	0.98 (0.62–1.26)	0.31	–	–
Anti–RNA polymerase III autoantibodies	0.83 (0.22–1.39)	0.55	–	–
Immunosuppressant without glucocorticoids	4.20 (0.82–2.50)	0.09	2.86 (0.27–30.74)	0.39
Glucocorticoids	0.66 (0.34–1.29)	0.23	–	–

* 95% CI = 95% confidence interval; lSSc = limited cutaneous SSc; MRSS = modified Rodnan skin thickness score; OR = odds ratio.

surgically. For 8 patients, no microbiological sampling was performed. All these patients were treated with antibiotics for 10 days to 6 weeks (penicillin ± inhibitors in 2 patients, fluoroquinolone in 4 patients, and rifampin in 2 patients). Additional local wounds were cared for with dressings until healing. Among the 8 patients, 2 required surgical management (amputation for both patients) due to a lack of clinical improvement. Six patients were healed at 6 months, and 2 had functional sequelae because of the amputation.

Associated factors for osteitis and osteitis healing. Associated factors for the occurrence of osteitis are presented in Table 6, with values shown as odds ratios (ORs) and 95% confidence intervals (95% CIs). Among the variables tested, digital tip ulcers (OR 4.26 [95% CI 1.43–12.66], $P = 0.009$) and calcinosis (OR 3.13 [95% CI 1.16–8.47], $P = 0.03$) were found as independent associated factors for osteitis.

Associated factors for the occurrence of osteitis on digital tip ulcers are presented in Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24530/abstract. Among all the tested variables, age (OR 1.06 [95% CI 1.01–1.12], $P = 0.03$) was found to be an independent associated factor for osteitis on digital tip ulcers. Among all the variables tested to identify factors with osteitis healing, none was found to be an independent associated factor (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24530/abstract).

DISCUSSION

This daily life study confirms that the presence of digital ulcers is an associated factor for osteitis in SSc patients. This is in accordance with the findings of Giuggioli et al in 2013, who showed that osteoarticular infections such as osteomyelitis were common in patients with SSc and skin ulcers (18). Moreover, digital ulcers predicted worsening of SSc and were complicated by osteitis in 5–23% of cases (7,19–22). In their 2016 prospective, observational cohort, Allanore et al sought the clinical features related to the occurrence of gangrene in patients with SSc (10). The authors found that digital ulcers, as well as smoking status, anti–Scl-70 antibodies, and a history of previous infection (gangrene, osteomyelitis) were associated with increased bone infections (10). The SCLEROS study confirmed that digital ulcers are associated with osteitis and highlighted calcinosis as another associated factor for osteitis.

Our work has highlighted the importance of microbiological documentation in the diagnosis and management of osteitis. In osteitis, the contribution of bone biopsy in the patient's diagnosis is unequivocal (23). Nevertheless, due to the underlying microangiopathy in SSc, physicians are often reluctant to perform surgical microbiological sampling in SSc patients. Our work demonstrated that both superficial sampling and bone biopsy can be performed safely in SSc patients and led to microbiological identification in 37 cases (77.1% of the entire cohort). Imaging is often proposed as an alternative to performing bone biopsy (24,25). As reported

by Giuggioli et al (18), MRI, CT, and radiography findings are already used in daily practice for the diagnosis of osteomyelitis, mostly in the context of diabetes mellitus (26–30). The SCLEROS study confirms that such examination can identify swelling or abscess of soft tissues with acro-osteolysis or localized collection and soft-tissue swelling without acro-osteolysis to confirm osteitis in SSc patients. Still, microbiological sampling remains necessary for the management of osteitis, as acro-osteolysis or bone inflammation due to calcinosis can be found without infection in SSc patients. As for any other infections, identification of the pathogens is relevant to avoid overtreatment of patients and to reduce the acquired antibiotic resistance (31). In the SCLEROS study, the identification of pathogens allowed reduction of the antibiotics' spectrum and decrease of the mean duration of antibiotic therapy (34 days versus 6 weeks in patients with sterile samples). In this study, *S aureus*, anaerobes and Enterobacteriaceae, and *P aeruginosa* were the most common pathogens. This is consistent with the work of Giuggioli et al and Bader, who reported in 2008 a frequent association with gram-positive bacteria in diabetic foot infection, another infection of the distal limb enhanced by microangiopathy (18,32). Still, the presence of 29% of bacterial floral culture was unexpected and could hypothetically be linked to specific microbiota, inefficient hand washing, or multiple use of antibiotics in SSc patients.

Treating osteitis is challenging, as the objectives of its management are to decrease the infection, to reconstruct bone structure and surrounding soft tissue, to preserve the affected joints, and from the patient's point of view, to return to a normal life as quickly as possible with minimum functional impairment (23). The SCLEROS study showed that treating osteitis in SSc patients is also extremely challenging. Although most patients had a favorable outcome, the study showed that 31.3% of SSc patients with osteitis had functional sequelae. Two patients died. Interestingly, the SCLEROS study suggested that after performing microbiological sampling, probabilistic antibiotic with fluoroquinolone or amoxicillin and β-lactamase inhibitor may be relevant. As it is efficient on *Staphylococcus* species, *Streptococcus* species, enterobacteria, and pyocyanin, such antibiotic therapy is already recommended in diabetic foot infections (33). In the SCLEROS study, patients with microbiological identification had pathogens for which such antibiotic therapy was or would have been efficient. Moreover, the conditions of 6 patients without microbiological identification were managed with such probabilistic antibiotic therapy and had a favorable outcome.

The SCLEROS study has some limitations. First, considering the methodology (i.e., a retrospective study), memory bias could have been present. The investigators involved may have reported difficult or severe clinical situations. This possible selection of severe cases might be an explanation for the reported death and sequelae. Moreover, patients with osteitis lacking any clinical symptoms might have been missed in this study. Indeed, some patients with osteomyelitis can be diagnosed with osteitis only on the basis of CT scan or MRI results. Second, also due to the retrospective methodology, radiologic examinations were not reanalyzed by a centralized radiologist for the purpose of the study, and a prospective study on radiologic examination in SSc osteitis should be performed. Third, possible differences regarding the modalities of local care and surgical cleansing may have had an impact on the osteitis recovery results. Nevertheless, it is difficult to standardize these treatments, as they depend on the local presentation of osteitis and physicians' practices. Last, no assessment of disability and quality of life due to osteitis in SSc was performed. Therefore, we were not able to describe the functional consequences of osteitis in SSc. Despite these weaknesses, this study is the first to describe osteitis in patients with SSc. Moreover, as this study was not tailored for prevalence calculation, the chosen design was appropriate to answer the research question (i.e., the description of osteitis in SSc patients). Besides, the data presented in SCLEROS reflected daily practice and stood for a homogeneous osteitis cohort.

In conclusion, this study characterized osteitis in SSc patients. It confirmed digital tip ulcers as an associated factor for osteitis and revealed a high rate of functional sequelae. It showed that antimicrobial bitherapy with either oral fluoroquinolone or intravenous amoxicillin and β-lactamase inhibitor are used as first-line antibiotic therapy in SSc patients with osteitis.

ACKNOWLEDGMENTS

The authors thank the members of the Groupe Francophone de Recherche sur la Sclérodermie, the Société Nationale Française de Médecine Interne, the Société de Pathologie Infectieuse de Langue Française, and the Club Rhumatismes et Inflammations for their participation in the study.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Chaigne 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. Cosse, Kernéis, Cerles, Le Jeunne, Mouthon, Chaigne.

Acquisition of data. Cosse, Lescoat, Pugnet, Truchetet, Priollet, Diot, Martin, Maurier, Viallard, Agard, Granel, Berthier, Fagedet, Watelet, Toquet, Paz, Giret, Dion, Nguyen, Raffray, Bertolino, Jourde, Le Jeunne, Mouthon, Chaigne.

Analysis and interpretation of data. Cosse, Le Jeunne, Mouthon, Chaigne.

REFERENCES

1. Cutolo M, Soldano S, Smith V. Pathophysiology of systemic sclerosis: current understanding and new insights. Expert Rev Clin Immunol 2019;15:753–64.
2. Dumoitier N, Lofek S, Mouthon L. Pathophysiology of systemic sclerosis: state of the art in 2014. Presse Med 2014;43:e267–78.
3. Denton CP, Khanna D. Systemic sclerosis. Lancet 2017;390: 1685–99.

4. Botou A, Bangeas A, Alexiou I, Sakkas LI. Acro-osteolysis. Clin Rheumatol 2017;36:9–14.
5. Sakkas LI, Bogdanos DP. Infections as a cause of autoimmune rheumatic diseases. Auto Immun Highlights 2016;7:13.
6. Sharp CA, Akram Q, Hughes M, Muir L, Herrick AL. Differential diagnosis of critical digital ischemia in systemic sclerosis: report of five cases and review of the literature. Semin Arthritis Rheum 2016;46: 209–16.
7. Becker M, Graf N, Sauter R, Allanore Y, Curram J, Denton CP, et al. Predictors of disease worsening defined by progression of organ damage in diffuse systemic sclerosis: a European Scleroderma Trials and Research (EUSTAR) analysis. Ann Rheum Dis 2019;78:1242–8.
8. Ea HZ, Zeller V, Lhotellier L, Ziza JM, Mamoudy P, Desplaces N. Chronic osteitis in adults: diagnosis and treatment. Antibiotiques 2007;9:120.
9. Norden CW, Budinsky A. Treatment of experimental chronic osteomyelitis due to Staphylococcus aureus with ampicillin/sulbactam. J Infect Dis 1990;161:52.
10. Allanore Y, Denton CP, Krieg T, Cornelisse P, Rosenberg D, Schwierin B, et al. Clinical characteristics and predictors of gangrene in patients with systemic sclerosis and digital ulcers in the Digital Ulcer Outcome Registry: a prospective, observational cohort. Ann Rheum Dis 2016;75:1736–40.
11. Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
12. Mader JT, Mohan D, Calhoun J. A practical guide to the diagnosis and management of bone and joint infections. Drugs 1997;54:253–64.
13. Berendt AR, Peters EJ, Bakker K, Embil JM, Eneroth M, Hinchliffe RJ, et al. Diabetic foot osteomyelitis: a progress report on diagnosis and a systematic review of treatment. Diabetes Metab Res Rev 2008;24 Suppl 1:S145–61.
14. Hamajima N, Hirose K, Inoue M, Takezaki T, Kuroishi T, Tajima K. Case-control studies: matched controls or all available controls? J Clin Epidemiol 1994;47:971–5.
15. Groenwold RH, van Smeden M. Efficient sampling in unmatched case-control studies when the total number of cases and controls is fixed. Epidemiology 2017;28:834–7.
16. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15:202–5.
17. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Int J Surg 2014;12:1495–9.
18. Giuggioli D, Manfredi A, Colaci M, Lumetti F, Ferri C. Osteomyelitis complicating scleroderma digital ulcers. Clin Rheumatol 2013;32: 623–7.
19. Mihai C, Landewe R, van der Heijde D, Walker UA, Constantin PI, Gherghe AM, et al. Digital ulcers predict a worse disease course in patients with systemic sclerosis. Ann Rheum Dis 2016;75:681–6.
20. Hachulla E, Clerson P, Launay D, Lambert M, Morell-Dubois S, Queyrel V, et al. Natural history of ischemic digital ulcers in systemic sclerosis: single-center retrospective longitudinal study. J Rheumatol 2007;34:2423–30.
21. Bogoch ER, Gross DK. Surgery of the hand in patients with systemic sclerosis: outcomes and considerations. J Rheumatol 2005;32:642–8.
22. Blagojevic J, Abignano G, Avouac J, Cometi L, Frerix M, Bellando-Randone S, et al. Use of vasoactive/vasodilating drugs for systemic sclerosis (SSc)-related digital ulcers (DUs) in expert tertiary centres: results from the analysis of the observational real-life DeSScipher study. Clin Rheumatol 2020;39:27–36.
23. Tiemann AH, Hofmann GO, Steen M, Schmidt R. Adult calcaneal osteitis: incidence, etiology, diagnostics and therapy. GMS Interdiscip Plast Reconstr Surg DGPW 2012;1:Doc11.
24. Zhou AY, Muir L, Harris J, Herrick AL. The impact of magnetic resonance imaging in early diagnosis of hand osteomyelitis in patients with systemic sclerosis. Clin Exp Rheumatol 2014;32 Suppl 86:S–232.
25. McQueen FM. Bone marrow edema and osteitis in rheumatoid arthritis: the imaging perspective. Arthritis Res Ther 2012;14:224.
26. Lauri C, Tamminga M, Glaudemans A, Juarez Orozco LE, Erba PA, Jutte PC, et al. Detection of osteomyelitis in the diabetic foot by imaging techniques: a systematic review and meta-analysis comparing MRI, white blood cell scintigraphy, and FDG-PET. Diabetes Care 2017;40:1111–20.
27. Dinh T, Snyder G, Veves A. Current techniques to detect foot infection in the diabetic patient. Int J Low Extrem Wounds 2010;9:24–30.
28. Lew DP, Waldvogel FA. Osteomyelitis. Lancet 2004;364:369–79.
29. Romano CL, Romano D, Logoluso N, Drago L. Bone and joint infections in adults: a comprehensive classification proposal. Eur Orthop Traumatol 2011;1:207–17.
30. Gentry LO. Management of osteomyelitis. Int J Antimicrob Agents 1997;9:37–42.
31. Maleb A, Frikh M, Lahlou YB, Chagar B, Lemnouer A, Elouennass M. Bacteriological aspects of chronic osteoarticular infections in adults: the influence of the osteosynthesis material. BMC Res Notes 2017; 10:635.
32. Bader MS. Diabetic foot infection. Am Fam Physician 2008;78:71–9.
33. Diamantopoulos EJ, Haritos D, Yfandi G, Grigoriadou M, Margariti G, Paniara O, et al. Management and outcome of severe diabetic foot infections. Exp Clin Endocrinol Diabetes 1998;106:346–52.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 818–827
DOI 10.1002/acr.24521

American College of Rheumatology
Empowering Rheumatology Professionals

Work Productivity and Economic Burden of Systemic Sclerosis in a Multiethnic Asian Population

Ling Xiang,[1] Sandra M. Y. Kua,[2] and Andrea H. L. Low[3]

Objective. To assess work productivity, identify associated factors and evaluate the economic burden of systemic sclerosis (SSc) in a multiethnic Asian population.

Methods. Data on employment status and work productivity loss were collected. Associations between demographic and disease characteristics and unemployment status, work productivity loss, and activity impairment were examined using logistic and linear regression analyses, as appropriate. Costs of unemployment and work productivity loss were estimated using the human capital approach.

Results. Of 111 patients with a mean disease duration of 9.1 years, 33 (29.7%) were unemployed. Their mean age at unemployment was 44.2 years, equating to 22.8 years of lost employment. No demographic and disease characteristics were significantly associated with unemployment status in multivariable analysis. Of 73 employed patients, 39 (53.4%) reported work productivity loss, accounting for 45.9% of the working week. The presence of hyperlipidemia (coefficient −19.01, $P = 0.03$) was associated with work productivity loss in multivariable analysis. In total, 37 of 78 employed patients (47.4%) and 19 of 33 unemployed patients (57.6%) reported activity impairment, accounting for 42.2% and 50.0%, respectively, of the preceding week. The presence of hyperlipidemia (coefficient −18.56, $P < 0.01$) was associated with activity impairment in multivariable analysis. Annual cost of unemployment and work productivity loss were estimated to be $53,244 and $13,045 (Singapore dollar) per patient, respectively.

Conclusion. SSc imposes significant unemployment and work productivity loss and causes a substantial economic burden to both affected individuals and society. Modifying the identified factors associated with unemployment and work productivity loss may reduce the burden of SSc.

INTRODUCTION

Systemic sclerosis (SSc) is a less common autoimmune rheumatic disease characterized by excessive production and accumulation of collagen in the skin and other tissues such as the joints, tendons, blood vessels, muscles, gastrointestinal tract, lungs, kidneys, and heart. SSc significantly impacts physical function and leads to a reduced capacity to participate in activities of daily life such as self-care/maintenance, household chores, and professional activities such as work and studies (1–5). Affected individuals in the workforce may experience work disability, namely reduced work productivity (presenteeism), absence from work (absenteeism), change of jobs/careers (work transitions), and work cessation (unemployment/early retirement) due to illness before reaching retirement age (6–9). Work disability has been reported among 18–89% of patients with SSc in published studies (10–12). Since SSc affects mostly patients of an economically productive age (10), it can lead to a substantial burden to both affected individuals and society (11,13,14).

The economic burden of disease has been mainly studied in the more common autoimmune rheumatic diseases, including rheumatoid arthritis (RA), axial spondyloarthritis, and systemic lupus erythematosus (SLE) (15–17). Studies on SSc have been mostly conducted in Europe, North America, and Australia (11,13,14,18–22). There is a paucity of reports from Asian countries, where the clinical manifestations of SSc have been shown to differ from those in other populations (23). The primary aim of this study was to assess work productivity and identify associated factors in patients with SSc in a multiethnic Asian population. The secondary aim was to evaluate the

[1]Ling Xiang, MBBS, MMed: Singapore General Hospital and Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [2]Sandra M. Y. Kua, BSc: Singapore General Hospital, Singapore; [3]Andrea H. L. Low, BMedSci, BMBS, MRCP, FAMS: Singapore General Hospital and Duke-NUS Medical School, Singapore.

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

Address correspondence to Andrea H. L. Low, BMedSci, BMBS, MRCP, FAMS, Department of Rheumatology and Immunology, Singapore General Hospital, Academia Building, Level 4, 20 College Road, Singapore, 169856. Email: andrea.low.h.l@singhealth.com.sg.

Submitted for publication April 19, 2020; accepted in revised form November 24, 2020.

SIGNIFICANCE & INNOVATIONS

- Nearly 30% of the patients with systemic sclerosis (SSc) were unemployed at a mean age of 44 years.
- Considerable work productivity loss was reported in over half of employed patients.
- SSc imposes a substantial economic burden to both affected individuals and society.
- Modifying identified factors for work productivity may reduce the burden of SSc.

economic burden of unemployment and work productivity loss in these patients.

PATIENTS AND METHODS

Patients. Patients who were enrolled in the Systemic Sclerosis Cohort in Singapore (SCORE) (24) and who were seen at the scleroderma clinic in Singapore General Hospital were consecutively recruited between October 2017 and January 2020. All patients fulfilled the 2013 American College of Rheumatology/European Alliance of Associations for Rheumatology classification criteria for SSc (25) or the Very Early Diagnosis of Systemic Sclerosis criteria (26), and were age ≥18 years. Written consent was obtained from all patients. This study was approved by the SingHealth Centralized Institutional Review Board (ref. 2007/011/E). Data on employment status, work productivity loss, and demographic and disease characteristics were collected from all patients.

Outcome measures. All patients were administered a standardized questionnaire during their clinic visit, with the following options for employment status: employed, unemployed due to medical condition, unemployed due to nonmedical condition, full-time student, homemaker, national service (compulsory service in Singapore), and retired. Unemployment was defined as being unemployed due to a medical condition or being unemployed due to a nonmedical condition in this study. Patients who were a full-time student, a homemaker, on national service, or retired were excluded from analysis.

Work productivity was measured using the Work Productivity and Activity Impairment (WPAI) specific health problem (SSc) questionnaire (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24521/abstract) (27). There are 6 items on the WPAI concerning employment status (item 1), hours missed from work due to SSc (item 2) and other reasons (item 3), total hours worked (item 4), and the impact of SSc on work productivity (item 5) and daily activities (item 6) during the preceding week. For employed patients, we calculated absenteeism (time absent from work), presenteeism (time at work with reduced productivity), and work productivity loss (overall work impairment calculated from absenteeism and presenteeism). For both employed and unemployed patients, we calculated activity impairment (activity limitation outside work for the employed). All these 4 scores from the WPAI were expressed as impairment percentages (possible range 0–100), with a higher score indicating greater impairment.

Independent variables. Demographic characteristics included age, sex, ethnicity, date of first non–Raynaud's phenomenon manifestation, date of WPAI survey, and smoking history. Disease duration at the time of the survey was derived from the date of the first non–Raynaud's phenomenon manifestation. Physician-diagnosed comorbidities, including hypertension, ischemic heart disease, hyperlipidemia, diabetes mellitus, renal impairment, stroke, and malignancy were extracted from the medical records.

Disease characteristics included SSc subtype, cumulative clinical manifestations, skin thickness score, pulmonary function, autoantibodies, and treatment received. Patients were classified as having limited cutaneous SSc (lcSSc), diffuse cutaneous SSc (dcSSc) according to the criteria established by LeRoy et al (28), or SSc overlap if SSc occurred with 1 or more of the following defined connective tissue diseases: SLE, RA, or inflammatory myositis.

Cumulative clinical manifestations captured in the SCORE cohort were described elsewhere (29) and included joint involvement (arthritis or joint contracture), tendon friction rub, mild peripheral vasculopathy (Raynaud's phenomenon, digital pitting, pulp atrophy, or telangiectasia), severe peripheral vasculopathy (digital ulcers or gangrene), calcinosis, inflammatory myositis, upper gastrointestinal involvement (reflux/dysphagia, vomiting, gastric antral vascular ectasia, or bloating/distension), lower gastrointestinal involvement (diarrhea, constipation, fecal soilage, or malabsorption), interstitial lung disease (ILD; diagnosed by high-resolution computed tomography) and pulmonary arterial hypertension (PAH; diagnosed by right heart catheterization), renal involvement (renal crisis, creatinine above the upper limit of normal on 3 or more occasions, and proteinuria) and cardiac involvement.

Skin thickness was measured using the maximum modified Rodnan skin score (30). Pulmonary function was quantified using the forced vital capacity (FVC). Treatment received included immunosuppression (methotrexate, azathioprine, cyclophosphamide, and mycophenolate mofetil) and treatment for peripheral vasculopathy and PAH (prostacyclin, phosphodiesterase type 5 inhibitor, and endothelin receptor antagonists).

Estimation of economic burden. In Singapore, employers are required to provide re-employment to employees who turn 62 (retirement age), until age 67 years (re-employment age) (31). In an epidemiologic survey conducted among the Singapore general population in 2017, only 1.8% of the residents

Table 1. Demographic and disease characteristics of SSc patients (n = 111)*

	Overall (n = 111)	Employed (n = 78)	Unemployed (n = 33)	*P*
Female sex	97 (87.4)	67 (85.9)	30 (90.9)	0.55
Age at survey, mean ± SD years	49.4 ± 10.5	48.2 ± 10.7	52.2 ± 9.5	0.07
Age at unemployment, mean ± SD years	–	–	44.2 ± 9.3	–
Disease duration from first non-RP symptom onset to survey, mean ± SD years	9.1 ± 6.7	8.8 ± 6.8	9.7 ± 6.8	0.52
Ethnicity				0.91
Chinese	78 (70.3)	55 (70.5)	23 (69.7)	–
Malay	13 (11.7)	10 (12.8)	3 (9.1)	–
Indian	6 (5.4)	4 (5.1)	2 (6.1)	–
Others	14 (12.6)	9 (11.5)	5 (15.2)	–
Smoking history (previous or current smoker)	10 (9.0)	7 (9.0)	3 (9.1)	>0.99
Comorbidities				
Hypertension	19 (17.1)	11 (14.1)	8 (24.2)	0.20
Ischemic heart disease	3 (2.7)	2 (2.6)	1 (3.0)	>0.99
Hyperlipidemia	21 (18.9)	16 (20.5)	5 (15.2)	0.51
Diabetes mellitus	4 (3.6)	4 (5.1)	0 (0.0)	0.32
Renal impairment	4 (3.6)	2 (2.6)	2 (6.1)	0.58
Stroke	2 (1.8)	2 (2.6)	0 (0.0)	>0.99
Malignancy	4 (3.6)	3 (3.9)	1 (3.0)	>0.99
Cumulative clinical manifestations				
SSc subtype				0.07
lcSSc	48 (43.2)	37 (47.4)	11 (33.3)	–
dcSSc	38 (34.2)	22 (28.2)	16 (48.5)	–
SSc overlap	24 (21.6)	19 (24.4)	5 (15.2)	–
Joint involvement	83 (74.8)	61 (78.2)	22 (66.7)	0.20
Tendon friction rub	5 (4.5)	3 (3.9)	2 (6.1)	0.83
Mild peripheral vasculopathy	105 (94.6)	74 (94.9)	31 (93.9)	>0.99
RP	89 (80.2)	64 (82.1)	25 (75.8)	0.45
Digital pitting	56 (50.5)	37 (47.4)	19 (57.6)	0.33
Pulp atrophy	75 (67.6)	51 (65.4)	24 (72.7)	0.45
Telangiectasia	64 (57.7)	42 (53.9)	22 (66.7)	0.21
Severe peripheral vasculopathy	18 (16.2)	8 (10.3)	10 (30.3)	0.01
Digital ulcers	17 (15.3)	8 (10.3)	9 (27.3)	0.02
Digital gangrene	6 (5.4)	1 (1.3)	5 (15.2)	0.01
Calcinosis	12 (10.8)	7 (9.0)	5 (15.2)	0.34
Inflammatory myositis	18 (16.2)	12 (15.4)	6 (18.2)	0.28
Gastrointestinal involvement	90 (81.1)	62 (79.5)	28 (84.9)	0.51
Upper	78 (70.3)	53 (68.0)	25 (75.8)	0.41
Lower	63 (56.8)	43 (55.1)	20 (60.6)	0.59
ILD	46 (41.4)	32 (41.0)	14 (42.4)	0.14
PAH	8 (7.2)	4 (5.1)	4 (12.1)	0.41
Renal involvement	15 (13.5)	7 (9.0)	8 (24.2)	0.03
Cardiac involvement	3 (2.7)	0 (0.0)	3 (9.1)	0.03
Maximum MRSS, median (IQR)	6.0 (2.0–15.0)	6.0 (2.0–12.0)	8.0 (3.0–21.0)	0.12
FVC, % predicted, median (IQR)	72.0 (61.0–79.0)	74.0 (62.0–81.0)	65.5 (54.0–76.0)	0.02
Anti–Scl-70	43 (38.7)	32 (41.0)	11 (33.3)	0.71
Anticentromere	16 (14.4)	11 (14.1)	5 (15.2)	0.74
Immunosuppressive treatment	83 (74.8)	60 (76.9)	23 (69.7)	0.42
Methotrexate	44 (39.6)	32 (41.0)	12 (36.4)	0.65
Azathioprine	10 (9.0)	8 (10.3)	2 (6.1)	0.72
Cyclophosphamide	26 (23.4)	17 (21.8)	9 (27.3)	0.53
Mycophenolate mofetil	44 (39.6)	32 (41.0)	12 (36.4)	0.65
Medication for peripheral vasculopathy	10 (9.0)	7 (9.0)	3 (9.1)	>0.99
Prostacyclin	1 (0.9)	0 (0.0)	1 (3.0)	0.30
PDE5i	10 (9.0)	7 (9.0)	3 (9.1)	>0.99
ERA	1 (0.9)	1 (1.3)	0 (0.0)	>0.99
Medication for PAH	4 (3.6)	1 (1.3)	3 (9.1)	0.08
Prostacyclin	1 (0.9)	0 (0.0)	1 (3.0)	0.30
PDE5i	4 (3.6)	1 (1.3)	3 (9.1)	0.08
ERA	1 (0.9)	0 (0.0)	1 (3.0)	0.30

* Values are the number (%) unless indicated otherwise. ERA = endothelin receptor antagonist; dcSSc = diffuse cutaneous systemic sclerosis; FVC = forced vital capacity; ILD = interstitial lung disease; IQR = interquartile range; lcSSc = limited cutaneous SSc; MRSS = modified Rodnan skin score; PAH = pulmonary arterial hypertension; PDE5i = phosphodiesterase type 5 inhibitor; RP = Raynaud's phenomenon.

age 60–74 years became unemployed, 42.9% continued to work, 31.6%, were retired, and 23.6% were homemakers (32). We therefore chose 67 as the upper cutoff age for the estimation of economic burden. Only patients up to age 67 years were included in this study.

We used the human capital approach to estimate the economic burden (33). Since data on monthly income were not collected, we used the median gross monthly income in Singapore in 2018 (Singapore dollar [SGD] $4,437) instead (34). For patients who were unemployed at the time of the survey, we estimated the cost of unemployment from their age at unemployment until the re-employment age. For patients who were employed at the time of the survey, we estimated the costs of absenteeism, presenteeism, and work productivity loss. We also estimated the economic burden of unemployment and work productivity loss to Singapore based on an estimated prevalence of 78–100 per million population and a reported resident population of 4,026,200 in 2019 (35). The prevalence of SSc in Singapore was estimated based on the prevalence in other Asian countries and our own hospital data (36,37). All estimated costs were expressed in 2018 value without inflation or discounting.

Statistical analysis. Continuous variables such as age, disease duration, and WPAI scores are expressed as mean ± SD or median (interquartile range), while categorical variables such as sex, ethnicity, and disease characteristics are expressed as frequency (percentage). We assessed differences between employed and unemployed patients using student's *t*-test/Wilcoxon's rank sum test for continuous variables and chi-square test/Fisher's exact test for categorical variables, as appropriate. We examined the associations between independent variables (demographic and disease characteristics) and unemployment status (dichotomous variable: unemployed/employed) in all patients using logistic regression analysis. We examined the associations between these independent variables and 1) work productivity loss (continuous variable,

Table 2. Factors associated with unemployment status in SSc patients (n = 111)*

	Univariate analysis		Multivariable analysis	
	OR (95% CI)	*P*	OR (95% CI)	*P*
Female sex	1.64 (0.43, 6.32)	0.47	2.28 (0.40, 12.82)	0.35
Age at survey	1.04 (1.00, 1.09)	0.07	1.04 (0.91, 1.07)	0.11
Disease duration, first non-RP symptom onset to survey	1.02 (0.96, 1.08)	0.52	0.99 (0.91, 1.07)	0.77
Smoking history	1.00 (0.24, 4.13)	>0.99	–	–
Comorbidities				
Hypertension	1.95 (0.70, 5.41)	0.20	–	–
Ischemic heart disease	1.19 (0.10, 13.57)	0.89	–	–
Hyperlipidemia	0.69 (0.23, 2.08)	0.51	–	–
Diabetes mellitus†	–	–	–	–
Renal impairment	2.45 (0.33, 18.19)	0.38	–	–
Stroke†	–	–	–	–
Malignancy	0.78 (0.08, 7.80)	0.83	–	–
SSc subtype (ref. lcSSc)				
dcSSc	2.45 (0.96, 6.21)	0.06	–	–
SSc overlap	0.89 (0.27, 2.92)	0.84	–	–
Joint involvement	0.56 (0.23, 1.37)	0.20	–	–
Tendon friction rub	1.63 (0.26, 10.31)	0.60	–	–
Mild peripheral vasculopathy	0.84 (0.15, 4.81)	0.84	–	–
Severe peripheral vasculopathy	3.80 (1.34, 10.79)	0.01	3.07 (0.79, 11.96)	0.11
Calcinosis	1.81 (0.53, 6.19)	0.34	–	–
Inflammatory myositis	1.27 (0.43, 3.74)	0.67	–	–
Gastrointestinal involvement	1.45 (0.48, 4.34)	0.51	–	–
ILD	0.64 (0.25, 1.64)	0.35	–	–
PAH	2.46 (0.57, 10.59)	0.23	–	–
Renal involvement	3.25 (1.07, 9.87)	0.04	3.68 (0.98, 13.76)	0.05
Cardiac involvement†	–	–	–	–
Maximum MRSS	1.04 (1.00, 1.09)	0.05	1.01 (0.95, 1.07)	0.71
FVC, % predicted	0.97 (0.94, 0.99)	0.02	0.97 (0.93, 1.00)	0.06
Anti–Scl-70	0.70 (0.30, 1.67)	0.42	–	–
Anticentromere	1.18 (0.37, 3.78)	0.78	–	–
Immunosuppressive treatment	0.69 (0.28, 1.71)	0.42	–	–
Medication for peripheral vasculopathy (ref. none)	1.01 (0.25, 4.19)	0.98	–	–
Medication for PAH (ref. none)	7.70 (0.77, 76.97)	0.08	0.66 (0.03, 13.04)	0.79

* 95% CI = 95% confidence interval; dcSSc = diffuse cutaneous systemic sclerosis; FVC = forced vital capacity; ILD = interstitial lung disease; lcSSc = limited cutaneous SSc; MRSS = modified Rodnan skin score; OR = odds ratio; PAH = pulmonary arterial hypertension; RP = Raynaud's phenomenon.
† Odds ratio was not calculated as no/all patients with the condition were unemployed.

Table 3. Work productivity loss and activity impairment in SSc patients*

	Patients, no.	Work time %, mean ± SD	Work time %, median (IQR)
Absenteeism			
All employed patients	73	6.2 ± 19.8	0.0 (0.0–0.0)
Employed reporting absenteeism due to SSc	9	50.5 ± 31.6	34.8 (32.7–60.0)
Presenteeism			
All employed patients	78	19.7 ± 25.5	5.0 (0.0–30.0)
Employed reporting presenteeism due to SSc	39	39.5 ± 22.8	30.0 (20.0–60.0)
Overall work impairment (work productivity loss)			
All employed patients	73	24.5 ± 30.2	10.0 (0.0–50.0)
Employed reporting absenteeism or presenteeism due to SSc	39	45.9 ± 26.9	46.2 (20.0–67.4)
Activity impairment			
All employed and unemployed patients	111	22.6 ± 28.8	10.0 (0.0–40.0)
Employed and unemployed reporting activity impairment due to SSc	56	44.8 ± 25.4	40.0 (25.0–70.0)
All employed patients	78	20.0 ± 27.4	0.0 (0.0–30.0)
Employed reporting activity impairment due to SSc	37	42.2 ± 25.4	30.0 (20.0–70.0)
All unemployed patients	33	28.8 ± 31.5	20.0 (0.0–50.0)
Unemployed reporting activity impairment due to SSc	19	50.0 ± 25.4	50.0 (30.0–70.0)

* IQR = interquartile range; SSc = systemic sclerosis.

possible range 0–100) in employed patients, and 2) activity impairment (continuous variable, possible range 0–100) in all patients using linear regression analysis. We reported odds ratios (ORs) from logistic regression and coefficients from linear regression. Independent variables with a *P* value less than 0.1 in the univariate analysis were included in the corresponding multivariable analysis, with adjustments for age, sex, and disease duration. All analyses were performed using Stata software, version 15.0.

RESULTS

Demographic and disease characteristics. Among 195 patients with SSc who consented and completed the WPAI questionnaire, 54 patients who were above the age of 67 years (re-employment age), and 30 patients who were a full-time student, a homemaker, on national service, or retired were excluded from analyses (Table 1). The remaining 111 patients had a mean ± SD age of 49.4 ± 10.5 years and a mean ± SD disease duration of 9.1 ± 6.7 years at the time of the survey. The majority were female (87.4%) and of Chinese ethnicity (70.3%). Approximately 18.9% of patients had hyperlipidemia and 17.1% had hypertension. The other comorbidities were present in <5.0% of patients. Overall, lcSSc (43.2%) was the most common subtype of SSc, followed by dcSSc (34.2%) and SSc overlap (21.6%).

Unemployment. Of 111 patients, 33 (29.7%) were unemployed at the time of the survey; 23 were unemployed due to their medical condition and 10 were unemployed due to a nonmedical condition. Unemployed patients did not differ from employed patients in terms of demographic characteristics (Table 1).

Factors associated with unemployment status are shown in Table 2, with values shown as the odds ratios (ORs) and 95% confidence intervals (95% CIs). The presence of severe peripheral vasculopathy (OR 3.80 [95% CI 1.34, 10.79]; *P* = 0.01), renal involvement (OR 3.25 [95% CI 1.07, 9.87]; *P* = 0.04), and FVC (OR 0.97 [95% CI 0.94, 0.99]; *P* = 0.02) were significantly associated with unemployment status in univariate analysis. However, they were not significant in the multivariable analysis after adjustment for age, sex, and disease duration.

Work productivity loss. Of 73 employed patients with relevant data on the WPAI, 9 (12.3%) reported absenteeism, accounting for 50.5% of the working week (Table 3). Of 78 employed patients, 39 (50.0%) reported presenteeism, accounting for 39.5% of the working week. Of 73 employed patients, 39 (53.4%) reported work productivity loss (incorporating both absenteeism and presenteeism), accounting for 45.9% of the preceding week.

Factors associated with work productivity loss are shown in Table 4. Smoking history was associated with work productivity loss in univariate analysis (coefficient 26.30 [95% CI 2.79, 49.81]; *P* = 0.03). The presence of hyperlipidemia was associated with work productivity loss in both univariate analysis (coefficient –21.64 [95% CI –38.47, –4.81]; *P* = 0.01) and multivariable analysis (coefficient –19.01 [95% CI –36.39, –1.63]; *P* = 0.03) after adjustment for age, sex, and disease duration.

Activity impairment. Of 78 employed and 33 unemployed patients, 37 (47.4%) and 19 (57.6%), respectively, reported activity impairment, accounting for 42.2% and 50.0% of the preceding week (Table 3). Factors associated with activity impairment are shown in Table 5. The presence of hyperlipidemia (coefficient –17.90 [95% CI –31.39, –4.42]; *P* = 0.01), severe peripheral vasculopathy (coefficient 17.44 [95% CI 3.04, 31.84]; *P* = 0.02), and calcinosis (coefficient 22.30 [95% CI 5.28, 39.32]; *P* = 0.01) were associated with activity impairment in univariate analysis. After adjustment for age, sex, and disease duration, the

Table 4. Factors associated with work productivity loss in employed SSc patients (n = 78)*

	Univariate analysis		Multivariable analysis	
	Coefficient (95% CI)	*P*	Coefficient (95% CI)	*P*
Female sex	–13.35 (–32.96, 6.27)	0.18	9.27 (–16.10, 34.64)	0.47
Age at survey	0.05 (–0.61, 0.71)	0.88	–0.08 (–0.72, 0.56)	0.81
Disease duration, first non-RP symptom onset to survey	–0.07 (–1.16, 1.02)	0.90	–0.01 (–1.11, 1.10)	0.99
Smoking history	26.30 (2.79, 49.81)	0.03	27.45 (–1.35, 56.25)	0.06
Comorbidities				
Hypertension	–4.62 (–24.46, 15.22)	0.64	–	–
Ischemic heart disease	43.46 (–16.82, 103.74)	0.16	–	–
Hyperlipidemia	–21.64 (–38.47, –4.81)	0.01	–19.01 (–36.39, –1.63)	0.03
Diabetes mellitus	10.19 (–20.94, 41.33)	0.52	–	–
Renal impairment	–4.59 (–65.73, 56.55)	0.88	–	–
Stroke	–25.22 (–68.35, 17.91)	0.25	–	–
Malignancy	–18.63 (–54.16, 16.90)	0.30	–	–
SSc subtype (ref. lcSSc)				
dcSSc	–4.19 (–21.14, 12.75)	0.62	–	–
SSc overlap	–0.73 (–18.53, 17.07)	0.94	–	–
Joint involvement	8.11 (–8.59, 24.82)	0.34	–	–
Tendon friction rub	–17.54 (–53.04, 17.96)	0.33	–	–
Mild peripheral vasculopathy	–15.23 (–50.85, 20.40)	0.40	–	–
Severe peripheral vasculopathy	10.35 (–12.27, 32.98)	0.37	–	–
Calcinosis	22.90 (–0.62, 46.43)	0.06	20.76 (–5.74, 47.27)	0.12
Inflammatory myositis	–8.82 (–28.58, 10.93)	0.38	–	–
Gastrointestinal involvement	–0.54 (–18.59, 17.52)	0.95	–	–
ILD	7.06 (–11.28, 25.39)	0.44	–	–
PAH	13.15 (–16.80, 43.11)	0.38	–	–
Renal involvement	15.13 (–10.51, 40.76)	0.24	–	–
Cardiac involvement†	–	–	–	–
Maximum MRSS	–0.40 (–1.27, 0.47)	0.36	–	–
FVC	–0.16 (–0.65, 0.33)	0.52	–	–
Anti–Scl-70	–0.68 (–15.55, 14.19)	0.93	–	–
Anticentromere	3.90 (–16.23, 24.04)	0.70	–	–
Immunosuppressive treatment (ref. none)	4.43 (–12.72, 21.58)	0.61	–	–
Medication for peripheral vasculopathy (ref. none)	13.22 (–12.47, 38.91)	0.31	–	–
Medication for PAH (ref. none)	–4.59 (–65.73, 56.55)	0.88	–	–

* 95% CI = 95% confidence interval; dcSSc = diffuse cutaneous systemic sclerosis; FVC = forced vital capacity; ILD = interstitial lung disease; lcSSc = limited cutaneous SSc; MRSS = modified Rodnan skin score; PAH = pulmonary arterial hypertension; RP = Raynaud's phenomenon.
† Coefficient was not calculated as all patients with the condition were unemployed.

presence of hyperlipidemia (coefficient –18.56 [95% CI –32.22, –4.90]; $P < 0.01$) was associated with activity impairment in multivariable analysis.

Economic burden of unemployment and work productivity loss. For unemployed patients, their mean ± SD age at unemployment was 44.2 ± 9.3 years, equating to 22.8 years of lost employment. The annual cost of unemployment (lost earning) was estimated to be SGD $53,244 per patient. The overall cost of unemployment in the remaining working life of unemployed patients was estimated to be SGD $1,213,963 per patient. For employed patients, the annual cost of work productivity loss was estimated to be SGD $13,045 per patient. The annual cost of absenteeism and presenteeism was estimated to be SGD $3,301 and $9,744 per patient, respectively.

Based on an estimated prevalence of 78–100 per million population in Singapore, the annual cost of unemployment and work productivity loss was estimated to be $5.0–6.4 million and $2.9–3.7 million, respectively. The overall cost of unemployment in the remaining working life of unemployed patients was estimated to be SGD $113–145 million.

DISCUSSION

This study examined the work productivity and economic burden of unemployment and work productivity loss among patients with SSc in a multiethnic Asian population. Nearly 30% of patients with SSc were unemployed in our study cohort. They averaged an age of 44.2 years at unemployment, which equates to 22.8 years of lost employment in Singapore. Work productivity loss was present in approximately half (53.4%) of the employed patients, while activity impairment was also reported in approximately half of both the employed patients (47.4%) and unemployed patients (57.6%), both affecting nearly 50% of the working week. This is the first study investigating the work productivity and economic burden of SSc in Southeast Asia. Findings

Table 5. Factors associated with activity impairment in all SSc patients (n = 111)*

	Univariate analysis		Multivariable analysis	
	Coefficient (95% CI)	*P*	Coefficient (95% CI)	*P*
Female	−2.73 (−19.13, 13.66)	0.74	8.36 (−8.78, 25.49)	0.34
Age at survey	0.08 (−0.44, 0.60)	0.77	0.09 (−0.42, 0.60)	0.73
Disease duration, first non-RP symptom onset to survey	0.34 (−0.50, 1.18)	0.43	0.05 (−0.80, 0.89)	0.91
Smoking history	2.60 (−16.51, 21.71)	0.79	–	–
Comorbidities				
Hypertension	6.37 (−8.03, 20.78)	0.38	–	–
Ischemic heart disease	−6.11 (−39.67, 27.45)	0.72	–	–
Hyperlipidemia	−17.90 (−31.39, −4.42)	0.01	−18.56 (−32.22, −4.90)	<0.01
Diabetes mellitus	12.85 (−16.27, 41.97)	0.38	–	–
Renal impairment	23.22 (−5.66, 52.11)	0.11	–	–
Stroke	−17.94 (−58.73, 22.86)	0.39	–	–
Malignancy	−2.71 (−31.92, 26.50)	0.85	–	–
SSc subtype (ref. dcSSc)				
lcSSc	−1.24 (−13.78, 11.30)	0.85	–	–
SSc overlap	3.96 (−10.48, 18.39)	0.59	–	–
Joint involvement	8.27 (−4.17, 20.71)	0.19	–	–
Tendon friction rub	0.17 (−26.41, 26.75)	0.99	–	–
Mild peripheral vasculopathy	−7.81 (−31.85, 16.23)	0.52	–	–
Severe peripheral vasculopathy	17.44 (3.04, 31.84)	0.02	14.82 (−0.95, 30.58)	0.07
Calcinosis	22.30 (5.28, 39.32)	0.01	18.92 (−0.03, 37.88)	0.05
Inflammatory myositis	−0.60 (−15.45, 14.24)	0.94	–	–
Gastrointestinal involvement	6.75 (−7.10, 20.59)	0.34	–	–
ILD	3.72 (−9.68, 17.12)	0.58	–	–
PAH	−2.97 (−24.46, 18.52)	0.78	–	–
Renal involvement	12.40 (−3.36, 28.15)	0.12	–	–
Cardiac involvement	21.74 (−11.32, 54.81)	0.20	–	–
Maximum MRSS	0.14 (−0.46, 0.75)	0.64	–	–
FVC	−0.26 (−0.63, −0.10)	0.16	–	–
Anti-Scl-70	1.38 (−9.97, 12.73)	0.81	–	–
Anticentromere	−0.55 (−17.04, 15.93)	0.95	–	–
Immunosuppressive treatment (ref. none)	−4.15 (−16.66, 8.37)	0.51	–	–
Peripheral vasculopathy medication (ref. none)	11.42 (−7.48, 30.31)	0.23	–	–
Medication for PAH (ref. none)	12.85 (−16.27, 41.97)	0.38	–	–

* 95% CI = 95% confidence interval; dcSSc = diffuse cutaneous systemic sclerosis; FVC = forced vital capacity; ILD = interstitial lung disease; lcSSc = limited cutaneous SSc; MRSS = modified Rodnan skin score; PAH = pulmonary arterial hypertension; ref. = reference; RP = Raynaud's phenomenon.

from this study could provide some insights into interventions to mitigate work impairment in patients with SSc and reduce the disease burden in the population.

The unemployment rate in our cohort of SSc patients (29.7%) was relatively lower than those reported in other populations (approximately one-third to two-thirds) (2,7,18,38,39). Work productivity loss was also reported by a smaller proportion of employed patients in our study (53.4%) compared to the study by Morrisroe et al in Australia (63.6%) (18), but the extent of work productivity loss (the proportion of the working week impaired) experienced by our patients was greater (45.9% versus 38.4%). The mean work productivity loss in all employed SSc patients (24.5%) was comparable to SSc patients in Australia (24.4%) (18) and axial spondyloarthritis patients in Singapore (27.6%) (40). While different disease characteristics, including disease duration and severity across study populations, may partly explain this variation, other factors including work and psychosocial characteristics may have also influenced a patient's employment status and work productivity (12,41). In addition to unemployment and work productivity loss, other measures of work disability, such as reduced work ability and work transitions/changes, have also been reported among patients with SSc. In the study by Sandqvist et al, 72.9% of the employed patients with SSc reported poor or moderate work ability after a mean disease duration of 13 years (38). In the studies by Nguyen et al (42), Bérezné et al (2), Decuman et al (7), and Morrisroe et al (18), 31.0–56.0% of patients with SSc reported work transitions/changes after their diagnosis of SSc.

As work characteristics, including employment status, occupation, and work productivity and ability, may change with different health status conditions and other contextual factors, longitudinal studies are required to capture the dynamic change of work disability and to explore its relationship with various contextual factors. Sandqvist et al found that the relative risk for sick leave or disability pension in patients with SSc compared to an age- and sex-matched reference group increased from 2.09 at 1 year from disease onset to 2.41 at 3 years (43). Sharif et al found that 26.7% of patients who were employed at baseline

self-reported work disability after an average of 4 years of follow-up (44). Hudson et al found that the odds of reporting work disability increased by ~15% with every 5 additional years of disease (45). However, in our study, we did not find an association between disease duration and unemployment or work productivity loss.

Exploring risk factors for unemployment and work productivity loss among patients with SSc can provide a basis for the development of interventions to prevent and mitigate the impact of disease on work. Currently identified factors for unemployment and productivity loss vary across studies, including demographic (2,38,39,43,44), disease-related (2,7,38,39,42,44–47), work-related (38,46), and sociopsychologic characteristics (42,44), many of which are potentially modifiable with appropriate interventions. In accordance with findings from previous studies (2,7,10,38,44,45), organ involvement such as kidney, severe peripheral vasculopathy, and severity of lung involvement (FVC) were associated with unemployment. Lung involvement (ILD and PAH) is the most common cause of mortality in SSc (48). Notably, while the presence of ILD and PAH was not associated with unemployment, the severity of disease with lower FVC was associated with unemployment. However, none of these associations were significant in multivariable analysis after adjustment for age, sex, and disease duration, possibly due to the small sample size.

Hyperlipidemia was negatively associated with work productivity loss and activity impairment, i.e., patients with hyperlipidemia compared to those without hyperlipidemia had better outcomes. This finding could be due to the fact that the majority of these patients (57%, data not shown) were treated with a statin, which has been shown to improve impaired endothelial function, lower high-sensitivity C-reactive protein and immune complex production (49), and improve skin microvascular function in SSc (50). Therefore, the treatment of hyperlipidemia with statins, rather than hyperlipidemia per se, may have been associated with better outcomes of work productivity and activity impairment. Disease duration, which was found to be associated with work disability in some studies (2,7,45,47) was not significant in our study. Decuman et al studied work participation and transition among SSc patients of working age and found that patients who reported work transitions had significantly longer disease duration compared to those with no work transition (6.3 versus 2.7 years) (7). Hence, patients may have received job accommodation/work transitions or have adjusted themselves to the demands at work, all of which may have offset the impact of more severe disease or disease damage due to longer disease duration.

Several systematic reviews have been conducted on the factors associated with work impairment of SSc, which yielded inconsistent findings. Schouffoer et al found only moderate evidence for the associations with disease-specific symptoms, functional disability, and quality of life (12). Decuman et al found that, among the factors with a strong association (global disability, income, muscle/skin involvement, and demands at work), only global disability was reported frequently in the literature (10). McCormick et al found that the productivity loss was mainly driven by more generalized factors such as pain, fatigue, depression, cognitive dysfunction, and being overweight (11), which we did not investigate in our study. While this inconsistency could be mainly attributed to the diversity of contextual factors examined in various studies, it also highlights the importance of defining the essential contextual factors in studying the impact of disease on work (41).

Significant economic burden of SSc has been reported in the literature, despite different components of the cost estimated and data resources used across studies (11,13). In this study, we estimated the cost of SSc associated with unemployment and work productivity loss among affected individuals. Nearly 30% of our patients with SSc were unemployed at the time of this study. They lost an average of 22.8 years of employment, which was much longer than those reported in other published studies and hence a much higher cost of unemployment (lost earning) to society (14,18–22,39). Our estimated cost of work productivity loss was also relatively higher compared to the cost reported in the study by Morrisroe et al, in which the same measurement of work productivity loss (WPAI) was used (18).

There are several limitations in this study. First, data on employment status before and at disease onset were not collected. Possibly some of our patients were unemployed before the onset of SSc. Thus, the impact of SSc on unemployment status might have been overestimated in our study. Second, other than the demographic and disease characteristics examined in this study, other factors, including work and sociopsychologic characteristics that were not examined, may have also played important roles in the unemployment and work impairment in our patients. This possibility would be the subject of future studies, which require a larger sample size to examine multiple factors. Last, we studied patients attending a tertiary referral center who might have more severe disease compared to those who were seeking treatment in nontertiary clinics and those who were not seeking treatment in the community. The impact and burden of SSc might thus have been overestimated in our population.

SSc imposes significant unemployment and work productivity loss and causes substantial economic burden to both affected individuals and society. Modifying the identified factors associated with unemployment and work productivity loss may reduce the burden of SSc.

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. Low 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. Xiang, Low.

Acquisition of data. Kua.

Analysis and interpretation of data. Xiang, Low.

REFERENCES

1. Castellvi I, Eguiluz S, Escudero-Contreras A, Rios JJ, Calvo-Alen J, Callejas-Rubio JL, et al. LAUDES study: impact of digital ulcers on hand functional limitation, work productivity and daily activities, in systemic sclerosis patients. Rheumatol Int 2019;39:1875–82.
2. Bérezné A, Seror R, Morell-Dubois S, de Menthon M, Fois E, Dzeing-Ella A, et al. Impact of systemic sclerosis on occupational and professional activity with attention to patients with digital ulcers. Arthritis Care Res (Hoboken) 2011;63:277–85.
3. Sandqvist G, Scheja A, Eklund M. Working ability in relation to disease severity, everyday occupations and well-being in women with limited systemic sclerosis. Rheumatology (Oxford) 2008;47:1708–11.
4. Sandqvist G, Eklund M. Daily occupations: performance, satisfaction and time use, and relations with well-being in women with limited systemic sclerosis. Disabil Rehabil 2008;30:27–35.
5. Sandqvist G, Akesson A, Eklund M. Daily occupations and well-being in women with limited cutaneous systemic sclerosis. Am J Occup Ther 2005;59:390–7.
6. Klak A, Raciborski F, Samel-Kowalik P. Social implications of rheumatic diseases. Reumatologia 2016;54:73–8.
7. Decuman S, Smith V, Verhaeghe S, Deschepper E, Vermeiren F, De Keyser F. Work participation and work transition in patients with systemic sclerosis: a cross-sectional study. Rheumatology (Oxford) 2012;51:297–304.
8. Brouwer WB, Meerding WJ, Lamers LM, Severens JL. The relationship between productivity and health-related QOL: an exploration. Pharmacoeconomics 2005;23:209–18.
9. Sandqvist JL, Henriksson CM. Work functioning: a conceptual framework. Work 2004;23:147–57.
10. Decuman S, Smith V, Verhaeghe ST, Van Hecke A, De Keyser F. Work participation in patients with systemic sclerosis: a systematic review. Clin Exp Rheumatol 2014;32 Suppl 86:S206–13.
11. McCormick N, Marra CA, Avina-Zubieta JA. Productivity losses and costs in the less-common systemic autoimmune rheumatic diseases. Curr Rheumatol Rep 2017;19:72.
12. Schouffoer AA, Schoones JW, Terwee CB, Vliet Vlieland TP. Work status and its determinants among patients with systemic sclerosis: a systematic review. Rheumatology (Oxford). 2012;51:1304–14.
13. Fischer A, Zimovetz E, Ling C, Esser D, Schoof N. Humanistic and cost burden of systemic sclerosis: a review of the literature. Autoimmun Rev 2017;16:1147–54.
14. Kawalec PP, Malinowski KP. The indirect costs of systemic autoimmune diseases, systemic lupus erythematosus, systemic sclerosis and sarcoidosis: a summary of 2012 real-life data from the Social Insurance Institution in Poland. Expert Rev Pharmacoecon Outcomes Res 2015;15:667–73.
15. Taylor PC, Moore A, Vasilescu R, Alvir J, Tarallo M. A structured literature review of the burden of illness and unmet needs in patients with rheumatoid arthritis: a current perspective. Rheumatol Int 2016;36:685–95.
16. Carter EE, Barr SG, Clarke AE. The global burden of SLE: prevalence, health disparities and socioeconomic impact. Nat Rev Rheumatol 2016;12:605–20.
17. Sieper J, Hu X, Black CM, Grootscholten K, van den Broek RW, Kachroo S. Systematic review of clinical, humanistic, and economic outcome comparisons between radiographic and non-radiographic axial spondyloarthritis. Semin Arthritis Rheum 2017;46:746–53.
18. Morrisroe K, Sudararajan V, Stevens W, Sahhar J, Zochling J, Roddy J, et al. Work productivity in systemic sclerosis, its economic burden and association with health-related quality of life. Rheumatology (Oxford) 2018;57:73–83.
19. Lopez-Bastida J, Linertova R, Oliva-Moreno J, Serrano-Aguilar P, Posada-de-la-Paz M, Kanavos P, et al. Social/economic costs and health-related quality of life in patients with scleroderma in Europe. Eur J Health Econ 2016;17 Suppl 1:109–17.
20. Chevreul K, Brigham KB, Gandre C, Mouthon L, Network BR. The economic burden and health-related quality of life associated with systemic sclerosis in France. Scand J Rheumatol 2015;44:238–46.
21. Furst DE, Fernandes AW, Iorga SR, Greth W, Bancroft T. Annual medical costs and healthcare resource use in patients with systemic sclerosis in an insured population. J Rheumatol 2012;39:2303–9.
22. Minier T, Pentek M, Brodszky V, Ecseki A, Karpati K, Polgar A, et al. Cost-of-illness of patients with systemic sclerosis in a tertiary care centre. Rheumatology (Oxford) 2010;49:1920–8.
23. Proudman SM, Huq M, Stevens W, Wilson ME, Sahhar J, Baron M, et al. What have multicentre registries across the world taught us about the disease features of systemic sclerosis? J Scleroderma Rel Disord 2017;2:169–82.
24. Ling AL, Gee TG, Giap LW, Cheng NS, Santosa A, Chan G, et al. Disease characteristics of the Singapore Systemic Sclerosis Cohort. Proc Singapore Healthcare 2013;22:8–14.
25. Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
26. Matucci-Cerinic M, Allanore Y, Czirjak L, Tyndall A, Muller-Ladner U, Denton C, et al. The challenge of early systemic sclerosis for the EULAR Scleroderma Trial and Research group (EUSTAR) community: it is time to cut the Gordian knot and develop a prevention or rescue strategy. Ann Rheum Dis 2009;68:1377–80.
27. Reilly MC, Gooch KL, Wong RL, Kupper H, van der Heijde D. Validity, reliability and responsiveness of the Work Productivity and Activity Impairment Questionnaire in ankylosing spondylitis. Rheumatology (Oxford) 2010;49:812–9.
28. LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger TA Jr, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15:202–5.
29. Santosa A, Tan CS, Teng GG, Fong W, Lim A, Law WG, et al. Lung and gastrointestinal complications are leading causes of death in SCORE, a multi-ethnic Singapore systemic sclerosis cohort. Scand J Rheumatol 2016;45:499–506.
30. Clements P, Lachenbruch P, Siebold J, White B, Weiner S, Martin R, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. J Rheumatol 1995;22:1281–5.
31. Ministry of Manpower Singapore. Responsible re-employment. 2019. URL: https://www.mom.gov.sg/employment-practices/re-employment.
32. Tan ME, Sagayadevan V, Abdin E, Picco L, Vaingankar J, Chong SA, et al. Employment status among the Singapore elderly and its correlates. Psychogeriatrics 2017;17:155–63.
33. Weisbrod BA. The valuation of human capital. J Polit Econ 1961;69:425–36.
34. Ministry of Manpower Singapore. Gross monthly income from work 2019. URL: https://stats.mom.gov.sg/Pages/Income-Summary-Table.aspx.
35. Department of Statistics Singapore. Population and population structure. 2019. URL: https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/latest-data.
36. Kang GW, Jung KH, Lee YS, Kim HJ, Yoon DY, Lee SH, et al. Incidence, prevalence, mortality and causes of death in systemic sclerosis in Korea: a nationwide population-based study. Br J Dermatol 2018;178:e37–9.
37. Kuo CF, See LC, Yu KH, Chou IJ, Tseng WY, Chang HC, et al. Epidemiology and mortality of systemic sclerosis: a nationwide population study in Taiwan. Scand J Rheumatol 2011;40:373–8.

38. Sandqvist G, Scheja A, Hesselstrand R. Pain, fatigue and hand function closely correlated to work ability and employment status in systemic sclerosis. Rheumatology (Oxford) 2010;49:1739–46.
39. Singh MK, Clements PJ, Furst DE, Maranian P, Khanna D. Work productivity in scleroderma: analysis from the University of California, Los Angeles Scleroderma Quality of Life Study. Arthritis Care Res (Hoboken) 2012;64:176–83.
40. Goh Y, Kwan YH, Leung YY, Fong W, Cheung PP. A cross-sectional study on factors associated with poor work outcomes in patients with axial spondyloarthritis in Singapore. Int J Rheum Dis 2019;22: 2001–8.
41. Tang K, Escorpizo R, Beaton DE, Bombardier C, Lacaille D, Zhang W, et al. Measuring the impact of arthritis on worker productivity: perspectives, methodologic issues, and contextual factors. J Rheumatol 2011; 38:1776–90.
42. Nguyen C, Poiraudeau S, Mestre-Stanislas C, Rannou F, Berezne A, Papelard A, et al. Employment status and socio-economic burden in systemic sclerosis: a cross-sectional survey. Rheumatology (Oxford) 2010;49:982–9.
43. Sandqvist G, Hesselstrand R, Petersson IF, Kristensen LE. Work disability in early systemic sclerosis: a longitudinal population-based cohort study. J Rheumatol 2015;42:1794–800.
44. Sharif R, Mayes MD, Nicassio PM, Gonzalez EB, Draeger H, McNearney TA, et al. Determinants of work disability in patients with systemic sclerosis: a longitudinal study of the GENISOS cohort. Semin Arthritis Rheum 2011;41:38–47.
45. Hudson M, Steele R, Lu Y, Thombs BD, Canadian Scleroderma Research Group, Baron M. Work disability in systemic sclerosis. J Rheumatol 2009;36:2481–6.
46. Morrisroe K, Huq M, Stevens W, Rabusa C, Proudman SM, Nikpour M, et al. Determinants of unemployment amongst Australian systemic sclerosis patients: results from a multicentre cohort study. Clin Exp Rheumatol 2016;34 Suppl 100:79–84.
47. Ouimet JM, Pope JE, Gutmanis I, Koval J. Work disability in scleroderma is greater than in rheumatoid arthritis and is predicted by high HAQ scores. Open Rheumatol J 2008;2:44–52.
48. Nikpour M, Baron M. Mortality in systemic sclerosis: lessons learned from population-based and observational cohort studies. Curr Opin Rheumatol 2014;26:131–7.
49. Timar O, Szekanecz Z, Kerekes G, Vegh J, Olah AV, Nagy G, et al. Rosuvastatin improves impaired endothelial function, lowers high sensitivity CRP, complement and immuncomplex production in patients with systemic sclerosis: a prospective case-series study. Arthritis Res Ther 2013;15:R105.
50. Rossi M, Bazzichi L, Ghiadoni L, Mencaroni I, Franzoni F, Bombardieri S. Increased finger skin vasoreactivity and stimulated vasomotion associated with simvastatin therapy in systemic sclerosis hypercholesterolemic patients. Rheumatol Int 2012;32:3715–21.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 828–832
DOI 10.1002/acr.24527

BRIEF REPORT

Finger Systolic Blood Pressure Index Measurement: A Useful Tool for the Evaluation of Arterial Disease in Patients With Systemic Sclerosis

Sophie Blaise,[1] Carine Boulon,[2] Marion Mangin,[2] Patricia Senet,[3] Isabelle Lazareth,[4] Bernard Imbert,[1] François-Xavier Lapebie,[5] Philippe Lacroix,[6] Joël Constans,[2] and Patrick Carpentier[1]

Objective. To evaluate the prevalence and clinical correlates of peripheral arterial disease of the upper limbs in patients with systemic sclerosis (SSc), as detected with finger brachial pressure index (FBPI) measurements.

Methods. This work is based on the baseline data of the SCLEROCAP multicenter cohort of SSc patients. Finger systolic blood pressure was measured with laser Doppler flowmetry, and the FBPI was obtained as its ratio over the ipsilateral brachial systolic blood pressure. An FBPI of <0.70 was used as the diagnostic criterion for occlusive arterial disease of the upper limbs. Thus, the prevalence of defined arterial disease as well as its clinical, biologic, and capillaroscopic correlates were evaluated.

Results. Among 326 enrolled patients, 177 (54.3%) met the criterion for arterial disease (FBPI <0.70). No association was found with the type of SSc nor with the type of associated antinuclear antibodies, but a significant association was found with the duration of the disease ($P < 0.001$), the capillaroscopic pattern ($P < 0.001$), and most strikingly with the presence of digital ulcers (42.9% versus 13.4%; $P < 0.001$). A quantitative relationship was found between the FBPI and the prevalence of digital ulcers and was shown to be independent from the capillaroscopic pattern.

Conclusion. This cross-sectional study shows a high prevalence of arterial disease of the upper limbs in patients with SSc. FBPI appears to be a strong and independent predictor of digital ulcers. This study suggests that both macro- and microangiopathy are contributing to the ischemic damage of the fingertips.

INTRODUCTION

Systemic sclerosis (SSc) is a life-threatening and disabling chronic disease involving the connective tissue and vasculature of the skin and internal organs. Vascular disorders of the extremities are a key feature of this disease that most often begins with severe Raynaud's phenomenon and is frequently associated with digital ulcers that are disabling and can lead to severe ischemic complications. Although a specific microangiopathy easily detected by capillaroscopy is thought to be the major cause of such ischemic digital ulcers, arterial involvement of the hand and finger arteries has also been found frequently and is associated with digital ulcers in a small series of patients using various imaging techniques from arteriography to Duplex ultrasound imaging (1–6). However, none of these techniques allows an easy quantification of the arterial functional deficit.

Distal blood pressure measurements, using various plethysmographic techniques or laser Doppler flowmetry (LDF), allow for a quantitative functional evaluation of the peripheral arterial system. The blood measurements are recommended and routinely performed at the toe level for the evaluation of patients with suspected critical limb ischemia of the lower limbs (7), and the toe brachial index has been proposed as a

Supported by the Société Française de Médecine Vasculaire, the Société Française de Microcirculation, and Actelion Pharmaceuticals, France.

[1]Sophie Blaise, MD, PhD, Bernard Imbert, MD, Patrick Carpentier, MD, PhD: Grenoble-Alpes University Hospital, Grenoble, France; [2]Carine Boulon, MD, Marion Mangin, Joël Constans, MD, PhD: Saint-André Hospital, Bordeaux, France; [3]Patricia Senet, MD: Tenon Hospital, Paris, France; [4]Isabelle Lazareth, MD: Saint Joseph Hospital, Paris, France; [5]François-Xavier Lapebie, MD: Rangueil Hospital, Toulouse, France; [6]Philippe Lacroix, MD, PhD: University Hospital, Limoges, France.

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

Address correspondence to Sophie Blaise, MD, PhD, Department of Vascular Medicine, Grenoble-Alpes University Hospital, CS 10217 38043, Grenoble CEDEX 09, France. Email: SBlaise@chu-grenoble.fr.

Submitted for publication April 29, 2020; accepted in revised form December 1, 2020.

SIGNIFICANCE & INNOVATIONS

- The finger brachial pressure index (FBPI) can be easily and noninvasively measured in patients with systemic sclerosis. It allows the detection, with a diagnostic threshold level of <0.70, and the quantification of arterial disease of the upper limbs.
- Arterial disease of the upper limbs, as detected with this method, is strongly associated with the prevalence of digital ulcers and with the severity of microangiopathy as evaluated by capillaroscopy.
- Both arterial disease assessed by FBPI and microangiopathy evaluated by capillaroscopy are independently associated with digital ulcers.

more reliable technique than the ankle brachial index for the diagnosis of peripheral arterial disease (PAD), especially in patients with diabetes mellitus, due to the frequent calcifications at the ankle level and to a more distal site of arterial involvement (7). Such blood measurements are also used in the upper limbs, in recent years, mostly for the prediction and evaluation of hemodialysis access–induced distal ischemia (HAIDI) (8,9).

The SCLEROCAP multicenter cohort study provided us with an opportunity to assess the potential of this approach in a large series of patients with SSc who benefited from a systematic clinical, biologic, and capillaroscopic examination. We also looked for the prevalence and clinical correlates of PAD of the upper limbs as detected and quantified by finger brachial pressure index (FBPI) measurements in these patients.

MATERIALS AND METHODS

This work is a predefined ancillary study of the SCLEROCAP cohort study on SSc, which is an observational prospective multicenter study aiming at evaluating the prognostic value of capillaroscopic classifications during a 3-year follow-up (10). SCLEROCAP was designed and performed in compliance with good clinical practices and the Declaration of Helsinki. Formal ethics committee approval was obtained on March 7, 2014.

Details of the methods are described in a previous publication (10). Briefly, consecutive patients age ≥18 years, presenting

Table 1. Systemic sclerosis (SSc) features associated with arterial disease of the upper limbs*

	Arterial disease FBPI <0.70 (n = 177)	No significant arterial disease FBPI ≥0.70 (n = 149)	Total population (n = 326)	*P*
Female/male, no.	153/24	123/26	276/50	NS
Age, mean ± SD years	58.6 ± 14.5	56.3 ± 12.5	57.5 ± 13/7	0.015
Smoker, current or past	30.5	32.9	31.6	NS
Body mass index, mean ± SD kg/m^2	23.8 ± 4.8	23.7 ± 4.4	23.8 ± 4.6	NS
Type of SSc				NS
Limited	18.1	26.8	22.1	NS
Cutaneous limited	71.2	66.4	69.0	NS
Diffuse cutaneous	10.7	6.7	8.9	NS
Raynaud's phenomenon	94.4	95.3	94.8	NS
Time since Raynaud's phenomenon onset, mean ± SD years	14.8 ± 13.3	12.0 ± 9.7	13.6 ± 11.9	0.002
Time since SSc diagnosis, mean ± SD years	7.8 ± 8.1	5.6 ± 5.8	6.8 ± 7.2	<0.001
Modified Rodnan score, mean ± SD	6.2 ± 5.6	4.0 ± 4.8	5.2 ± 5.5	0.032
Pitting scars/active ulcers	42.9	13.4	29.4	<0.001
Active ulcers only	15.3	1.3	8.9	<0.001
Antinuclear antibodies				NS
Anticentromere	58.2	55.7	57.1	NS
Anti–Scl-70	22.0	14.2	18.7	NS
Anti-RNase III	3.6	4.2	4.2	NS
Capillaroscopy: Maricq classification				
No SSc pattern	11.3	18.1	14.4	<0.001
Slow pattern	50.3	65.1	57.1	<0.001
Active pattern	38.4	16.8	28.5	<0.001
Capillaroscopy: Cutolo classification				
No SSc pattern	11.3	18.1	14.4	<0.001
Early	18.1	30.9	23.9	<0.001
Active	39.0	37.6	38.3	<0.001
Late	31.6	13.4	23.3	<0.001

* Values are the percentage unless indicated otherwise. FBPI = Finger Brachial Pressure Index; NS = nonsignificant.

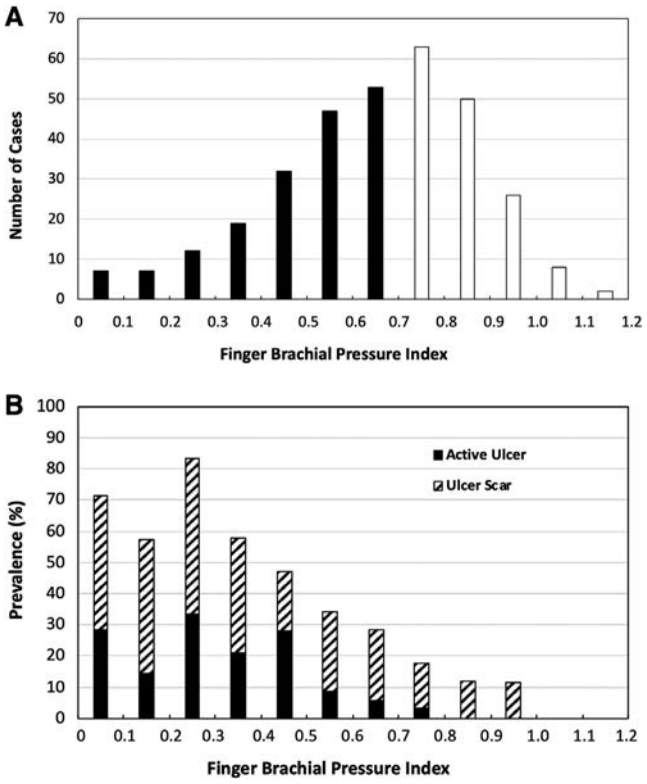

Figure 1. **A,** Histogram of finger brachial pressure index (FBPI), as measured in 326 patients with systemic sclerosis (SSc). An FBPI of <0.70 in at least 1 upper limb was the criterion used for the diagnosis of occlusive arterial disease of the upper limbs. **B,** Prevalence of active ulcers and ulcer scars according to the level of FBPI in 326 patients with systemic sclerosis.

with SSc according to the LeRoy and Medsger's criteria (in order to include early SSc) (11), were enrolled for a 3-year follow-up in 10 French medical departments with different specialties (vascular medicine, rheumatology, internal medicine, and dermatology). In addition to the SSc work-up and capillaroscopy rating according to the Maricq (12) and Cutolo (13) classifications, a systematic measurement of digital arterial pressures with calculation of the FBPI was performed at baseline in 5 of the 10 centers.

Briefly, Maricq's classification distinguishes 2 different patterns: a slow one, with giant capillaries but no or minimal avascular areas, and an active one, with few or absent giant capillaries but with large avascular areas and architectural disorganization (12). In Cutolo's classification, 3 different patterns are considered. The early pattern contained few giant capillaries and hemorrhage and preserved capillary organization with no evident capillary loss. The active pattern contained frequent giant capillaries and hemorrhage and moderate capillary loss. The late pattern consisted of few or absent giant capillaries, the absence of hemorrhage, severe capillary loss with large avascular areas, disorganization of capillary loops, and the presence of bushy capillaries (13).

Finger systolic blood pressure (FSBP) was measured in all centers using LDF (Periflux System 5000; Perimed) on 8 fingers, all but the thumbs. The patient was resting for 15 minutes, before the pressure measurements, in a room with a comfortable temperature (between 20° and 25°C) according to the season, and after warming the fingers using different techniques according to each center (14,15). A cuff of suitable size was fitted on the proximal or medial phalanx. Before inflating the cuff, the pulp of the finger was emptied by gentle manual compression. A suprasystolic pressure (200 mm Hg) was applied, and after a baseline recording of 3 seconds, the cuff was automatically and gradually deflated at a rate of 3.4 mm Hg/second. The system also provided automatic detection of the reappearance of the flow and the corresponding pressure value as the systolic pressure. Each measurement was repeated 3 times for a given finger and the median value recorded as the FSBP. The FBPI was calculated as the ratio of the lowest FSBP of the limb over the ipsilateral brachial systolic blood pressure. The lowest of the 2 upper limbs' FBPI was selected as the patient FBPI. In accordance with the lower limbs' most used criterion (7), an FBPI of <0.70 was considered the threshold level for significant occlusive arterial disease of the upper limbs. Statistical analysis was performed using SPSS software, version 25, for descriptive statistics, chi-square tests (categorical variables) and variance analysis (quantitative variables) for hypothesis testing. A P value less than 0.05 was considered significant.

RESULTS

A total of 326 patients, of 387 in the whole SCLEROCAP study group, were included in the current study by the 5 centers where FBPI measurements were systematically recorded. The patients' main characteristics are reported in Table 1. The FBPI histogram is shown in Figure 1A. Our criterion of an FBPI of <0.70 for significant occlusive arterial disease was met in as many as 177 patients (54.3%); notably, 45 patients (13.8%) were found to have an FBPI of <0.40.

The clinical, immunologic, and capillaroscopic correlates of the presence of an arterial disease of the upper limbs (FBPI <0.70) are reported in Table 1. No significant association was found with age, sex, body mass index, or smoking habits, or with the type of SSc and type of associated antinuclear antibodies. By contrast, a significant relationship was found with the duration of the disease (delays from Raynaud's phenomenon onset and from SSc diagnosis; $P < 0.001$ for both parameters). But most interestingly, a strong and highly significant association was found with an aggressive capillaroscopic pattern ($P < 0.001$ for both Maricq and Cutolo classifications), as well as with the presence of all digital ulcers (42.9% versus 13.4%; $P < 0.001$) and active digital ulcers (15.3% versus 1.3%; $P < 0.001$).

Furthermore, the prevalence of both all ulcers and active ulcers progressively and strongly increased with lower levels of FBPI, as shown in Figure 1B, and this finding remained significant

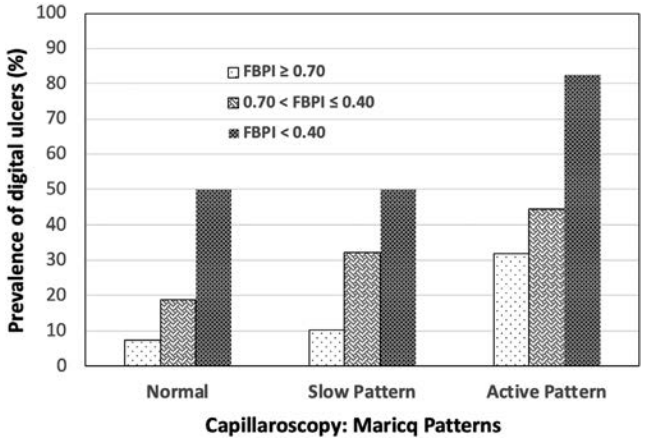

Figure 2. Prevalence of digital ulcers (active or scar) according to finger brachial pressure index (FBPI) level and capillaroscopic status (Maricq pattern).

after adjusting for the capillaroscopic pattern (variance analysis Maricq or Cutolo). A low FBPI and an aggressive capillaroscopic pattern were strong independent factors for digital ulcers (Figure 2).

DISCUSSION

Several studies have reported the frequent presence of PAD of the upper limbs in patients with SSc and its association with digital ulcers, using arteriography (1), magnetic resonance arterial imaging (2), computed tomography (3), plethysmography (4), and ultrasound duplex (5,6), as well as finger pressure measurements in a small series (1). However, our study is the first to evaluate the prevalence of PAD in a large multicenter cohort of patients with SSc, showing its high magnitude and strength of association with digital ulcers. This study is also the first exploring the association with capillaroscopic findings on a large scale and to show that both arterial and microvascular angiopathies are statistically independent factors for digital ulcers in patients with SSc.

Compared to the other arterial evaluation techniques, the distal arterial pressure measurements have the advantage of allowing for a quantification of the arterial functional deficit. They have been extensively studied in the lower limbs, where toe blood pressure measurements are now used routinely in patients with severe PAD for prognostic classification and clinical decision-making. The toe brachial pressure index is also increasingly recommended for the diagnosis of PAD instead of the ankle brachial index, with a diagnostic threshold level of 0.70, especially in patients with diabetes mellitus, due to their more distal site of arterial lesions and the frequent occurrence of calcified ankle arteries, hampering the segmental pressure measurements at this level.

We chose to use the same threshold level for the upper limbs, in the absence of specific recommendations or relevant studies. Indeed, very few studies have evaluated upper limb arterial function using arterial pressure measurements until now, with the exception of studies in HAIDI (8,9). These studies used plethysmographic techniques, which are less sensitive than LDF (16,17) but are well-suited for this indication, where there is no need to accurately measure very low pressures, as is the case in patients with SSc. In the HAIDI studies, the contralateral brachial pressure was used as a reference for the FBPI calculation. Either 0.70 or 0.60 were found to be the best thresholds for the HAIDI risk, and 0.40 for the diagnosis of hand ischemia. The methodology used in that study showed an excellent repeatability of the measurements, whatever the warming method used and the site of measurement (15). The prevalence of arteriopathy we found could even be slightly underestimated due to the positioning of the cuff on the proximal phalanx in some centers.

In this study, arterial disease, remarkably, is strongly associated with both capillaroscopic abnormalities and digital ulcers, but with none of the nonvascular characteristics of SSc except its duration and a weak association with the Rodnan score that can very well be related to the disease duration. No association with smoking habits was found. By contrast, the strong and independent association of both microangiopathic and arterial disease with digital ulcers is in favor of their combined responsibility in the occurrence of such complications. This relationship will be explored further in the longitudinal follow-up of these patients, as well as the individual predictive values of the different parameters. However, we already propose that not only microvascular but also arterial factors should be evaluated in patients with SSc for a comprehensive understanding and management of their vascular complications, especially digital ulcers. The best way to estimate the risk of digital ulceration is probably to evaluate the risk factors with the macro- and microcirculatory status, and thus to combine finger pressure measurements and capillaroscopy. Systematic screening for arteriopathy of the upper limbs, particularly in patients with a long history of SSc, could lead to a less negative evolution toward digital ulcers.

ACKNOWLEDGMENTS

We thank all the investigators of the SCLEROCAP study and their skilled nurses who performed all the pressure measurements.

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. Blaise 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. Blaise, Constans, Carpentier.

Acquisition of data. Blaise, Boulon, Mangin, Senet, Lazareth, Imbert, Lapebie, Lacroix, Constans, Carpentier.

Analysis and interpretation of data. Blaise, Constans, Carpentier.

ROLE OF THE STUDY SPONSOR

Actelion had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision

to submit the manuscript for publication. Publication of this article was not contingent upon approval by Actelion.

REFERENCES

1. Hirai M. Arterial insufficiency of the hand evaluated by digital blood pressure and arteriographic findings. Circulation 1978;58:902–8.
2. Allanore Y, Seror R, Chevrot A, Kahan A, Drapé J.L. Hand vascular involvement assessed by magnetic resonance angiography in systemic sclerosis. Arthritis Rheum 2007;56:2747–54.
3. Emad Y, Al-Sherbeni H, Ragab Y, Abo-El-Youn I, El-Shaarawy N, Nassar DY, et al. Arterial vasculopathy in systemic sclerosis: computerized tomography (CT) angiographic features of macrovascular and microvascular upper limbs arteries. Joint Bone Spine 2014;81:433–7.
4. Aïssou L, Meune C, Avouac J, Meunier M, Elhaïd M, Sorbets E, et al. Small, medium but not large arteries are involved in digital ulcers associated with systemic sclerosis. Joint Bone Spine 2016;83:444–7
5. Lescoat A, Coiffier G, de Carlan M, Droitcourt C, Ballerie A, Cazalets C, et al. Combination of capillaroscopic and ultrasonographic evaluations in systemic sclerosis: results of a cross-sectional study. Arthritis Care Res (Hoboken) 2018;70:938–43.
6. Schioppo T, Orenti A, Boracchi P, De Lucia O, Murgo A, Ingegnoli F. Evidence of macro- and micro-angiopathy in scleroderma: an integrated approach combining 22-MHz power Doppler ultrasonography and video-capillaroscopy. Microvasc Res 2019;122:125–30.
7. Tehan PE, Santos D, Chuter VH. A systematic review of the sensitivity and specificity of the toe–brachial index for detecting peripheral artery disease. Vasc Med 2016;21:382–9.
8. Schanzer A, Nguyen LL, Owens CD, Schanzer H. Use of digital pressure measurements for the diagnosis of AV access-induced hand ischemia. Vasc Med 2006;11:227–31.
9. Beathard GA, Jennings WC, Wasse H, Shenoy S, Hentschel DM, Abreo K, et al. ASDIN white paper: assessment and management of hemodialysis access-induced distal ischemia by interventional nephrologists. J Vasc Access 2020;21:543–53.
10. Boulon C, Aiouaz S, Blaise S, Mangin M, Decamps-Le Chevoir J, Senet P, et al. Correlation between capillaroscopic classifications and disease severity in systemic sclerosis: results from SCLEROCAP study at inclusion. Clin Exp Rheumatol 2019;37 Suppl 119:S63–8
11. LeRoy EC, Medsger TA Jr. Criteria for the classification of early systemic sclerosis. J Rheumatol 2001;28:1573–6
12. Maricq HR, Weinberger AB, LeRoy EC. Early detection of scleroderma-spectrum disorders by in vivo capillary microscopy: a prospective study of patients with Raynaud's phenomenon. J Rheumatol 1982;9:289–91.
13. Cutolo M, Sulli A, Pizzorni C, Sccardo S. Nailfold videocapillaroscopy assessment of microvascular damage in systemic sclerosis. J Rheumatol 2000;27:155–60.
14. Beinder E, Hoffmann U, Franzeck UK, Huch A, Huch R, Bollinger A. Laser Doppler technique for the measurement of digital and segmental systolic blood pressure. Vasa 1992;21:15–21.
15. Blaise S, Constans J, Pellegrini L, Senet P, Lazareth I, Cracowski JL, et al. Optimizing finger systolic blood pressure measurements with laser Doppler: validation of the second phalanx site. Microvasc Res 2020;131:104029.
16. Ubbink DT. Toe blood pressure measurements in patients suspected of leg ischaemia: a new laser Doppler device compared with photoplethysmography. Eur J Vasc Endovasc Surg 2004;27:629–34
17. Hoyer C, Paludan JP, Pavar S, Biurrun Manresa JA, Petersen LJ. Reliability of laser Doppler flowmetry curve reading for measurement of toe and ankle pressures: intra- and inter-observer variation. Eur J Vasc Endovasc Surg 2014;47:311–8.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 833–840
DOI 10.1002/acr.24525

American College of Rheumatology
Empowering Rheumatology Professionals

Electromyographic Muscle Activity and Three-Dimensional Scapular Kinematics in Patients With Multidirectional Shoulder Instability: A Study in the Hypermobile Type of the Ehlers-Danlos Syndrome and the Hypermobility Spectrum Disorders

Valentien Spanhove,[1] Patrick Calders,[1] Kelly Berckmans,[1] Tanneke Palmans,[1] Fransiska Malfait,[2] Ann Cools,[1] and Inge De Wandele[3]

Objective. To investigate differences in electromyography (EMG), muscle activity, and scapular kinematics during elevation in the scapular plane between healthy controls, participants with multidirectional shoulder laxity (MDL), and patients with multidirectional shoulder instability (MDI) who are diagnosed with hypermobile Ehlers-Danlos syndrome (hEDS) or hypermobility spectrum disorder (HSD).

Methods. Twenty-seven women with hEDS/HSD and MDI, 27 female healthy control subjects, and 28 female subjects with MDL participated in this study. Scapular 3-dimensional kinematic data were obtained using 8 Oqus Qualisys cameras. Simultaneously, surface EMG was used to measure muscle activity of the upper, middle, and lower trapezius, infraspinatus, latissimus dorsi, serratus anterior, posterior deltoid, and pectoralis major during arm elevation in the scapular plane. Group differences were assessed using statistical parametric mapping.

Results. Regarding scapular kinematics, significantly less upward rotation was observed in hEDS/HSD patients with MDI compared to both healthy controls and MDL subjects. Significantly less posterior tilt was seen in hEDS/HSD patients compared to MDL subjects. Furthermore, significantly higher EMG activity of the infraspinatus, middle trapezius, and posterior deltoid was found in hEDS/HSD patients with MDI.

Conclusion. hEDS/HSD patients with MDI demonstrate altered scapular kinematics and increased EMG muscle activity compared to subjects without MDI. These findings could serve as a stepping stone for future research regarding treatment strategies in patients whose conditions belong to the hypermobility spectrum.

INTRODUCTION

Although multiple factors affect shoulder stability, capsular laxity plays a major role in the development of several clinical conditions, ranging from asymptomatic shoulder laxity to multidirectional shoulder instability (MDI) (1). Shoulder laxity is an asymptomatic condition characterized by increased translations of the articular surfaces (2). If this increased laxity occurs in ≥2 directions, it is referred to as multidirectional shoulder laxity (MDL) (3). By contrast, MDI is a symptomatic, pathologic condition associated with involuntary shoulder luxations and subluxations (1). Individuals with MDI experience symptoms that interfere with daily life activities such as shoulder pain and apprehension tension (4). Hence, some individuals with increased glenohumeral translations develop symptoms while others do not (5).

Shoulder instability may occur in isolation or as part of a more generalized joint hypermobility condition, such as the hypermobility spectrum disorder (HSD) or the hypermobile Ehlers-Danlos syndrome (hEDS). hEDS is a heritable connective tissue disorder, in which patients experience a generalized joint hypermobility with joint instability, accompanied by other systemic signs of tissue fragility, such as atrophic scarring, herniations of the abdominal wall, or organ prolapses (6,7). Individuals diagnosed with HSD

Dr. Spanhove's work was supported by the Special Research Fund of Ghent University (BOF17/DOC/220). Dr. Malfait's work was supported by the Research Foundation Flanders, Belgium.

[1]Valentien Spanhove, PT, PhD, Patrick Calders, PT, PhD, Kelly Berckmans, PT, Tanneke Palmans, Ing, Ann Cools, PT, PhD: Ghent University, Ghent, Belgium; [2]Fransiska Malfait, MD, PhD: Ghent University Hospital, Ghent, Belgium; [3]Inge De Wandele, PT, PhD: Ghent University and Ghent University Hospital, Ghent, Belgium.

Drs. Cools and De Wandele contributed equally to this article.

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

Address correspondence to Valentien Spanhove, PT, PhD, Department of Rehabilitation Sciences, Ghent University, Corneel Heymanslaan 10, B-9000 Ghent, Belgium. Email: valentien.spanhove@ugent.be.

Submitted for publication May 14, 2020; accepted in revised form November 24, 2020.

SIGNIFICANCE & INNOVATIONS

- This is the first study to provide evidence for an altered scapular movement and increased electromyography (EMG) muscle activity in hypermobile Ehlers-Danlos syndrome (hEDS)/hypermobility spectrum disorder (HSD) patients with multidirectional shoulder instability (MDI).
- Since alterations in scapular movement and EMG activity were found in hEDS/HSD patients and not in multidirectional shoulder laxity subjects, factors other than increased shoulder laxity may play a role.
- This study could serve as a stepping stone for future research regarding functional exercise therapy in hEDS/HSD patients with MDI.

experience symptomatic joint hypermobility with secondary musculoskeletal manifestations but do not meet the full criteria for hEDS regarding the systemic tissue fragility (6). Despite the difference in soft tissue fragility, musculoskeletal manifestations of hEDS and HSD mostly overlap and consist of recurrent joint dislocations, overload injuries, and chronic pain (6).

Although hEDS/HSD patients frequently experience recurrent shoulder dislocations (8), knowledge regarding MDI in patients with hEDS/HSD is limited and mainly derived from studies investigating MDI in individuals without generalized joint hypermobility. Additionally, findings regarding muscle activity in patients with MDI have been contradictory. While some studies have reported an increased and/or prolonged activation of rotator cuff muscles (9,10), other studies were unable to confirm this finding (3,11). Results regarding scapular movement in MDI are more consistent and indicate insufficient scapular upward rotation and increased scapular internal rotation (5,12). Although a considerable number of studies have investigated MDI, knowledge regarding MDL is scarce. Morris et al. reported that the activity of the rotator cuff muscles was similar in MDI, MDL, and healthy shoulders (3). However, increased activity of the posterior deltoid was observed in MDL, which was explained as a compensatory strategy for the increased shoulder laxity (3).

Nevertheless, current findings of muscle activity and scapular kinematics in MDI and MDL might not be fully transferable to hEDS/HSD patients for several reasons. First, both myotendinous as well as myofascial force transmission might be disturbed in EDS, as multiple studies have shown evidence for an abnormal connective tissue composition of the muscle's extracellular matrix (13), which might influence electromyography (EMG) measurements (14). Furthermore, evidence exists for the involvement of central sensitization in hEDS/HSD patients, contributing to widespread, chronic pain (15). Therefore, pain might negatively interfere with muscle performance. Additionally, fear of pain might trigger patients to avoid painful muscle contractions, possibly affecting movement as well (16).

Therefore, we wanted to examine whether changes in muscle activity and scapular motion occur in patients with hEDS/HSD, and if so, whether these changes are comparable with alterations seen in MDI or MDL. The purpose of this study was to determine differences in EMG activity and 3-dimensional (3D) scapular kinematics during elevation in the scapular plane between hEDS/HSD patients with MDI, participants with MDL, and healthy controls. We hypothesized that muscle activity and scapular kinematics are altered in hEDS/HSD patients with MDI compared to both MDL subjects and healthy controls and that these changes may be present to a lesser extent in MDL subjects.

MATERIALS AND METHODS

Subjects. Three groups of female participants were recruited: 27 patients with MDI, who had previously been diagnosed with hEDS (n = 14) or HSD (n = 13); 28 participants with MDL, and 27 healthy controls. Patients with hEDS/HSD were recruited at the Center for Medical Genetics of Ghent University Hospital. All patients were diagnosed with hEDS or generalized HSD, according to the classification of the international EDS consortium (6,7). For this study, only hEDS/HSD patients with MDI were selected. MDI was diagnosed based on the following inclusion criteria (17): 1) shoulder pain for at least 3 months prior to the study; 2) symptoms of shoulder instability (e.g., involuntary recurrent subluxations/dislocations, apprehensive muscle tension, sensation of shoulder giving way) in daily life, without a traumatic onset; and 3) shoulder laxity in at least 2 directions confirmed on clinical examination (Table 1). Additionally, participants had to be able to elevate their arm up to 120° without shoulder luxation or subluxation. Patients were excluded if they had a history of frozen shoulder or shoulder surgery.

Healthy controls and MDL subjects were volunteers recruited from the local community via advertisement on social media. Controls and MDL participants were excluded if they had symptoms of shoulder instability, shoulder pain, any generalized disease diagnosed by a medical doctor (e.g., hEDS, HSD, diabetes

Table 1. Shoulder laxity tests

Test	Direction of laxity	Cutoff value
Sulcus sign	Inferior	>2 cm (44)
Gagey hyperabduction test	Inferior	>105° (45)
Anterior/posterior load and shift	Anterior/ posterior	Grade 2 or 3 (46)
Passive external rotation in supine	Anterior	>90° (47)
Active external rotation in standing	Anterior	>85° (47)
Posterior jerk	Posterior	Positive if a clunk is felt (48)

mellitus, multiple sclerosis, etc.), a history of shoulder pathology (confirmed on imaging), if they used any medication (except for contraceptives), or if they participated in overhead sports for >3 hours per week. In addition, controls were excluded if they had MDL or generalized joint hypermobility (Beighton score ≥4 of 9) (7). Participants were included in the MDL group if they scored positive on shoulder laxity tests in at least 2 directions (Table 1) but had no shoulder complaints that interfered with daily life activities. All clinical shoulder examinations were performed by the same team of physical therapists (VS and IDW) with knowledge of hEDS, HSD, and MDI. Only women were recruited for this study, given the large predominance of joint hypermobility conditions in females (14,18). This study was approved by the Ethics Committee of Ghent University Hospital. An informed consent was obtained from all participants.

Instrumentation. Surface EMG data were collected from the upper trapezius, middle trapezius, lower trapezius, infraspinatus, latissimus dorsi, serratus anterior, posterior deltoid, and pectoralis major (sternal part) using the Noraxon Ultium ESP System. Bipolar surface electrodes (Ag/AgCl; Ambu BlueSensor N, type N-00-S/25; 30 × 22 mm) were placed with an interelectrode distance of 1 cm, in line with the muscle fibers. Prior to electrode placement, the skin surface was shaved, scrubbed, and degreased with alcohol to minimize skin impedance. Surface electrodes were placed according to the Surface Electromyography for the Non-Invasive Assessment of Muscles guidelines on the posterior deltoid and the upper, middle, and lower trapezius. Electrode placement for the other muscles is described in Table 2.

3D kinematic data were obtained using 8 Oqus Qualisys cameras (6 Oqus 3+ and 2 Oqus 4, 200Hz). After calibration, spherical infrared reflective markers (12.5-mm lightweight markers) were placed on 34 anatomical landmarks, based on International Society of Biomechanics recommendations (19). Simultaneous collection of EMG and 3D kinematic data was possible using the Qualisys Track Manager Software (QTM 2019.1).

Table 2. Electromyography electrode placement*

Muscle	Electrode placement
IS	2.5 to 3 cm inferior to the scapular spine, between the LT and PD. The borders of LT and PD were palpated during contraction to ensure correct electrode placement (11)
LD	3 cm lateral and inferior to the inferior angle of the scapula (49)
SA	At the seventh rib, below the axilla, anterior to the LD and posterior to the PM (50)
PM	One-third medially of the greater tuberosity, on an imaginary line determined by the greater tuberosity and the xiphoid process (51)

* IS = infraspinatus; LD = latissimus dorsi; LT = lower trapezius; PD = posterior deltoid; PM = pectoralis major; SA = serratus anterior.

Testing procedure. The shoulders of all participants were clinically examined. First, shoulder laxity tests were evaluated (Table 1). Next, the apprehension and relocation tests were performed to evaluate symptoms of shoulder instability, such as subluxation or apprehensive muscle tension (20). Lastly, the Beighton score was evaluated.

After clinical examination, EMG electrodes were placed. To allow for relevant EMG comparisons between subjects and studies, it is recommended to normalize EMG activity to a maximum voluntary isometric contraction (MVIC) (21). Most hEDS/HSD patients were unable to perform the previously described MVICs (22), since many of these positions are challenging for unstable shoulders and provoke pain symptoms and apprehension. Consequently, for this study, a set of 4 alternative MVICs was performed in random order: isometric shoulder abduction, adduction, external rotation, and internal rotation, with the humerus positioned next to the thorax to minimize the risk of dislocating. Each MVIC was maintained for 5 seconds and repeated 3 times. Participants were verbally encouraged to achieve maximal effort during all trials. Shoulder pain was rated on a 10-cm visual analog scale (VAS; 0 = no pain and 10 = worst pain) just before and during the MVIC tests.

Next, all participants performed 5 repetitions of elevation in the scapular plane (30° anterior to the frontal plane). Elevation in the scapular plane was chosen since it is a functional movement frequently studied in patients with MDI (10,23). First, a visual demonstration was given by the investigator. Subsequently, each participant was instructed to elevate the arm in the scapular plane up to 120° to a count of 8 seconds: 4 seconds to elevate and 4 seconds to lower the arm. Each repetition was preceded and followed by 5 seconds of rest. To ensure standardization, external feedback was provided by a stick with marked limits that indicated the position of the scapular plane and the end range of motion (120°). Verbal and visual feedback was provided by the investigator (VS). Additionally, a metronome was used to control movement velocity (60 beats/minute).

Signal processing and data analysis. For the signal processing of the 3D-kinematic data, anatomical markers were digitally labeled within the Qualisys Track Manager interface. Subsequently, the labeled c3d files were imported in Visual 3D (v6.01.36, C-motion Inc.). Data were filtered using a fourth-order Butterworth low-pass filter at 6Hz (25), and joint angles were calculated based on the Upper Limb Pelvis Model 2018, which uses an inverse kinematics constraint model based on ISB recommendations of Wu et al (19,25).

Surface EMG signals were first processed using the MyoResearch 3 3.14.16 (Noraxon) software program. Electrocardiogram reduction was performed, followed by full-wave rectification and signal smoothing (root mean square, 100-millisecond window). For each muscle and MVIC test, the mean amplitude of each

trial was calculated using a 3-second analysis window (26). Next, the mean amplitudes of each trial were averaged. As a result, 4 averaged mean amplitudes for each muscle were recorded. The MVIC value (100%) was defined as the highest averaged mean amplitude for that muscle obtained during 1 of the 4 MVIC tests.

To perform the statistical parametric mapping (SPM) analysis, all data were exported to Matlab (Matlab R2016b; 9.1.0.441655). Subsequently, data were time normalized within Matlab to 101 data frames. This procedure means that at 50% of the elevation movement, the highest elevation angle (120°) was reached.

Statistical analysis. Demographic characteristics were analyzed using IBM SPSS statistics software, version 25.0. For each muscle, intraclass correlation coefficients (ICCs; 2-way random, absolute agreement) were calculated over the 3 trials performed to determine the trial-to-trial reliability of MVIC measurements.

To statistically compare scapular kinematics and EMG data between groups, a curve analysis was conducted using SPM (27). All SPM analyses were performed using the open-source spm1d code (www.spm1d.org) in Matlab. Specifically, a 1-way independent analysis of variance (ANOVA), SPM(F), was used to compare scapular kinematics and muscle activity between groups. If statistical significance was reached in the ANOVA, post hoc analyses were done using 2-sample *t*-tests of SPM, conducted on all group pairs. The Bonferroni correction was used to correct for multiple comparisons across the 3 groups. First, the SPM(F) curve was computed for the complete time series. Second, the temporal smoothness of the data was estimated based on the average temporal gradient. Third, to test our null hypothesis (no differences in scapular kinematics and EMG activity between groups), the critical threshold of SPM(F) was calculated, above which only 5% of the data (alpha) would be expected to traverse in case of an equally smooth random process. This method is based on the random field theory (28). Due to interdependence of neighboring points and waveform smoothness, multiple adjacent points of the SPM(F) curve may exceed the critical threshold, forming a suprathreshold cluster. Fourth, cluster-specific *P* values were calculated for each suprathreshold cluster. For all analyses, alpha was set at 0.05. This statistical method has been previously described in detail elsewhere (27,28) (www.spm1d.org).

RESULTS

Demographic characteristics. Demographic characteristics are shown in Table 3. No statistically significant differences between groups were found, except for age and Beighton score.

Scapular kinematics. Statistical analysis showed significantly less posterior tilt in the hEDS/HSD group compared to the MDL group during the descending phase of elevation (60–79%; $P = 0.02$) (Figure 1). Furthermore, hEDS/HSD patients demonstrated significantly less upward rotation compared to both controls and MDL participants during the first half of the ascending (0–25% [$P = 0.005$] and 0–7.5% [$P = 0.04$], respectively) and the final part of the descending phase of elevation (65–100% [$P < 0.001$] and 77–100% [$P = 0.01$], respectively) (Figure 2). Regarding scapular internal/external rotation, no significant group differences were found.

Surface EMG. ICCs for trial-to-trial reliability of EMG activity during the MVICs varied from 0.961 to 0.991, showing excellent reliability of MVIC measurements (29). ANOVA analysis in SPM identified significant group differences in EMG activity during elevation in the scapular plane. An overview of the significant results of the post hoc analyses is available in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24525/abstract. Mean trajectories with SD clouds for all muscles are shown in Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24525/abstract.

hEDS/HSD patients versus controls. Patients showed higher infraspinatus activity during the full range of motion. In addition, higher activity of the middle trapezius was observed in patients at different stages of elevation. During the final part of the descending phase, patients showed higher activity of the lower trapezius and posterior deltoid. During the first part of the ascending phase, higher latissimus dorsi activity was observed in patients. No significant differences between patients and controls were

Table 3. Demographic characteristics*

Variable	hEDS/HSD patients (n = 27)	Healthy controls (n = 27)	MDL subjects (n = 28)
Age, years†	34.7 ± 13.22	31.9 ± 9.61	24.2 ± 3.49
Height, cm	168.5 ± 7.64	170.0 ± 5.34	169.9 ± 4.34
Body mass, kg	64.9 ± 13.40	63.0 ± 8.18	60.0 ± 5.75
Beighton score‡	5.7 ± 2.14	1.5 ± 1.48	3.1 ± 2.13

* Values are the mean ± SD. hEDS = hypermobile Ehlers-Danlos syndrome; HSD = hypermobility spectrum disorder; MDL = multidirectional shoulder laxity.
† The MDL group was significantly younger than the hEDS/HSD ($P < 0.001$) and control groups ($P = 0.01$).
‡ The hEDS/HSD group had a significantly higher Beighton score than the MDL ($P < 0.001$) and control ($P < 0.001$) groups. Moreover, the MDL group had a significantly higher Beighton score than the control group ($P = 0.008$).

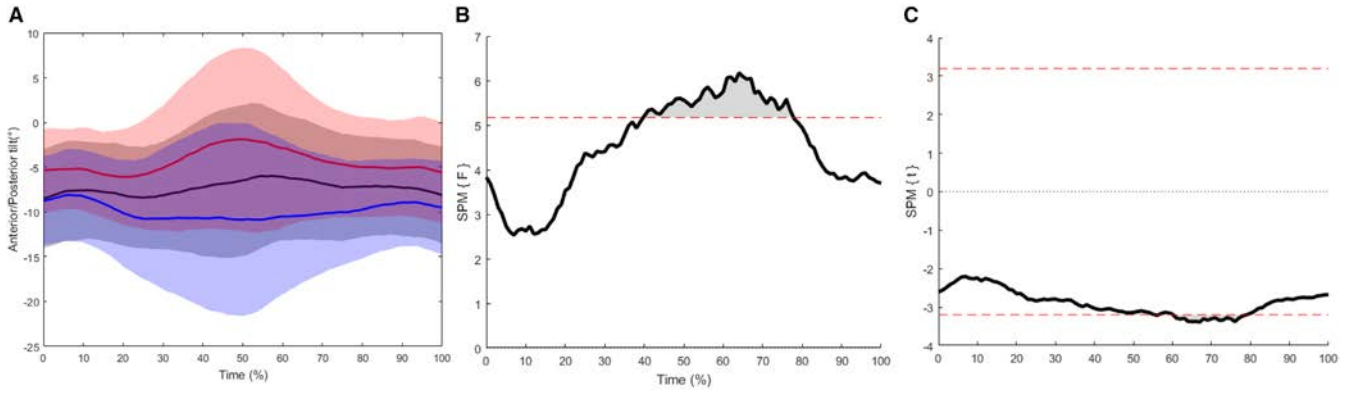

Figure 1. **A**, Mean trajectories with SD clouds for anterior/posterior tilt in controls (black), multidirectional shoulder laxity (MDL) participants (red), and hypermobile Ehlers-Danlos syndrome (hEDS)/hypermobility spectrum disorder (HSD) patients (blue); **B**, Statistical parametric mapping (SPM) 1-way independent analysis of variance test statistic (SPM[F]). The critical threshold of 5.173 (red broken line) was exceeded at time 40–78% ($P = 0.001$), indicating a significant group difference; **C**, Post hoc SPM *t*-test (SPM[t]) comparing MDL subjects with hEDS/HSD patients. The critical threshold of –3.2 (red broken line) was exceeded at time 60–79% ($P = 0.02$).

found for the serratus anterior, pectoralis major, and upper trapezius.

hEDS/HSD patients versus MDL participants. Patients had an overall higher infraspinatus, middle trapezius, and posterior deltoid activity throughout the greater part of elevation. During the first part of the ascending phase, higher activity of latissimus dorsi, serratus anterior, and pectoralis major was found in patients. No significant differences between patients and MDL participants were found for the upper and lower trapezius. Differences in EMG activity between MDL participants and controls did not reach statistical significance.

DISCUSSION

This is the first study to identify differences in both EMG activity and 3D scapular kinematics during elevation in the scapular plane between 1) patients with hEDS/HSD and MDI, 2) MDL participants, and 3) healthy controls, using SPM. hEDS/HSD patients

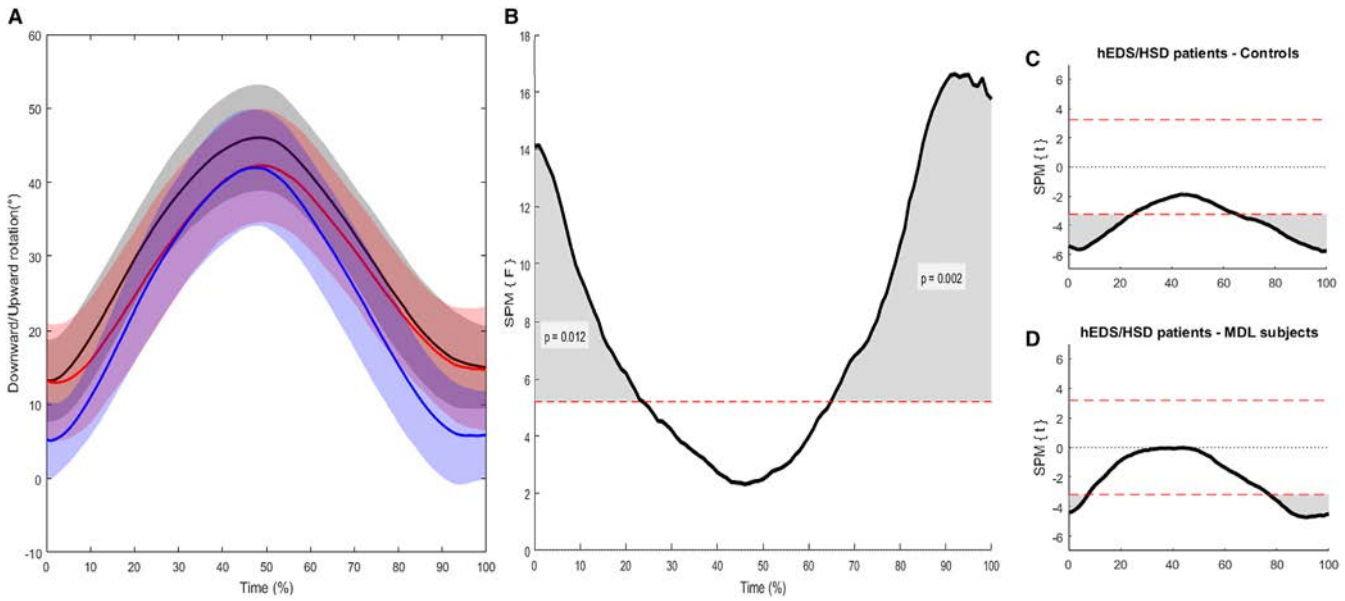

Figure 2. **A**, Mean trajectories with SD clouds for downward/upward rotation in controls (black), multidirectional shoulder laxity (MDL) participants (red), and hypermobile Ehlers-Danlos syndrome (hEDS)/hypermobility spectrum disorder (HSD) patients (blue); **B**, Statistical parametric mapping (SPM) 1-way independent analysis of variance test statistic (SPM[F]). The critical threshold of 5.195 (red broken line) was exceeded at time 0–24% ($P = 0.012$) and 65–100% ($P = 0.002$), indicating a significant group difference; **C**, Post hoc SPM *t*-test (SPM[t]) comparing hEDS/HSD patients with healthy controls. Two suprathreshold clusters exceeded the critical threshold of –3.23 (red broken line) at time 0–25% ($P = 0.005$) and 65–100% ($P < 0.001$); **D**, Post hoc SPM(t) comparing hEDS/HSD patients with MDL subjects. Two suprathreshold clusters exceeded the critical threshold of –3.18 (red broken line) at time 0–7.5% ($P = 0.04$) and 77–100% ($P = 0.01$).

demonstrated significantly less scapular upward rotation compared to both other groups. Additionally, significantly less posterior tilt was seen in hEDS/HSD patients compared to MDL participants. Finally, significantly higher EMG activity of the infraspinatus, middle trapezius, and posterior deltoid was found in hEDS/HSD patients.

Our finding that hEDS/HSD patients with MDI demonstrate less scapular upward rotation during elevation in the scapular plane aligns with other studies investigating MDI patients (5,12,30). However, while some studies (5,30) have reported a decreased upward rotation in higher angles of elevation, we identified less upward rotation in the first half of the ascending and the final part of the descending phase of elevation. Nevertheless, our finding is in good agreement with the clinical presentation of MDI, characterized by instability symptoms that usually arise in the mid ranges of motion (31). Less upward rotation leads to less inclination of the glenoid cavity, which consecutively facilitates caudal translation, or even subluxation, of the humeral head (32).

Our result of a decreased posterior tilt in hEDS/HSD patients with MDI deviates from the result of Ogston and Ludewig (5), who found significantly more posterior tilt in MDI patients, but our result is in accordance with other studies that have reported a decreased posterior tilt in patients with impingement (33). We must emphasize that the difference in posterior tilt in our study was only present between the MDI and MDL group. As shown in Figure 1, MDL subjects tend to have more posterior tilt than controls, albeit not significantly. Hypothetically, this tendency toward more posterior tilt in MDL subjects represents a positive compensation for increased capsuloligamentous laxity, which is lost in hEDS/HSD patients. While the elevation was painless in controls and MDL subjects, almost all hEDS/HSD patients reported shoulder pain while moving the arm. Therefore, pain or other symptoms typical for hEDS/HSD, such as fatigue, might play a role. However, this possibility is a hypothesis only, since we did not examine the extent to which these symptoms affect scapular kinematics or EMG activity.

Finally, we could not find significant group differences in scapular internal/external rotation. This finding deviates from other studies that have reported an increased internal rotation in MDI (5,12). These contrasting results may be due to high variability in scapular internal/external rotation across subjects (33). Nevertheless, both the reduced upward rotation and the impression of less posterior tilt argue for the presence of scapular dyskinesis in hEDS/HSD patients with MDI. Furthermore, mean between-group differences in upward rotation and posterior tilt are ≥5°. These differences are unlikely to be due to measurement error since these are greater than previously reported standard error of measurement values (34).

Interestingly, no significant differences in scapular kinematics were found between the MDL and healthy control groups. By contrast, MDL subjects did significantly differ from MDI patients with respect to upward rotation and posterior tilt. Since MDL is an asymptomatic condition while MDI is not, these results support the hypothesis that scapular dyskinesis in hEDS/HSD patients may be associated with factors other than capsuloligamentous laxity. In this view, whether instability symptoms are the cause or rather the result of abnormal scapular movement remains unknown (5,35). Therefore, scapular dyskinesis may also be a reaction to a painful shoulder condition and may emerge as an antalgic movement pattern (31). Since the patients in this study experienced a variety of symptoms that are all attributed to hEDS/HSD with MDI, identifying the exact reason for the altered scapular movement is difficult. Soft tissue fragility, muscle weakness, decreased endurance, fatigue, and/or chronic pain could all be possible underlying mechanisms that may elicit or worsen scapular dyskinesis. These underlying mechanisms should be primarily targeted in treatment. Nevertheless, we believe that scapular dyskinesis should also be addressed in the treatment of MDI, since it can easily aggravate existing symptoms and jeopardize normal shoulder function (31,35).

The result of increased infraspinatus activity in hEDS/HSD patients is in accordance with previous MDI research (10,36). Middle trapezius and posterior deltoid activity was found to be higher in hEDS/HSD patients as well. Since these muscles contribute to scapulothoracic (37) and glenohumeral stabilization (11), these results reinforce the hypothesis of an increased need for dynamic stabilization in patients with MDI (36).

Notably, mean between-group differences of infraspinatus, middle trapezius, and posterior deltoid activity are ≥10% of MVIC. These differences are likely to represent true differences, since they are greater than previously reported standard error of measurement values (38) and the previous reported cutoff (10%) in terms of muscle strengthening purposes (39).

Since no differences in EMG activity were found between MDL subjects and controls, we believe the following reasons might account for the increased muscle activity in hEDS/HSD patients with MDI. First, fear of injury, fatigue, and pain are identified as commonly reported barriers for hEDS patients to exercise (40). Possibly, hEDS/HSD patients experience apprehensive muscle tension because they are afraid of moving the unstable shoulder. Since MDL subjects did not experience shoulder pain during elevation, there was possibly no need for apprehension to occur. Second, both hEDS and HSD are known to be associated with muscle weakness and decreased muscle endurance (8,14). We hypothesize that the increased EMG amplitudes could also be a result of a compensatory mechanism in which muscle fibers attempt to maximize their force output, albeit unsuccessful due to the increased compliance of the tendons and disturbed composition of the extracellular matrix, which makes transmitting the force generated by the muscle difficult (13). Thus, the increased EMG amplitudes could by clarified by an additional motor unit recruitment and an increased firing frequency, both potential

strategies to overcome muscle weakness (41,42). Further research is necessary to confirm or refute these hypotheses.

Whereas Morris et al (3) stated that in MDL subjects, a muscular compensatory strategy is present that fails in subjects with MDI, we could not support this theory. In contrast, we hypothesize that hEDS/HSD patients with MDI are much more in need of a muscular compensatory strategy than subjects with MDL. However, in the study of Morris et al, none of the MDI subjects had hEDS or HSD. Therefore, comparison is difficult.

This study has several strengths. First, SPM allowed us to detect statistically significant differences over the entire elevation movement. Second, this research provided groundwork in investigating both scapular kinematics as well as EMG muscle activity in patients with hEDS/HSD and MDI.

Nevertheless, our findings should be interpreted in the light of the study limitations. First, since maximal muscle activation can cause discomfort and overestimation of muscle strength (21), caution is recommended when normalizing to an MVIC in patients with shoulder pain. To avoid overestimation of the percentage of muscle activation, MVIC tests should happen in a pain-free or reduced-pain condition (21). Despite our effort of using alternative MVICs, creating a pain-free condition was not always possible. However, we observed that in hEDS/HSD patients, baseline pain ratings (mean VAS baseline = 1.4) did not increase considerably during the MVIC (mean VAS during MVICs = 2.2), since the mean difference is smaller than the minimal detectable change for pain (34). We therefore presume that differences in normalized EMG activity between MDI patients and MDL/controls are not related to pain-related underperformance of the patients during the MVIC procedure. However, normalized EMG should still be interpreted with caution in patients with shoulder pain.

Second, MDL subjects were significantly younger than controls and hEDS/HSD patients. However, as joint laxity reduces with increasing age (43), the age of the MDL group may be representative for the study population. Third, only patients who were able to elevate their arm up to 120° without luxation or subluxation were included. This limitation could induce a selection bias toward patients at the better end of the hypermobility spectrum. Therefore, we should be careful with generalizing our results to all MDI patients. Fourth, hEDS and HSD patients were seen as 1 group, since musculoskeletal manifestations of hEDS and HSD largely overlap (6). However, this grouping may have led to a substantial within-group heterogeneity. Future research could implement cluster analysis, preferrably in a large patient cohort, to unravel whether subgroups exist within the hEDS/HSD population, exhibiting different muscle patterns that lead to insufficient scapular upward rotation.

In conclusion, hEDS/HSD patients with MDI demonstrate altered scapular kinematics and EMG muscle activity. These findings extend previous research and could serve as a stepping stone for developing treatment strategies in patients whose conditions belong to the hypermobility spectrum.

ACKNOWLEDGMENT

We express our gratitude toward all study participants.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Ms. Spanhove 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. Spanhove, Berckmans, Cools, De Wandele.
Acquisition of data. Spanhove, Palmans, Malfait, Cools, De Wandele.
Analysis and interpretation of data. Spanhove, Calders, Palmans, Cools, De Wandele.

REFERENCES

1. Best MJ, Tanaka MJ. Multidirectional instability of the shoulder: treatment options and considerations. Sports Med Arthrosc Rev 2018; 26:113–9.
2. Navlet MG, Asenjo-Gismero CV. Multidirectional instability: natural history and evaluation. Open Orthop J 2017;11:861–74.
3. Morris AD, Kemp GJ, Frostick SP. Shoulder electromyography in multidirectional instability. J Shoulder Elbow Surg 2004;13:24–9.
4. Bahu MJ, Trentacosta N, Vorys GC, Covey AS, Ahmad CS. Multidirectional instability: evaluation and treatment options. Clin Sports Med 2008;27:671–89.
5. Ogston JB, Ludewig PM. Differences in 3-dimensional shoulder kinematics between persons with multidirectional instability and asymptomatic controls. Am J Sports Med 2007;35:1361–70.
6. Castori M, Tinkle B, Levy H, Grahame R, Malfait F, Hakim A. A framework for the classification of joint hypermobility and related conditions. Am J Med Genet C Semin Med Genet 2017;175:148–57.
7. Malfait F, Francomano C, Byers P, Belmont J, Berglund B, Black J, et al. The 2017 international classification of the Ehlers-Danlos syndromes. Am J Med Genet C Semin Med Genet 2017;175:8–26.
8. Rombaut L, Deane J, Simmonds J, De Wandele I, De Paepe A, Malfait F, et al. Knowledge, assessment, and management of adults with joint hypermobility syndrome/Ehlers-Danlos syndrome hypermobility type among Flemish physiotherapists. Am J Med Genet C Semin Med Genet 2015;169c:76–83.
9. Auge WK II, Morrison DS. Assessment of the infraspinatus spinal stretch reflex in the normal, athletic, and multidirectionally unstable shoulder. Am J Sports Med 2000;28:206–13.
10. Illyes A, Kiss J, Kiss RM. Electromyographic analysis during pull, forward punch, elevation and overhead throw after conservative treatment or capsular shift at patient with multidirectional shoulder joint instability. J Electromyogr Kinesiol 2009;19:e438–47.
11. Barden JM, Balyk R, Raso VJ, Moreau M, Bagnall K. Atypical shoulder muscle activation in multidirectional instability. Clin Neurophysiol 2005;116:1846–57.
12. Von Eisenhart-Rothe R, Mayr HO, Hinterwimmer S, Graichen H. Simultaneous 3D assessment of glenohumeral shape, humeral head centering, and scapular positioning in atraumatic shoulder instability: a magnetic resonance-based in vivo analysis. Am J Sports Med 2010;38:375–82.
13. Voermans NC, Altenburg TM, Hamel BC, de Haan A, van Engelen BG. Reduced quantitative muscle function in tenascin-X deficient Ehlers-Danlos patients. Neuromuscul Disord 2007;17:597–602.
14. Rombaut L, Malfait F, De Wandele I, Mahieu N, Thijs Y, Segers P, et al. Muscle–tendon tissue properties in the hypermobility type of Ehlers-Danlos syndrome. Arthritis Care Res (Hoboken) 2012;64: 766–72.

15. Di Stefano G, Celletti C, Baron R, Castori M, Di Franco M, La Cesa S, et al. Central sensitization as the mechanism underlying pain in joint hypermobility syndrome/Ehlers-Danlos syndrome, hypermobility type. Eur J Pain 2016;20:1319–25.
16. Scheper MC, de Vries JE, Verbunt J, Engelbert RH. Chronic pain in hypermobility syndrome and Ehlers-Danlos syndrome (hypermobility type): it is a challenge. J Pain Res 2015;8:591–601.
17. Spanhove V, De Wandele I, Kjær BH, Malfait F, Vanderstukken F, Cools A. The effect of five isometric exercises on glenohumeral translations in healthy subjects and patients with the hypermobility type of the Ehlers-Danlos syndrome (heds) or hypermobility spectrum disorder (hsd) with multidirectional shoulder instability: an observational study. Physiotherapy 2019;107:11–8.
18. Tinkle B, Castori M, Berglund B, Cohen H, Grahame R, Kazkaz H, et al. Hypermobile Ehlers-Danlos syndrome (a.k.a. Ehlers-Danlos syndrome type III and Ehlers-Danlos syndrome hypermobility type): clinical description and natural history. Am J Med Genet C Semin Med Genet 2017;175:48–69.
19. Wu G, van der Helm FC, Veeger HE, Makhsous M, Van Roy P, Anglin C, et al. ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion. Part II: shoulder, elbow, wrist and hand. J Biomech 2005;38:981–92.
20. Cools AM, Cambier D, Witvrouw EE. Screening the athlete's shoulder for impingement symptoms: a clinical reasoning algorithm for early detection of shoulder pathology. Br J Sports Med 2008;42:628–35.
21. Ettinger L, Weiss J, Shapiro M, Karduna A. Normalization to maximal voluntary contraction is influenced by subacromial pain. J Appl Biomech 2016;32:433–40.
22. Castelein B, Cagnie B, Parlevliet T, Danneels L, Cools A. Optimal normalization tests for muscle activation of the levator scapulae, pectoralis minor, and rhomboid major: an electromyography study using maximum voluntary isometric contractions. Arch Phys Med Rehabil 2015;96:1820–7.
23. Ellenbecker TS, Davies GJ. The application of isokinetics in testing and rehabilitation of the shoulder complex. J Athl Train 2000;35:338–50.
24. Robertson DG, Dowling JJ. Design and responses of Butterworth and critically damped digital filters. J Electromyogr Kinesiol 2003;13:569–73.
25. Palmans T, Kepple T. The Ghent University upper limb and pelvis model. Ghent: Ghent University Press; 2018.
26. Lopes AD, Timmons MK, Grover M, Ciconelli RM, Michener LA. Visual scapular dyskinesis: kinematics and muscle activity alterations in patients with subacromial impingement syndrome. Arch Phys Med Rehabil 2015;96:298–306.
27. Pataky TC. Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. J Biomech 2010;43:1976–82.
28. Pataky T. rft1d: Smooth one-dimensional random field upcrossing probabilities in Python. J Stat Softw 2016;71:1–22.
29. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–63.
30. Ozaki J. Glenohumeral movements of the involuntary inferior and multidirectional instability. Clin Orthop Relat Res 1989;238:107–11.
31. Kibler WB, Sciascia A. The role of the scapula in preventing and treating shoulder instability. Knee Surg Sports Traumatol Arthrosc 2016;24:390–7.
32. Itoi E. Pathophysiology and treatment of atraumatic instability of the shoulder. J Orthop Sci 2004;9:208–13.
33. Ludewig PM, Reynolds JF. The association of scapular kinematics and glenohumeral joint pathologies. J Orthop Sports Phys Ther 2009;39:90–104.
34. Michener LA, Elmore KA, Darter BJ, Timmons MK. Biomechanical measures in participants with shoulder pain: intra-rater reliability. Man Ther 2016;22:86–93.
35. Cools AM, Struyf F, De Mey K, Maenhout A, Castelein B, Cagnie B. Rehabilitation of scapular dyskinesis: from the office worker to the elite overhead athlete. Br J Sports Med 2014;48:692–7.
36. Illyes A, Kiss RM. Kinematic and muscle activity characteristics of multidirectional shoulder joint instability during elevation. Knee Surg Sports Traumatol Arthrosc 2006;14:673–85.
37. Contemori S, Panichi R, Biscarini A. Effects of scapular retraction/protraction position and scapular elevation on shoulder girdle muscle activity during glenohumeral abduction. Hum Mov Sci 2019;64:55–66.
38. Seitz AL, Uhl TL. Reliability and minimal detectable change in scapulothoracic neuromuscular activity. J Electromyogr Kinesiol 2012;22:968–74.
39. Andersen LL, Andersen CH, Mortensen OS, Poulsen OM, Bjørnlund IB, Zebis MK. Muscle activation and perceived loading during rehabilitation exercises: comparison of dumbbells and elastic resistance. Phys Ther 2010;90:538–49.
40. Simmonds JV, Herbland A, Hakim A, Ninis N, Lever W, Aziz Q, et al. Exercise beliefs and behaviours of individuals with joint hypermobility syndrome/Ehlers-Danlos syndrome: hypermobility type. Disabil Rehabil 2019;41:445–55.
41. Robbins SM, Cossette-Levasseur M, Kikuchi K, Sarjeant J, Shiu YG, Azar C, et al. Neuromuscular activation differences during gait in patients with Ehlers-Danlos syndrome and healthy adults. Arthritis Care Res (Hoboken) 2020;72:1653–62.
42. Hug F, Tucker K. Surface electromyography to study muscle coordination. In: Müller B, Wolf SI, Brueggemann GP, Deng Z, McIntosh A, Miller F, et al, editors. Handbook of human motion. New York: Springer International Publishing; 2017. p. 1–21.
43. Johnson SM, Robinson CM. Shoulder instability in patients with joint hyperlaxity. J Bone Joint Surg Am 2010;92:1545–57.
44. Tzannes A, Paxinos A, Callanan M, Murrell GA. An assessment of the interexaminer reliability of tests for shoulder instability. J Shoulder Elbow Surg 2004;13:18–23.
45. Gagey OJ, Gagey N. The hyperabduction test. J Bone Joint Surg Br 2001;83:69–74.
46. Eshoj H, Ingwersen KG, Larsen CM, Kjaer BH, Juul-Kristensen B. Intertester reliability of clinical shoulder instability and laxity tests in subjects with and without self-reported shoulder problems. BMJ Open 2018;8:e018472.
47. Whitehead NA, Mohammed KD, Fulcher ML. Does the Beighton score correlate with specific measures of shoulder joint laxity? Orthop J Sports Med 2018;6:2325967118770633.
48. Tennent TD, Beach WR, Meyers JF. A review of the special tests associated with shoulder examination. Part II: laxity, instability, and superior labral anterior and posterior (SLAP) lesions. Am J Sports Med 2003;31:301–7.
49. Danneels LA, Vanderstraeten GG, Cambier DC, Witvrouw EE, Stevens VK, De Cuyper HJ. A functional subdivision of hip, abdominal, and back muscles during asymmetric lifting. Spine 2001;26:E114–21.
50. Ludewig PM, Hoff MS, Osowski EE, Meschke SA, Rundquist PJ. Relative balance of serratus anterior and upper trapezius muscle activity during push-up exercises. Am J Sports Med 2004;32:484–93.
51. Pontillo M, Orishimo KF, Kremenic IJ, McHugh MP, Mullaney MJ, Tyler TF. Shoulder musculature activity and stabilization during upper extremity weight-bearing activities. N Am J Sports Phys Ther 2007;2:90–6.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 841–848
DOI 10.1002/acr.24526

American College of Rheumatology
Empowering Rheumatology Professionals

Altered Multisegment Ankle and Foot Kinematics During Gait in Patients With Hypermobile Ehlers-Danlos Syndrome/ Hypermobility Spectrum Disorder: A Case–Control Study

Stefan Vermeulen,[1] Sophie De Mits,[2] Roel De Ridder,[1] Patrick Calders,[1] Joris De Schepper,[3] Fransiska Malfait,[4] and Lies Rombaut[4]

Objective. Ankle-foot problems have a considerable impact on daily functioning in patients with hypermobile Ehlers-Danlos syndrome and hypermobility spectrum disorder (hEDS/HSD). Therefore, the objective of this study was to identify alterations in multisegment ankle and foot kinematics during gait and to assess foot function and pain in these patients.

Methods. Twenty-three women with hEDS/HSD and 23 healthy controls participated in this 3-dimensional gait analysis. Multisegment ankle and foot kinematics were collected using the Ghent Foot Model and analyzed with Statistical Parametric Mapping. Foot function and pain were assessed using visual analog scale scores, the Margolis Pain Diagram, and the Foot Function Index.

Results. Levels of pain and foot dysfunction were significantly higher in subjects with hEDS/HSD ($P < 0.001$). Kinematic curve analyses provide evidence for a hypermobile first ray, represented by a significantly increased eversion position of the medial forefoot during stance phase ($P < 0.001$) in subjects with hEDS/HSD compared to controls. In addition, significantly more dorsiflexion was found in the medial and lateral forefoot and the rearfoot ($P < 0.001$). At the midfoot, an increased plantar flexion ($P < 0.001$) and at the level of the hallux a decreased dorsiflexion ($P = 0.037$) and increased inversion ($P < 0.001$) and abduction ($P = 0.016$) were found in subjects with hEDS/HSD.

Conclusion. This study is the first to apply a multisegment foot model during gait in hEDS/HSD, which confirms the characteristic hypermobility throughout the foot, especially the hypermobile first ray.

INTRODUCTION

Ehlers-Danlos syndrome (EDS) comprises a heterogeneous group of hereditary connective tissue disorders caused by different defects in an array of genes with diverse biologic functions linked to collagen biosynthesis, assembly, and organization. EDS is characterized by skin hyperextensibility and fragility, joint hypermobility, and generalized connective tissue fragility (1). The latest EDS classification recognizes 13 distinct EDS types caused by defects in 19 different genes (1). However, the genetic basis of hypermobile EDS (hEDS), which is the most common type, is still unknown (1). Consequently, hEDS remains a clinically established diagnosis based on a combination of criteria, i.e., generalized joint hypermobility, multiple signs of extraarticular soft tissue fragility, positive family history, recurrent joint luxations or subluxations or joint instability, and musculoskeletal pain (1). Patients who do not fulfill the criteria for hEDS in the new EDS nosology, which is more stringent than the previous Villefranche nosology, are now labeled as having hypermobility spectrum disorder (HSD) (2).

In addition to generalized joint hypermobility and recurrent joint dislocations, patients with hEDS/HSD present with multiple other musculoskeletal symptoms and problems. Many of the musculoskeletal complaints are located in the lower extremity, among which foot and ankle pain and ankle distortions are highly present (3). On a structural level, patients with hEDS/HSD seem to exhibit more foot anomalies, such as hallux valgus, claw toes, and pes planus (4–6). By contrast, pes cavus has also been reported (7). These foot alterations are part of a multifactorial process in which lower-limb joint hypermobility, lower-limb pain, lower-limb muscle weakness (8), and fear of falling (9,10) likely interact with

Dr. Malfait is a Senior Clinical Investigator at the Research Foundation Flanders (grant 1842318N).

[1]Stefan Vermeulen, MSc, Roel De Ridder, PhD, Patrick Calders, PhD: Ghent University, Ghent, Belgium; [2]Sophie De Mits, PhD: Ghent University Hospital and Artevelde University of Applied Sciences, Ghent, Belgium; [3]Joris De Schepper, MSc: Artevelde University of Applied Sciences, Ghent, Belgium; [4]Fransiska Malfait, PhD, MD, Lies Rombaut, PhD: Ghent University Hospital, Ghent, Belgium.

Drs. Vermeulen and De Mits contributed equally to this work.

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

Address correspondence to Lies Rombaut, PhD, Ghent University Hospital, Corneel Heymanslaan 10, Entrance 81, 9000 Ghent, Belgium. Email: Lies.rombaut@ugent.be.

Submitted for publication May 12, 2020; accepted in revised form December 1, 2020.

SIGNIFICANCE & INNOVATIONS

- This study is the first to use a multisegment foot model to provide better insight into the foot kinematics of the different parts of the foot in patients with hypermobile Ehlers-Danlos syndrome and hypermobility spectrum disorder (hEDS/HSD).
- This study demonstrates that hypermobility is observed in the smaller parts of the foot, specifically the first ray, influencing motion pattern of different foot segments during gait in hEDS/HSD.
- Since the multisegment foot model demonstrated hypermobility throughout the whole foot, therapy such as custom-made foot orthoses can be more specifically targeted based on the current findings in the fore- and midfoot.

each other. Moreover, patients with hEDS/HSD reported walking and bending as the activity of daily life with the greatest physical impairment (11). Since the foot acts as the only point of contact during walking, investigating foot and ankle gait kinematics seems essential to better understand ankle and foot pain and disfunction in this population.

Until now, several studies have investigated gait in patients with hEDS/HSD (9,10,12–16). Patients have been shown to walk with a reduced gait velocity, step length, and/or stride length (9,10,12,13,15,16). In addition, hEDS individuals show altered kinematics of the pelvis, hip, knee, and ankle during the different phases of gait (9,12,13,14,16). However, for the latter, current findings are limited to the foot modeled as 1 single rigid segment, oversimplifying the complexity of the ankle-foot complex or even contradicting true ankle joint kinematics (17). Foot kinematics in patients with hEDS/HSD remain poorly understood due to the 1-segment approach, whereas the use of multisegment foot models, acknowledging the functional entities within the foot, are emerging in other patient populations, e.g., with psoriatic arthritis (18). Moreover, during 50% of stance phase (terminal stance and pre-swing), the rearfoot is not in contact with the ground anymore, emphasizing the importance of the forefoot in gait biomechanics (19). Consequently, there is a need to obtain a better insight into the biomechanics of the ankle-foot complex of hEDS/HSD in order to develop evidence-based therapy guidelines for this patient group.

Therefore, the main objective of this study was to explore the 3-dimensional (3D) foot and ankle kinematics of hEDS/HSD patients in comparison with healthy controls using the 6-segment Ghent Foot Model. The secondary objective was to evaluate foot function and foot pain in these patients.

PATIENTS AND METHODS

Patients. This study protocol was reviewed and approved by the Ethical Committee of Ghent University Hospital, and written, informed consent was obtained from all participants. As >90% of the hEDS/HSD patients are female (20), the study included only women. After a routine visit at the Center for Medical Genetics at the Ghent University Hospital, consecutive patients who fulfilled the inclusion criteria for the study were informed verbally and in written form and invited to participate in the study. Patient selection was performed on the basis of Villefranche criteria for EDS hypermobility type (at that time abbreviated as EDS-HT) because recruitment started prior to publication of the latest EDS nosology. Afterward, during a routine follow-up, all patients have been relabeled according to the current 2017 EDS criteria into hEDS/HSD (by clinical geneticist FM). After all patients participated, healthy volunteers were recruited by oral inquiry. Since hypermobility decreases with age, is more prevalent in women, and differs between ethnic groups, the control subjects were individually matched for sex, age, and ethnicity. Only controls with no foot complaints, injuries to the lower limbs, or generalized joint hypermobility by the Beighton scale were included. The Beighton scale consists of 4 bilateral tests, i.e., elbow hyperextension >10°, knee hyperextension >10°, first finger opposition, fifth finger extension >90°, and forward bending with hands flat on the floor. A score of ≥5 of 9 indicates the presence of generalized joint hypermobility (21). Surgery of the lower limb or trunk, a systemic disease with potential impact on feet and/or gait pattern (e.g., diabetes mellitus or rheumatic disorders) and pregnancy or delivery in the last year were exclusion criteria for all subjects. In addition, patients were excluded if they had an injury to the lower limb (currently or in the past 6 months) or if they were not able to walk barefoot and without a walking device for 5 meters. Since this is the first study that examines multisegment ankle and foot kinematics, no sample size calculation could be performed.

Procedure. Each participant was individually evaluated using a standardized protocol. First, subject characteristics including age, height, and weight were collected. Then, the Margolis Pain Diagram and the Foot Function Index (FFI) were completed, followed by the 3D gait analysis. Visual analog scale (VAS) scores were gathered both before and after gait analysis.

The Margolis Pain Diagram objectifies foot function and pain. Subjects have to shade on 2 body outlines (front and back) the body parts where they felt pain lasting for >24 hours in the past 4 weeks. A weighted final score reflects the total percentage of body surface shaded as painful (22). Foot and lower-limb pain intensity were assessed just before and after gait analysis by VAS scores, where 0 = no pain and 10 = unbearable pain. The FFI was used to measure foot pain, disability, and functional limitations due to foot pathology (23). This foot-specific questionnaire consists of 23 items (range 0–100),

Table 1. Foot function scores and pain scores*

	hEDS/HSD	Control	P†
FFI			
Total	38.8 ± 16.44	10.1 ± 1.08	<0.001
Pain	44.7 ± 18.45	10.0 ± 2.33	<0.001
Physical disability	41.7 ± 20.98	10.2 ± 0.72	<0.001
Activity limitation	22.9 ± 14.51	10.0 ± 0.00	<0.001
VAS pre			
Lower limb	1.9 ± 2.11	0.0 ± 0.00	<0.001
Foot	3.5 ± 2.55	0.0 ± 0.00	<0.001
VAS post			
Lower limb	3.1 ± 2.69	0.0 ± 0.21	<0.001
Foot	5.0 ± 2.86	0.1 ± 0.52	<0.001

* Values are the mean ± SD unless indicated otherwise. hEDS = hypermobile Ehlers-Danlos syndrome; HSD = hypermobility spectrum disorder; FFI = Foot Function Index; VAS = visual analog scale for pain.
† All *P* values are significant.

with higher values corresponding to greater pain, disability, and activity limitation.

For the kinematic gait analysis, all participants walked barefoot at a self-selected comfortable walking speed over a 12-meter long instrumented walkway. Kinematic data were collected with an 8-camera optoelectronic system (Oqus 3, Qualysis) at 500 Hz, and a force plate (Accugait, AMTI, 1,000 Hz) was embedded in the walkway for stance phase event detection. The marker configuration for the multisegment foot model was in accordance with the Ghent Foot Model (24). This 6-segment foot model is defined by the shank, rearfoot, midfoot, medial forefoot, lateral forefoot, and the hallux. During the gait measurements, a midgait protocol was used. Participants were first allowed to familiarize themselves with the test procedure by performing a few practice trials, followed by the actual measurements. Trials were discarded if 2 feet touched the force plate, or if subjects were seen to show an adaptation in stride length in an attempt to target the force plate.

Data analysis and statistical analysis. Kinematic and kinetic data were processed in Qualisys (Qualisys Track Manager) and subsequently in Visual 3D (version 5, C-motion) software. Marker data were filtered using a fourth order Butterworth low-pass filter at 10 Hz. Euler rotations (X-Y-Z, representing respectively dorsi/plantar flexion [sagittal plane], eversion/inversion [frontal plane] and ab-/adduction [transverse plane]) were used to calculate 3D multisegment foot joint kinematics. Stance phase was determined using the vertical component of the ground reaction force with a threshold set at 10 N. Afterward, each point in the time series was normalized to 100%, and for each individual an average joint curve, combining 3 trials of the left and right foot, was exported for curve analysis.

The analysis of anthropometric data and foot function and pain was performed using IBM SPSS statistics, version 26. Normality was checked with the Shapiro-Wilk test and corresponding normality plots. Independent *t*-tests were used to compare anthropometrics, pain scores, and the FFI total and domain scores between patients with hEDS/HSD and healthy controls. Statistical Parametric Mapping (SPM) was used to compare ankle and foot kinematics during the entire stance phase of walking between the patients and controls. SPM was performed in Matlab (version 2019a) using a 2-tailed 2-sample *t*-test. SPM allows the calculation of the traditional *t* statistics over the entire normalized time-series, which is commonly used in biomechanical literature (25). Overall, the level of significance was set at $\alpha = 0.05$.

RESULTS

Subjects, anthropometrics, foot function, and pain. A total of 24 patients performed the gait analysis. One subject's

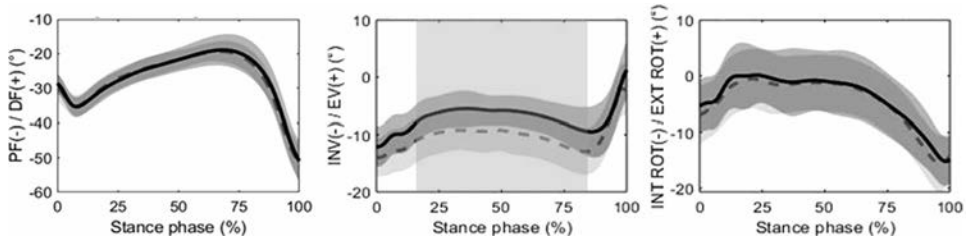

Figure 1. Kinematic between-group comparison of the rigid foot in the sagittal, frontal, and transverse plane during the stance phase of walking. Kinematic trajectories are presented as mean and SD clouds (solid line = hypermobile Ehlers-Danlos syndrome and hypermobility spectrum disorder; broken line = control). The gray area indicates those time frames of the kinematic curve that are significantly different between study groups. DF = dorsiflexion; EV = eversion; EXT ROT = external rotation; INT ROT = internal rotation; INV = inversion; PF = plantar flexion.

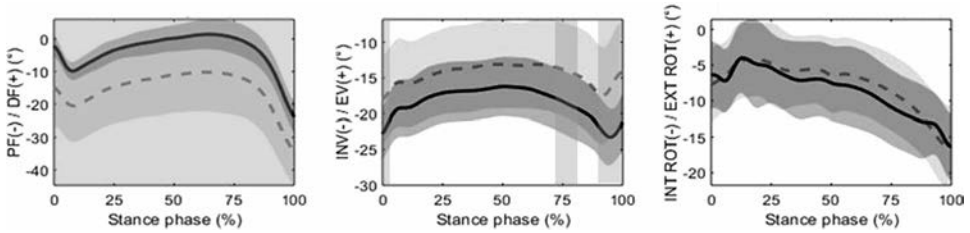

Figure 2. Kinematic between-group comparison of the rearfoot in the sagittal, frontal, and transverse plane during the stance phase of walking. Kinematic trajectories are presented as mean and SD clouds (solid line = hypermobile Ehlers-Danlos syndrome and hypermobility spectrum disorder; broken line = control). The gray area indicates those time frames of the kinematic curve that are significantly different between study groups. DF = dorsiflexion; EV = eversion; EXT ROT = external rotation; INT ROT = internal rotation; INV = inversion; PF = plantar flexion.

data could not be used due to technical issues. Twenty-three patients (mean ± SD age 41 ± 11.0 years) and 23 healthy control subjects (age 41 ± 11.5 years) were analyzed. Of the patients, 16 are currently reclassified according to the 2017 EDS criteria as having HSD, and 7 retained the diagnosis of hEDS. For readability, the patient group is referred to as hEDS/HSD in this article. The data fulfilled the assumptions of normal distribution. No significant between-group differences ($P > 0.05$) were found for anthropometric properties between the hEDS/HSD group (mean ± SD weight 72.1 ± 19.27 kg; height 165 ± 9.0 cm; body mass index [BMI] 26.4 ± 6.69 kg/m^2) and the healthy control group (weight 66.4 ± 10.80 kg; height 166 ± 5.4 cm; BMI 24.1 ± 3.57 kg/m^2).

The Margolis Pain Diagram revealed a significantly higher mean body percentage of pain in the hEDS/HSD group (37.46%) when compared to the control group (1.02%) ($P < 0.001$). Furthermore, 91.30% of the patients reported pain at the lower extremity, with 73.91% at the shank, 69.57% at the foot, and 60.87% at the ankle joint. VAS scores for pain at the lower limb and foot showed significant between-group differences, with higher values in the hEDS/HSD group, both just before and after 3D gait analysis ($P < 0.001$) (Table 1). No within-group differences were found when pain VAS scores were compared before and after testing ($P > 0.05$). Furthermore, patients with hEDS/HSD demonstrated significantly higher total and domain scores for the FFI, indicating more pain, physical disability, and activity limitation due to foot pathology ($P < 0.001$) (Table 1).

Ankle and foot kinematics. The data fulfilled the assumptions of normal distribution. Ankle and foot kinematics during stance phase are presented as joint angles (degrees) for each of the 3 motion planes. Dorsiflexion in the sagittal plane, eversion in the frontal plane, and external rotation/abduction in the transverse plane are represented as positive values. We refer to Supplementary Figures 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24526/abstract, for the detailed SPM results of the single-segment and multisegment ankle and foot kinematics.

Single-segment kinematics. Kinematics of the foot as a single, rigid segment revealed a significantly increased eversion, from 16% to 84% ($P < 0.001$) of stance phase in patients with hEDS/HSD, when compared to healthy controls. No significant differences were found for kinematics in the sagittal and transverse plane between the patient and control group (Figure 1).

Multisegment kinematics. Patients with hEDS/HSD walked with a significantly increased dorsiflexion of the rearfoot ($P < 0.001$) during whole-stance phase when compared to healthy controls. Significantly increased inversion occurred from 0% to 3% ($P = 0.049$), 73% to 84% ($P = 0.044$), and 90% to 100% ($P = 0.046$) of stance phase for the hEDS/HSD-group. No significant differences were found for kinematics in the transverse plane between the patient and control group (Figure 2).

Kinematics of the midfoot demonstrated a significantly increased plantar flexed position ($P < 0.001$) in the hEDS/HSD-group in contrast to healthy controls during whole-stance phase.

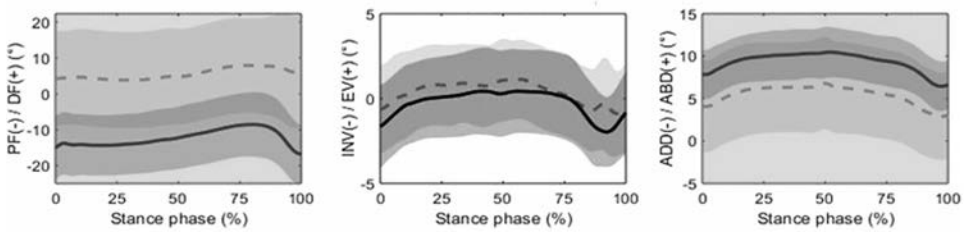

Figure 3. Kinematic between-group comparison of the midfoot in the sagittal, frontal, and transverse plane during the stance phase of walking. Kinematic trajectories are presented as mean and SD clouds (solid line = hypermobile Ehlers-Danlos syndrome and hypermobility spectrum disorder; broken line = control). The gray area indicates those time frames of the kinematic curve that are significantly different between study groups. ABD = abduction; ADD = adduction; DF = dorsiflexion; EV = eversion; INV = inversion; PF = plantar flexion.

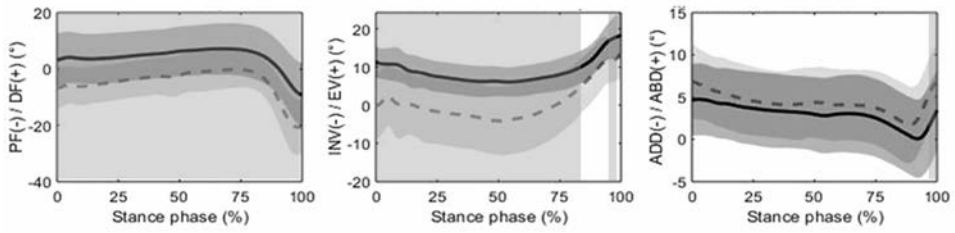

Figure 4. Kinematic between-group comparison of the medial forefoot in the sagittal, frontal, and transverse plane during the stance phase of walking. Kinematic trajectories are presented as mean and SD clouds (solid line = hypermobile Ehlers-Danlos syndrome and hypermobility spectrum disorder; broken line = control). The gray area indicates those time frames of the kinematic curve that are significantly different between study groups. ABD = abduction; ADD = adduction; DF = dorsiflexion; EV = eversion; INV = inversion; PF = plantar flexion.

No significant between-group differences were found in the frontal plane. Abduction was significantly increased ($P < 0.001$) during whole-stance phase for the patient group (Figure 3).

Significantly increased dorsiflexion ($P < 0.001$) at the lateral forefoot was shown in the hEDS/HSD-group during whole-stance phase when compared to healthy controls. No significant differences were found for frontal and transverse plane kinematics between the patient and control group (see Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24526/abstract).

Kinematics of the medial forefoot of patients with hEDS/HSD demonstrated significantly increased dorsiflexion ($P < 0.001$) during whole-stance phase when compared to healthy controls. Significantly increased eversion occurred from 0% to 84% ($P < 0.001$) and 94% to 99% ($P = 0.046$) of stance phase for the hEDS/HSD-group. Significantly decreased abduction was shown from 95% to 100% ($P = 0.049$) of stance phase for the patient group (Figure 4).

Patients with hEDS/HSD walked with a significantly decreased dorsiflexion of the hallux from 44% to 53% ($P = 0.037$) of stance phase when compared to healthy controls. Significantly increased inversion occurred from 0% to 4% ($P = 0.047$) and 6% to 65% ($P < 0.001$) of stance phase for the hEDS/HSD-group. Finally, there was a significantly increased abduction from 0% to 20% ($P = 0.016$) of stance phase for the patient group (Figure 5).

DISCUSSION

This study is the first to apply a multisegment foot model during gait to provide a better insight into the kinematics of the ankle-foot complex associated with hEDS/HSD. The results of the multisegment foot model suggesting hypermobility throughout the foot and the reported foot and ankle pain likely play an important role in the altered gait kinematics in patients with hEDS/HSD. First, we will reflect on the actual findings based on the multisegment foot model and subsequently emphasize the added value when compared to single-segment foot models.

Our multisegment foot model findings suggest that the characteristic hypermobility in patients with hEDS/HSD during gait is reflected in the first ray, represented by the medial forefoot. This first ray plays a key function in normal gait, alternating between a mobile adaptor and rigid lever. To comply with the function of a rigid lever, the medial forefoot everts during the mid and late stance toward a locked position (26). Hypermobility of the first ray, however, due to ligamentous laxity, is said to be related to an excessive eversion and has been associated with lower-extremity injuries (27). An increased eversion, a delay in the time-to-peak eversion or an extended duration of the eversion may signify a postponed locking mechanism, compromising an effective rigid lever during push off (28). Our study results confirm that subjects with hEDS/HSD display an increased everted position of the medial forefoot throughout stance phase. This dysfunctional locking mechanism is also represented by the increased

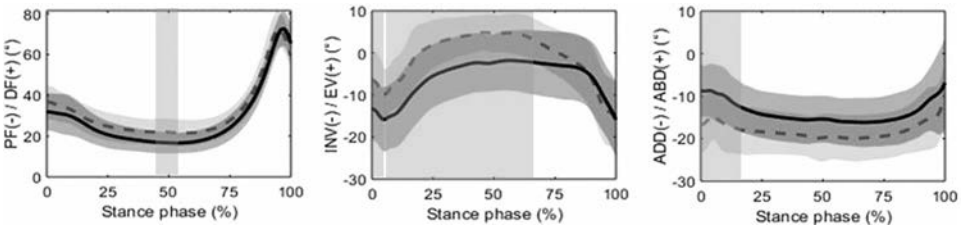

Figure 5. Kinematic between-group comparison of the hallux in the sagittal, frontal, and transverse plane during the stance phase of walking. Kinematic trajectories are presented as mean and SD clouds (solid line = hypermobile Ehlers-Danlos syndrome and hypermobility spectrum disorder; broken line = control). The gray area indicates those time frames of the kinematic curve that are significantly different between study groups. ABD = abduction; ADD = adduction; DF = dorsiflexion; EV = eversion; INV = inversion; PF = plantar flexion.

dorsiflexion mobility of the medial forefoot in subjects with hEDS/HSD compared to healthy controls. These altered push-off mechanics might increase metabolic costs during gait, potentially overloading the calf muscles (29). Because subjects with hEDS/HSD have a smaller cross-sectional area of the calf muscles compared to healthy subjects already, this mechanism probably contributes to accelerated fatigue, another key feature within this patient group (30).

Furthermore, our results suggest that the hypermobile first ray might influence the whole foot, where toward the proximal direction of the foot, the more dorsiflexed medial and lateral forefoot is compensated by an increased plantar flexion of the midfoot and in its turn by a more dorsiflexed position of the rearfoot segment. Toward the distal direction of the foot, the hallux demonstrates less dorsiflexion, more inversion, and more abduction in the hEDS/HSD group, being a counter rotation for the medial forefoot segment (the hypermobile first ray). We hypothesized that this possibility occurs to keep the hallux to the floor to maximize the stability of the hypermobile foot. Additionally, these deviating kinematics of the hallux could also be linked to the development of a hallux valgus, which is highly prevalent within hEDS/HSD patients (4–6).

As mentioned, the use of ankle joint kinematics computed with a single-segment foot model is subject to caution. Previous research has demonstrated opposite results for ankle joint kinematics for inversion and eversion during gait when measured with a single- or a multisegment model (17). Similar results were found in our study with a significantly more everted ankle position in subjects with hEDS/HSD compared to controls during midstance using the single-segment model, whereas a more inverted rearfoot position was noted throughout stance phase using a multisegment foot model. This is due to the fact that mid- and forefoot kinematics are ignored in the single-segment model and the resulting movement, including the mid- and forefoot, is only attributed to the ankle joint. This again emphasizes the importance of using multisegment foot models.

When comparing our single-segment foot results to the ones described in a previous study in subjects with hEDS/HSD (12), sagittal plane results differ. Those authors have identified a more plantar flexed position at initial contact and a reduced dorsiflexion during stance phase, whereas our single-segment model did not identify sagittal plane differences. A possible explanation might be the methodologic difference in marker setting to define the foot segment, as those authors only put markers on the lateral side of the foot (12), ignoring the influence and/or contribution by the medial part of the foot on kinematics. Furthermore, there is again a discrepancy between our single-segment and our multisegment model results in the sagittal plane. Although no significant differences were found in our single-segment foot results, a more dorsiflexed rearfoot position was observed throughout stance phase. This finding could be part of the hypermobile mechanism as presented above.

The second purpose of this case–control study was to evaluate foot function and pain. In line with other reports (31,32), our results show, by means of the Margolis Pain Diagram, that hEDS/HSD patients have generalized pain. Approximately two-thirds of the patients showed significant foot/ankle pain, which, together with foot hypermobility, likely plays an important role in the altered gait kinematics demonstrated in this study. However, other factors such as fatigue, muscle weakness, altered muscle activation, proprioception deficit, and kinesiophobia in hEDS/HSD could also impact kinematics (8,15,33–35). Furthermore, disability related to foot pain and activity limitations due to foot pathology were highly reported by our patients by means of the FFI. This finding is in accordance with the results of a recent study (36) and is in line with reports of difficulties with mobility due to foot problems in hEDS/HSD (11,37). Moreover, foot pain and other foot problems are reported to make daily life problematic to manage for individuals with hEDS/HSD. Untreated foot hypermobility and foot pain can be disabling and can result in significant difficulties for the patient and clinicians (7).

From a clinical perspective, precise and quantitative characterization of foot kinematics during gait in hEDS/HSD as performed in this study is important to be able to identify, develop, and enhance the rehabilitative options. An observational study showed that the use of custom-made foot orthoses for 3 months improved foot pain, disability related to foot pain, and foot functionality in patients with EDS (36). Weight-bearing phenolic foam molds of the feet were obtained to create custom-made foot orthoses adjusted from heel to just proximal to the metatarsal heads. Although we agree that custom-made foot orthoses are important, we suggest on the basis of the multisegment kinematic results of our study that custom-made foot orthoses for patients with EDS and HSD should be designed for the entire foot, inclusive of the forefoot and midfoot, and should not be limited to the rearfoot.

This study has certain limitations. First, a possible bias of the study is the relatively small sample size, resulting in limited strength of these findings. However, this gap may not have affected the statistical power, as multiple statistically significant findings were demonstrated. Second, this is a first exploratory study with the focus on ankle-foot kinematics during gait in hEDS/HSD. More research is necessary to further elucidate the underlying multisegment foot and ankle hypermobility and its potential impact on wider spatiotemporal, kinematic, and kinetic lower-limb parameters. Electromyographic evaluations can be a valuable additive to extract muscle activation patterns during walking. Third, all subjects walked barefoot. Further research could investigate the influence of shoes and custom-made functional foot orthotics on gait parameters in HSD, hEDS, and other types of EDS.

In conclusion, this study is, to the best of the authors' knowledge, the first one to use a multisegment foot model to provide better insight in the kinematics of the different parts of the foot in

hEDS/HSD. The hypermobility is also reflected in the different smaller parts of the foot, influencing the motion pattern of the different segments during gait. The validity of a single-segment foot model is of concern when describing foot kinematics. More in-depth studies are needed to better understand all the mechanisms involved within the kinematics between the different foot segments.

ACKNOWLEDGMENTS

The authors thank the patients with hEDS/HSD who participated in this study. We also thank S. Verstraelen, J. Carbon, and G. George for assistance in collecting the data and T. Palmans for technical assistance.

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. Rombaut 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. De Mits, Calders, De Schepper, Malfait, Rombaut.

Acquisition of data. Vermeulen, De Mits, Calders, De Schepper, Malfait, Rombaut.

Analysis and interpretation of data. Vermeulen, De Mits, De Ridder, Rombaut.

REFERENCES

1. Malfait F, Francomano C, Byers P, Belmont J, Berglund B, Black J, et al. The 2017 international classification of the Ehlers–Danlos syndromes. Am J Med Genet Part C Semin Med Genet 2017;175:8–26.
2. Castori M, Tinkle B, Levy H, Grahame R, Malfait F, Hakim A. A framework for the classification of joint hypermobility and related conditions. Am J Med Genet Part C Semin Med Genet 2017;175:148–57.
3. Rombaut L, Malfait F, Cools A, De Paepe AD, Calders P. Musculoskeletal complaints, physical activity and health-related quality of life among patients with the Ehlers-Danlos syndrome hypermobility type. Disabil Rehabil 2010;32:1339–45.
4. Beighton P, Horan FT. Surgical aspects of the Ehlers-Danlos syndrome: a survey of 100 cases. Br J Surg 1969;56:255–59.
5. Tompkins MH, Bellacosa R A. Podiatric surgical considerations in the Ehlers-Danlos patient. J Foot Ankle Surg 1997;36:381–7.
6. Pau M, Galli M, Celletti C, Morico G, Leban B, Albertini G, et al. Plantar pressure patterns in women affected by Ehlers-Danlos syndrome while standing and walking. Res Dev Disabil 2013;34:3720–6.
7. Cimolin V, Galli M, Celletti C, Pau M, Castori M, Morico G, et al. Foot type analysis based on electronic pedobarography data in individuals with joint hypermobility syndrome/Ehlers-Danlos syndrome hypermobility type during upright standing. J Am Podiatr Med Assoc 2014;104:588–93.
8. Rombaut L, Malfait F, De Wandele I, Taes Y, Thijs Y, De Paepe A, et al. Muscle mass, muscle strength, functional performance, and physical impairment in women with the hypermobility type of Ehlers-Danlos syndrome. Arthritis Care Res (Hoboken) 2012;64:1584–92.
9. Cimolin V, Galli M, Vismara L, Grugni G, Camerota F, Celletti C, et al. Gait pattern in two rare genetic conditions characterized by muscular hypotonia: Ehlers-Danlos and Prader-Willi syndrome. Res Dev Disabil 2011;32:1722–8.
10. Rombaut L, Malfait F, De Wandele I, Thijs Y, Palmans T, De Paepe A, et al. Balance, gait, falls, and fear of falling in women with the hypermobility type of Ehlers-Danlos syndrome. Arthritis Care Res (Hoboken) 2011;63:1432–9.
11. Rombaut L, Malfait F, De Paepe A, Rimbaut S, Verbruggen G, De Wandele I, et al. Impairment and impact of pain in female patients with Ehlers-Danlos syndrome: a comparative study with fibromyalgia and rheumatoid arthritis. Arthritis Rheum 2011;63:1979–87.
12. Galli M, Cimolin V, Rigoldi C, Castori M, Celletti C, Albertini G, et al. Gait strategy in patients with Ehlers-Danlos syndrome hypermobility type: a kinematic and kinetic evaluation using 3D gait analysis. Res Dev Disabil 2011;32:1663–8.
13. Rigoldi C, Galli M, Cimolin V, Camerota F, Celletti C, Tenore N, et al. Gait strategy in patients with Ehlers-Danlos syndrome hypermobility type and Down syndrome. Res Dev Disabil 2012;33:1437–42.
14. Celletti C, Galli M, Cimolin V, Castori M, Tenore N, Albertini G, et al. Use of the gait profile score for the evaluation of patients with joint hypermobility syndrome/Ehlers-Danlos syndrome hypermobility type. Res Dev Disabil 2013;34:4280–5.
15. Robbins SM, Cossette-Levasseur M, Kikuchi K, Sarjeant J, Shiu YG, Azar C, et al. Neuromuscular activation differences during gait in patients with Ehlers-Danlos syndrome and healthy adults. Arthritis Care Res (Hoboken) 2020;72:1653–62.
16. Alsiri N, Cramp M, Barnett S, Palmer S. Gait biomechanics in joint hypermobility syndrome: a spatiotemporal, kinematic and kinetic analysis. Musculoskeletal Care 2020;18:301–14.
17. Pothrat C, Authier G, Viehweger E, Berton E, Rao G. One- and multi-segment foot models lead to opposite results on ankle joint kinematics during gait: implications for clinical assessment. Clin Biomech (Bristol, Avon) 2015;30:493–9.
18. Hyslop E, Woodburn J, McInnes IB, Semple R, Newcombe L, Hendry G, et al. A reliability study of biomechanical foot function in psoriatic arthritis based on a novel multi-segmented foot model. Gait Posture 2010;32:619–26.
19. Perry J. Phases of gait. In: Gait analysis, normal and pathological function. Thorofare (NJ); Slack: 1992. p. 9–19.
20. Beighton P, De Paepe A, Steinmann B, Tsipouras P, Wenstrup RJ. Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Am J Med Genet 1997;77:31–7.
21. Beighton P, Solomon L, Soskolne CL. Articular mobility in an African population. Ann Rheum Dis 1973;32:413–8.
22. Margolis R, Tait R, Krause S. A rating system for use with patient pain drawings. Pain 1986;24:57–65.
23. Budiman-Mak E, Conrad KJ, Roach KE. The foot function index: a measure of foot pain and disability. J Clin Epidemiol 1991;44:561–70.
24. De Mits S, Segers V, Woodburn J, Elewaut D, De Clercq D, Roosen P. A clinically applicable six-segmented foot model. J Orthop Res 2012;30:655–61.
25. Pataky TC. Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. J Biomech 2010;43:1976–82.
26. Perez HR, Reber LK, Christensen JC. The effect of frontal plane position on first ray motion: forefoot locking mechanism. Foot Ankle Int 2008;29:72–6.
27. Allen MK, Cuddeford TJ, Glasoe WM, DeKam LM, Lee PJ, Wagner KJ, et al. Relationship between static mobility of the first ray and first ray, midfoot, and hindfoot motion during gait. Foot Ankle Int 2004;25:391–6.
28. Donatelli RA. Normal biomechanics of the foot and ankle. J Orthop Sports Phys Ther 1985;7:91–5.
29. Tiberio D. Pathomechanics of structural foot deformities. Phys Ther 1988;68:1840–9.
30. Banica T, Coussens M, Verroken C, Calders P, De Wandele I, Malfait F, et al. Higher fracture prevalence and smaller bone size in

patients with hEDS/HSD: a prospective cohort study. Osteoporos Int 2019;31:849–56.

31. Rombaut L, Scheper M, De Wandele I, De Vries J, Meeus M, Malfait F et al. Chronic pain in patients with the hypermobility type of Ehlers-Danlos syndrome: evidence for generalized hyperalgesia. Clin Rheumatol 2015;34:1121–9.
32. Scheper MC, Pacey V, Rombaut L, Adams RD, Tofts L, Calders P, et al. Generalized hyperalgesia in children and adults diagnosed with hypermobility syndrome and Ehlers-Danlos syndrome hypermobility type: a discriminative analysis. Arthritis Care Res (Hoboken) 2017; 69:421–9.
33. Celletti C, Castori M, La Torre G, Camerota F. Evaluation of kinesiophobia and its correlations with pain and fatigue in joint hypermobility syndrome/Ehlers-Danlos syndrome hypermobility type. Biomed Res Int 2013;2013:1–7.
34. Hakim A, De Wandele I, O'Callaghan C, Pocinki A, Rowe P. Chronic fatigue in Ehlers-Danlos syndrome-hypermobile type. Am J Med Genet C Semin Med Genet 2017;175:175–80.
35. Rombaut L, De Paepe A, Malfait F, Cools A, Calders P. Joint position sense and vibratory perception sense in patients with Ehlers-Danlos syndrome type III (hypermobility type). Clin Rheumatol 2010;29: 289–95.
36. Reina-Bueno M, Vázquez-Bautista C, Palomo-Toucedo IC, Domínguez-Maldonado G, Castillo-López JM, Munuera-Martínez PV. Custom-made foot orthoses reduce pain and fatigue in patients with Ehlers-Danlos syndrome: a pilot study. Int J Environ Res Public Health 2020; 17:1359.
37. Berglund B, Nordström G, Hagberg C, Mattiasson AC. Foot pain and disability in individuals with Ehlers-Danlos syndrome (EDS): impact on daily life activities. Disabil Rehabil 2005;27:164–9.

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 849–857
DOI 10.1002/acr.24522

American College of Rheumatology
Empowering Rheumatology Professionals

Developing and Validating Methods to Assemble Systemic Lupus Erythematosus Births in the Electronic Health Record

April Barnado,[1] Amanda M. Eudy,[2] Ashley Blaske,[1] Lee Wheless,[1] Katie Kirchoff,[3] Jim C. Oates,[3] and Megan E. B. Clowse[2]

Objective. Electronic health records (EHRs) represent powerful tools to study rare diseases. Our objective was to develop and validate EHR algorithms to identify systemic lupus erythematosus (SLE) births across centers.

Methods. We developed algorithms in a training set using an EHR with over 3 million subjects and validated the algorithms at 2 other centers. Subjects at all 3 centers were selected using ≥1 code for SLE International Classification of Diseases, Ninth Revision (ICD-9) or SLE International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification (ICD-10-CM) and ≥1 ICD-9 or ICD-10-CM delivery code. A subject was a case if diagnosed with SLE by a rheumatologist and had a birth documented. We tested algorithms using SLE ICD-9 or ICD-10-CM codes, antimalarial use, a positive antinuclear antibody ≥1:160, and ever checked double-stranded DNA or complement, using both rule-based and machine learning methods. Positive predictive values (PPVs) and sensitivities were calculated. We assessed the impact of case definition, coding provider, and subject race on algorithm performance.

Results. Algorithms performed similarly across all 3 centers. Increasing the number of SLE codes, adding clinical data, and having a rheumatologist use the SLE code all increased the likelihood of identifying true SLE patients. All the algorithms had higher PPVs in African American versus White SLE births. Using machine learning methods, the total number of SLE codes and an SLE code from a rheumatologist were the most important variables in the model for SLE case status.

Conclusion. We developed and validated algorithms that use multiple types of data to identify SLE births in the EHR. Algorithms performed better in African American mothers than in White mothers.

INTRODUCTION

Systemic lupus erythematosus (SLE) is an autoimmune disease that primarily affects women of childbearing age. Studying pregnancy outcomes in SLE is difficult given its relative rarity. SLE pregnancy studies are typically limited to a single-center cohort (1,2) that may not reflect real-world pregnancy outcomes. Population-based studies have investigated SLE pregnancies, but mainly in European populations (3–5).

Electronic health records (EHRs) contain rich, longitudinal data and serve as a powerful research tool (6). To harness the power of EHR data, validated methods are needed to identify subjects accurately. Using billing codes alone does not accurately identify SLE patients in the EHR (7,8). Adding clinical data to billing codes (8) and using noncoded data (9) improves the accuracy of identifying SLE patients in the EHR. These methods did not focus on identifying SLE births. There is a paucity of EHR studies in SLE pregnancy outcomes, with only 1 study using delivery data from multiple EHRs (10). That study used a 1-time SLE International Classification of Diseases, Ninth Revision (ICD-9) billing code at delivery or discharge to identify SLE births and did not conduct chart review to confirm SLE case status. Building upon our prior SLE algorithms that incorporate clinical data with billing codes, we incorporated ICD-9 or International Statistical Classification

Vanderbilt University Medical Center was supported by the NIH/National Center for Research Resources (UL1-RR-024975), and NIH/National Center for Advancing Translational Sciences (NCATS: UL-TR-000445). Duke University Medical Center was supported by the NIH/NCATS (UL1-TR-002553). Medical University of South Carolina was supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS: P60-AR-062755 and AR-072582), and by NIH/NCATS (UL1-TR-001450). Dr. Barnado's work was supported by the NIH/NIAMS (1K08-AR-072757-01). Dr. Eudy's work was supported by the NIH/NCATS (1KL2-TR-002554).

[1]April Barnado, MD, MSCI, Ashley Blaske, MD, Lee Wheless, MD, PhD: Vanderbilt University Medical Center, Nashville, Tennessee; [2]Amanda M. Eudy, PhD, Megan E. B. Clowse, MD, MPH: Duke University Medical Center, Durham, North Carolina; [3]Katie Kirchoff, MSHI, Jim C. Oates, MD: Medical University of South Carolina, Charleston.

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

Address correspondence to April Barnado, MD, MSCI, 1161 21st Avenue South, T3113 MCN, Nashville, TN 37232. Email: april.barnado@vumc.org.

Submitted for publication April 1, 2020; accepted in revised form November 24, 2020.

SIGNIFICANCE & INNOVATIONS

- To the best of our knowledge, we are the first to assemble an electronic health record (EHR)–based systemic lupus erythematosus (SLE) cohort and SLE birth cohort across multiple centers in the US.
- We developed, validated, and successfully deployed EHR-based algorithms to identify a subset of patients with a rare disease across multiple centers.
- We demonstrated key factors of clinical data, case definition, coding provider, and subject race that impact EHR algorithm performance and portability.
- We employed both traditional, rule-based algorithm methods as well as machine learning techniques, including extreme gradient boosting.
- The performance of the SLE delivery algorithms varied by race, with higher positive predictive values in African American mothers compared to White mothers.

of Diseases and Related Health Problems, Tenth Revision, Clinical Modification (ICD-10-CM) codes to identify SLE deliveries in the EHR. We also validated algorithms at multiple centers and investigated the impact of patient and provider factors on algorithm performance and portability. We then applied these algorithms to assemble a large, multicenter EHR cohort of SLE deliveries at 3 tertiary care centers in the Southeastern US.

PATIENTS AND METHODS

Patient selection. This study was approved by the institutional review board for each center. Due to center differences in how EHR data are stored and accessed, the methods for identifying SLE deliveries were slightly different, as described below. An overview of our approach is illustrated in Figure 1. Vanderbilt University Medical Center (VUMC) served as a training set. Duke University Medical Center (DUMC) and Medical University of South Carolina (MUSC) served as external validation sets. Chart review rules were consistent across all 3 centers. A subject was defined as a case if diagnosed with SLE by a rheumatologist and had a delivery documented at the institution after SLE diagnosis. Remaining subjects were classified as not a case, probable case, or missing. Subjects with cutaneous or drug-induced lupus or other autoimmune diseases were counted as not cases. Subjects who were given an SLE diagnosis by a nonrheumatology provider were counted as not cases. Probable cases were subjects who had a probable SLE diagnosis by a rheumatologist or who were labeled as having undifferentiated connective tissue disease or mixed connective disease by a rheumatologist. Probable subjects were counted as not cases in the primary analysis. Missing subjects who had no clinical documentation to determine case status were excluded. Delivery status was assessed on chart review at all 3 centers.

VUMC. We used a deidentified version of VUMC's EHR called the Synthetic Derivative (6), which contains over 3.2 million subjects. We searched for potential SLE deliveries restricting to female subjects between ages 12 to 65 years using ≥1 count of the SLE ICD-9 code (710.0) or SLE ICD-10-CM codes (M32.1*, M32.8, M32.9) and ≥1 ICD-9 or ICD-10-CM code for delivery-related diagnoses. The ICD-9 delivery codes have been validated with positive predictive values (PPVs) >90% (11) and have been used to assess pregnancy outcomes in other chronic diseases (12,13) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). Of these potential SLE cases, we randomly selected 100 for chart review to identify case status

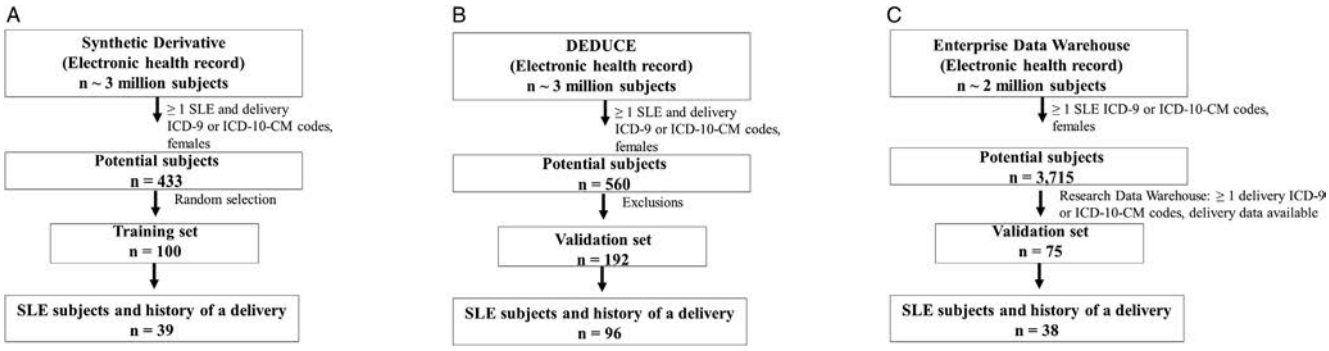

Figure 1. Training and validation sets. **A**, Training set formed at Vanderbilt University Medical Center starting with the Synthetic Derivative, a deidentified electronic health record, and applying ≥1 systemic lupus erythematosus (SLE) and delivery International Classification of Diseases, Ninth Revision (ICD-9) or International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification (ICD-10-CM) code to female subjects, resulting in 433 potential SLE subjects with deliveries. **B**, Validation set created at Duke University Medical Center by applying ≥1 SLE and ≥1 delivery ICD-9 or ICD-10-CM code to females subjects in a deidentified electronic health record called DEDUCE, resulting in 560 potential SLE subjects (for full list of exclusions, see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). **C**, Validation set created at Medical University of South Carolina (MUSC) by applying ≥1 SLE ICD-9 or ICD-10-CM code while restricting to female subjects in the Enterprise Data Warehouse, resulting in 3,715 subjects.

and to serve as a training set for algorithm development (Figure 1A).

DUMC. At DUMC, potential patients with ≥1 SLE ICD-9 or ICD-10-CM code and ≥1 ICD-9 or ICD-10-CM delivery code restricting to female subjects between ages 12 to 65 years were selected from the Duke Enterprise Data Unified Content Explorer (DEDUCE) data set (Figure 1B). Of those potential patients, exclusions (i.e., no delivery at Duke or unknown pregnancy outcome) were applied to facilitate chart review. A full list of exclusions is in Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract.

MUSC. At MUSC, female patients between ages 12 and 65 years with ≥1 SLE ICD-9 or ICD-10-CM code were selected from the Enterprise Data Warehouse from 2007 to 2017. As delivery data are stored in a different data warehouse (Research Data Warehouse), a second step was performed where subjects were selected who had ≥1 ICD-9 or ICD-10-CM delivery code and delivery data available (Figure 1C). Of these potential subjects, chart review was conducted.

Algorithm development and validation. A priori, we selected clinically important criteria that would be available in the EHR. We selected SLE ICD-9 and ICD-10-CM code counts, ever documented antimalarials, a positive antinuclear antibody (ANA) ≥1:160, and ever checked double-stranded DNA (dsDNA) or complement (C3 or C4). Occurrences of billing codes represent distinct days. Antimalarials included were hydroxychloroquine, Plaquenil, chloroquine, quinacrine, and Aralen. We tested algorithms using ≥1, ≥2, ≥3, and ≥4 code counts of the SLE ICD-9 code, SLE ICD-10-CM codes, or "SLE ICD-9 or SLE ICD-10-CM" codes. We then combined these codes with the above clinical data using "and" or "or" for possible algorithms. The PPV was calculated as the number of subjects who fit the algorithm and were confirmed cases on chart review, divided by the total number of subjects who fit the algorithm. Sensitivity was calculated as the number of subjects who fit the algorithm and were confirmed cases on chart review, divided by the total number of confirmed cases. To fit the algorithm, the subject had to have available data for that particular algorithm's criteria. If laboratory results were not checked at the center, they were considered missing. The F score, which is the harmonic mean of the PPV and sensitivity: (2 × PPV × sensitivity)/(PPV + sensitivity), was calculated for all algorithms.

Cohort assembly. The algorithm with the highest F score (≥4 counts of the SLE ICD-9 or ICD-10-CM codes) was applied across all centers to identify potential deliveries. All the subjects who fit the algorithm were chart reviewed to determine SLE case status, defined as SLE diagnosis by a rheumatologist. Only pregnancies delivered at the center with available outcomes that occurred after SLE diagnosis were included. Data were available for VUMC from 1993 to 2017, DUMC from 2007 to 2018, and MUSC from 2007 to 2017.

Sensitivity analysis. The primary analysis defined cases as diagnosed with SLE by a rheumatologist on chart review and allowed ICD-9 or ICD-10-CM codes to be billed by any provider. One sensitivity analysis changed the case definition to also include probable SLE cases. A second sensitivity analysis included only SLE ICD-9 or ICD-10-CM codes billed by a rheumatology provider. Additionally, we determined differences in algorithm performance by maternal race.

Machine learning methods. In addition to rule-based algorithms, we used machine learning methods, random forest and extreme gradient boosting (XGB) for algorithm development. Random forest builds multiple classification trees (a "forest") using a random sample of input variables for each tree (14,15). The final classification is an average of the forest. XGB is an ensemble method that is the summation of multiple models where each successive model attempts to correct errors in the previous model to improve overall performance. Data across 3 centers were randomly divided in training sets (80%) and testing sets (20%). For race-stratified analyses, to increase sample size, training and testing sets were 70% and 30%, respectively. Models were constructed using the training set with 5-fold cross validation, and were tuned using the caret package (16,17). Final model performance was assessed using the test set. The ranger package was used for random forest models (18), and the xgboost package for XGB models using the method = "xgbTree" in the caret framework (19). We reported algorithms with the highest PPVs in the test set and identified the most important variables in the models. Model input variables including the following: the total number of SLE ICD codes, SLE code from a rheumatologist, ever antimalarial use, ANA positive, ever checked dsDNA, ever checked C3, ever checked C4, age, race, SLE duration defined as the first SLE code to delivery date, EHR duration defined as the first code for any condition to delivery date, and center. All analyses were conducted in R software, version 3.5.1.

RESULTS

Description of the training set. An overview of our approach is illustrated in Figure 1. A training set was created at VUMC by applying at least 1 SLE and 1 delivery ICD-9 or ICD-10-CM codes to the Synthetic Derivative, resulting in 433 potential SLE deliveries. Of the 433, 100 were randomly selected for chart review. Of these, 40 subjects were SLE cases, with 39 subjects having a delivery documented after SLE diagnosis. There were 37 subjects who were not SLE cases, 16 with a probable SLE diagnosis, and 7 with missing clinic notes. Of the 37 subjects not classified as SLE, 21 had alternative autoimmune diagnoses,

with the most common being a subject with a positive autoantibody (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract).

Description of the validation sets. A validation set was created at DUMC by applying at least 1 SLE and 1 delivery ICD-9 or ICD-10-CM codes to DEDUCE, resulting in 560 potential SLE deliveries. Of these, 192 had deliveries that occurred after an SLE diagnosis. On chart review of these 192, 95 were an SLE case and 36 were probable SLE. Of the remaining subjects, 61 did not have SLE, of which 31 had alternative autoimmune diagnoses, with the most common being cutaneous lupus (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract).

A second validation set was created at MUSC by applying at least 1 SLE ICD-9 or ICD-10-CM code and selecting for female subjects in the EHR. Of these 3,715 potential SLE subjects, subjects with at least 1 delivery ICD-9 or ICD-10-CM code

Table 1. Summary of algorithm performances*

Algorithm	VUMC: training set			DUMC: validation set			MUSC: validation set		
	PPV	Sensitivity	F score	PPV	Sensitivity	F score	PPV	Sensitivity	F score
ICD-9 or ICD-10-CM code counts									
≥1	56	100	72	50	100	67	51	100	68
≥4	81	95	87	71	88	79	64	84	73
ICD-9 or ICD-10-CM code counts and ever antimalarial documented									
≥1	68	85	76	67	76	71	72	94	82
≥4	83	83	83	76	70	73	74	81	77
ICD-9 or ICD-10-CM code counts and ANA positive†									
≥1	64	76	69	56	96	71	64	91	75
≥4	79	76	77	72	86	78	76	83	79
ICD-9 or ICD-10-CM code counts and ever laboratory results checked‡									
≥1	62	95	75	59	91	72	61	87	72
≥4	84	93	88	74	81	77	68	74	71
Case definition: definite and probable SLE, ICD-9 or ICD-10-CM code counts									
≥1	81	100	90	63	100	77	75	100	86
≥4	91	67	77	78	78	78	90	80	85
Billing code provider: require rheumatology to use SLE ICD-9 or ICD-10-CM code counts									
≥1	79	78	78	82	81	81	77	54	63
≥4	85	73	79	96	56	71	100	24	39
Maternal race: African American, ICD-9 or ICD-10-CM code count									
≥1	71	100	83	60	100	75	61	100	76
≥4	84	94	89	81	85	83	73	88	80
Maternal race: White, ICD-9 or ICD-10-CM code count									
≥1	48	100	65	33	100	50	34	100	51
≥4	69	90	78	88	55	68	41	70	52
Maternal race: African American, ICD-9 or ICD-10-CM code count, require rheumatology to use codes									
≥1	93	76	84	88	76	82	80	67	73
≥4	92	71	80	97	51	67	100	33	50
Maternal race: White, ICD-9 or ICD-10-CM code count, require rheumatology to use codes									
≥1	68	75	71	71	83	77	60	30	40
≥4	82	70	76	93	54	68	100	10	18

* Values are the percentage. Systemic lupus erythematosus (SLE) International Classification of Diseases, Ninth Revision (ICD-9) code: 710.0. SLE International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification (ICD-10-CM) codes: M32.1*, M32.8, M32.9. VUMC = Vanderbilt University Medical Center; DUMC = Duke University Medical Center; MUSC = Medical University of South Carolina; PPV = positive predictive value.
† Antinuclear antibody (ANA) positive (≥1:160).
‡ Ever laboratory results checked included double-stranded DNA, C3, or C4.

and a delivery documented at MUSC after SLE diagnosis resulted in 75 potential SLE deliveries. Of these, 38 were an SLE case and 11 were probable SLE. Of the remaining subjects, 26 did not have SLE, of which 15 had alternative autoimmune diagnoses, with the most common being cutaneous lupus (see Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract).

Algorithms using ICD-9 or ICD-10-CM codes. Algorithm performances using counts of the SLE ICD-9, ICD-10-CM, and ICD-9 or ICD-10-CM codes in the training (VUMC) and validation (DUMC, MUSC) sets are shown as a summary in Table 1, with full data in Supplementary Tables 6–8 (available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). PPVs and sensitivities for algorithms using only ICD-9 codes were similar across centers. Requiring more code counts increased PPVs but decreased sensitivities. Across 3 centers, increasing the number of ICD-9 code counts increased PPVs from 54–57% for ≥1 code to 74–77% for ≥4 codes. As data duration for ICD-9 codes differed in training versus validation centers, we limited the training set data duration to 2007–2017 to match the 2 validation centers. Within the training set, the ICD-9 code algorithm performances for the restricted 2007–2017 duration were similar to the algorithm performances for the full data duration 1993–2017 (see Supplementary Table 9, at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract).

Algorithms using only ICD-10-CM codes performed differently between centers, with higher PPVs in the training set but lower PPVs in the validation sets. For algorithms using either ICD-9 or ICD-10-CM codes, PPVs were in between the algorithms using ICD-9 codes and algorithms using ICD-10-CM codes.

Algorithms incorporating clinical data. We investigated adding ever antimalarial documented, ever checked laboratory results (dsDNA, C3, or C4), and a positive ANA (≥1:160) to ICD-9 and ICD-10-CM codes. Algorithms that incorporated clinical data had higher PPVs compared to algorithms using only codes (Table 1 and Supplementary Tables 6–8, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). This addition, however, lowered sensitivities, as some SLE patients did not have data documented in the center's EHR. One exception was in the training set where adding clinical data to the ICD-10-CM codes did not significantly increase PPVs, as PPVs for ICD-10-CM codes were already high. Across all 3 centers, adding antimalarials to the codes improved PPVs most robustly. Adding ever checked laboratory results to the codes increased PPVs slightly. Adding a positive ANA did not significantly increase PPVs but decreased sensitivities.

Case definition. For the above analyses, probable SLE subjects were counted as not cases. We examined algorithm performance with counting probable SLE subjects as cases (Table 1 and Supplementary Tables 10–12, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). With this alternative case definition, PPVs increased substantially while sensitivities decreased for all algorithms.

Billing code provider. For the above analyses, we allowed for any provider to use the SLE codes. We investigated whether a rheumatology provider using the SLE codes impacted algorithm performance (Table 1 and Supplementary Tables 13–15, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). With requiring a rheumatology provider to use an SLE ICD-9 code, PPVs significantly increased at all centers with a decrease in sensitivities. Adding clinical data did not significantly improve PPVs, as PPVs were already relatively high. By requiring a rheumatologist to use the ICD-10-CM codes, PPVs at the validation centers more closely approximated the high PPVs at the training center.

Subject race. We evaluated the impact of subject race on algorithm performance. The prevalence for African American subjects was 31% in the training set and 51% and 55% in the validation sets. While sensitivities were similar, PPVs were significantly higher in African American subjects compared to White subjects (Table 1). Pooling data from the 3 centers, the algorithm with ≥4 counts of the SLE ICD-9 or ICD-10-CM codes had a PPV of 78% in African American subjects compared to 54% in White subjects (see Supplementary Table 16, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). Adding clinical data to the codes increased PPVs in White subjects but not in African American subjects (see Supplementary Table 16). Requiring rheumatology to use the codes increased PPVs significantly in both White subjects and African American subjects (see Supplementary Table 17, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract).

Results of machine learning methods. We performed random forest and XGB models for algorithm development. For random forest, the algorithm with the highest PPV included 500 trees and sampled 2 random variables per tree, with PPV of 79%, sensitivity of 80%, F score of 80%, negative predictive value (NPV) of 81%, and area under the curve of 87% in the training set (see Supplementary Tables 18 and 19, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). The most important variables in the model were the total number of SLE ICD codes and rheumatology using the SLE codes. Model performance varied by race, with an F score of 0.87 in African American subjects versus 0.67

in White subjects. For XGB, the highest-performing model had PPV of 79%, sensitivity of 82%, F score of 80%, NPV of 82%, and area under the curve of 84% in the training set (see Supplementary Tables 20 and 21, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24522/abstract). The most important variables in the model were the total number of SLE ICD codes and rheumatology using the SLE codes. Model performance varied by race, with an F score of 0.89 in African American subjects versus 0.71 in White subjects.

Highest performing algorithms. We assessed algorithm performance with PPV, sensitivity, and F score, a measure that accounts for both PPV and sensitivity. Algorithms' performances varied across the 3 centers, leading to different high-performing algorithms at each center. Algorithms with the highest PPVs included higher SLE code counts, incorporated clinical data, expanded the case definition to include probable and definite SLE cases, and required rheumatology to use the SLE codes (Figure 2). Algorithms that incorporated ≥4 SLE codes by rheumatology along with ever antimalarial documented had PPVs from 90% to 100% across the 3 centers. Algorithms with the highest sensitivities were those that used fewer SLE code counts and incorporated either SLE ICD-9 or ICD-10-CM codes (Figure 2). The algorithm with the highest F score across the 3 centers used ≥4 counts of the SLE ICD-9 or ICD-10-CM codes and was 87% at VUMC, 79% at DUMC, and 73% at MUSC (Table 1).

Results of cohort assembly. Deploying the algorithm with the highest F score (≥4 counts of the SLE ICD-9 or ICD-10-CM codes) resulted in 438 possible SLE deliveries across the 3 centers (Table 2). In this cohort, the mean ± SD age at first delivery was 29.5 ± 1.2 years, with White patient deliveries at 51%, African American at 42%, Asian at 3%, and 4% other. Only 5% of deliveries were of Hispanic ethnicity.

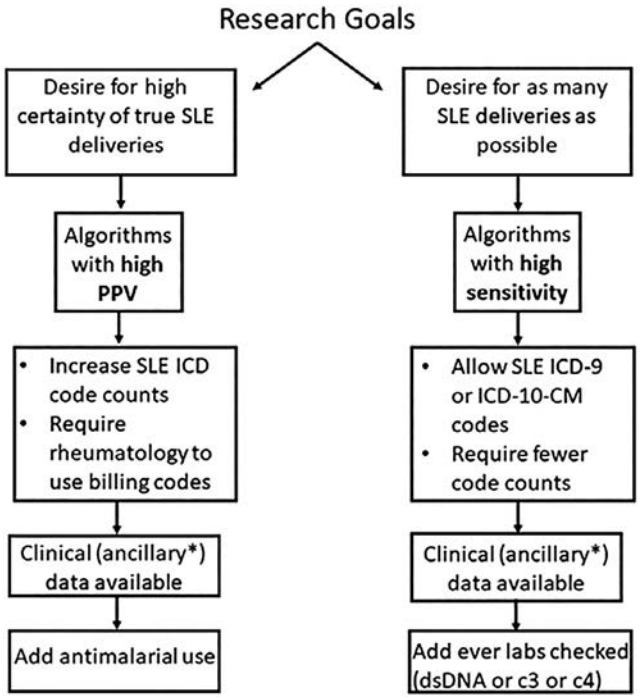

Figure 2. Guide to selecting algorithms to identify systemic lupus erythematosus (SLE) deliveries in the Electronic Health Record. Algorithms can be selected based on the researcher's goals and available clinical or ancillary data. If there is a desire for high certainty for true SLE deliveries, one would select an algorithm with a high positive predictive value (PPV). If chart review is not available or possible to confirm case status, one would also want to select an algorithm with a high PPV. Alternatively, if there is a desire to assemble as many SLE deliveries as possible, one would select an algorithm with a high sensitivity. If clinical or ancillary data are available, such as structured electronic health record data including laboratory values or medications, those data will further influence algorithm selection. dsDNA = double-stranded DNA; ICD-9 = International Classification of Diseases, Ninth Revision; ICD-10-CM = International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification.

DISCUSSION

We have harnessed the power of the EHR to develop, validate, and deploy algorithms that assemble a rare event across multiple EHRs. To the best of our knowledge, this is one of the first successful applications of assembling SLE and SLE deliveries from several centers using EHR data in the US. This work is important because it establishes valuable methods for researchers to not only identify SLE and SLE deliveries but also other outcomes across EHRs. In summary, increasing the number of SLE codes, adding ever antimalarial documented to codes, expanding case definition to probable and definite SLE cases, and requiring rheumatology to use SLE codes all improved the algorithm PPVs. Subject race had a significant impact on algorithm performance, with significantly higher PPVs in African American subjects compared to White subjects.

While there are validated algorithms to identify SLE in the EHR (8,9), there are no studies on accurately identifying SLE deliveries in the EHR. Literature in sickle cell anemia (12) has evaluated and validated delivery codes in studying pregnancy outcomes (11,13,20,21). Our study used these validated delivery codes and built upon our work in SLE EHR algorithms (8,9). As ICD codes transitioned from ICD-9 to ICD-10 in the US on October 1, 2015, we developed algorithms that incorporated either ICD-9 or ICD-10-CM codes to capture both historical and newly diagnosed SLE births. To the best of our knowledge, there are no published studies that investigate the performance of SLE ICD-10-CM codes. We developed multiple algorithms that incorporate different types of data to meet researchers' diverse goals. We also performed validation and found good portability of the algorithms. Lastly, we identified key factors such as clinical data, case

Table 2. Characteristics of systemic lupus erythematosus deliveries using a high-performing algorithm across 3 centers*

	≥4 SLE ICD-9/ICD-10-CM and delivery codes				Probable and definite SLE deliveries				Definite SLE deliveries			
	Overall	VUMC	DUMC	MUSC	Overall	VUMC	DUMC	MUSC	Overall	VUMC	DUMC	MUSC
	(n = 438)	(n = 269)	(n = 119)	(n = 50)	(n = 369)	(n = 231)	(n = 93)	(n = 45)	(n = 286)	(n = 170)	(n = 84)	(n = 32)
Age at delivery, mean ± SD (range)	29.5 ± 1.2 (16–46)	28.4 ± 5.6 (16–41)	30.8 ± 5.9 (19–46)	29.4 ± 5.2 (19–41)	29.5 ± 1.4 (16–46)	28.3 ± 5.6 (16–41)	31.0 ± 5.9 (19–46)	29.3 ± 5.2 (19–41)	29.2 ± 1.4 (16–46)	28.1 ± 5.7 (16–41)	30.8 ± 5.9 (19–46)	28.8 ± 5.6 (19–41)
Race												
African American	175 (42)	83 (33)	62 (62)	30 (64)	157 (44)	74 (34)	53 (57)	30 (71)	133 (48)	61 (38)	50 (6)	22 (76)
White	215 (51)	160 (63)	38 (32)	17 (36)	171 (48)	134 (61)	25 (27)	12 (29)	121 (44)	93 (57)	21 (25)	7 (24)
Asian	14 (3)	7 (3)	7 (6)	0 (0)	12 (4)	6 (3)	6 (6)	0 (0)	11 (4)	6 (4)	5 (6)	0 (0)
Other	17 (4)	5 (2)	12 (10)	0 (0)	14 (4)	5 (2)	9 (10)	0 (0)	10 (4)	2 (1)	8 (10)	0 (0)
Unknown, no.	17	14	0	3	15	12	0	3	11	8	0	3
Hispanic ethnicity	24 (5)	17 (6)	7 (6)	0 (0)	21 (6)	17 (7)	4 (4)	0 (0)	14 (5)	10 (6)	4 (5)	0 (0)

* Values are the number (%) unless indicated otherwise. Applying the algorithm of ≥4 systemic lupus erythematosus (SLE) International Classification of Diseases, Ninth Revision (ICD-9) or SLE International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification (ICD-10-CM) and delivery codes resulted in 438 possible deliveries across the 3 centers. Of these 438 deliveries, 369 were probable and definite SLE deliveries based on chart review, and 286 were definite SLE deliveries. Electronic health record delivery data available at Vanderbilt University Medical Center (VUMC) 1993–2017, Duke University Medical Center (DUMC) 2007–2018, and Medical University of South Carolina (MUSC) 2014–2017.

definition, subject race, and coding provider that all significantly impact algorithm performance.

Performance of algorithms using ICD-10-CM codes was significantly better in the training versus validation centers. Because of high performance in the training center, adding clinical data or requiring rheumatologists to use codes did not impact the algorithms' PPVs. In contrast, at the validation centers, requiring the rheumatologist to use SLE ICD-10-CM codes dramatically improved the algorithms' PPVs. We hypothesize that at the training center, where the ICD-10-CM algorithms performed well, rheumatologists were early adopters and frequent users of SLE ICD-10-CM codes. In contrast, at the validation sets, we suspect that nonrheumatology providers may have incorrectly assigned SLE ICD-10-CM codes. A rheumatology provider using an SLE code was one of the most important variables in machine learning models for SLE case status. Thus, this difference in provider coding likely explains the difference in ICD-10-CM algorithm performance in training versus validation sets.

As expected, requiring higher counts of SLE ICD-9 or ICD-10-CM codes resulted in algorithms with higher PPVs but lower sensitivities. The more visits a potential SLE patient has, the more times an SLE code is used, with the clinician feeling confident with the diagnosis. In general, adding clinical data to the codes improved the algorithms' PPVs but decreased sensitivities. For example, some SLE patients did not have clinical data, such as a positive ANA within the center's EHR, as they were followed by an outside rheumatologist. Using SLE ICD-9 or ICD-10-CM codes from only rheumatologists resulted in algorithms with high PPVs without requiring clinical data. These algorithms would be useful if clinical data are not available, but this method limits the sample to SLE women managed by center rheumatologists.

As expected, broadening the SLE case definition to include probable patients increased PPVs while decreasing sensitivities. Compared to definite SLE patients, probable SLE patients were more likely to have fewer SLE codes. Algorithms requiring higher counts of the SLE codes would then exclude more of these probable SLE patients, resulting in lower sensitivities. We used a specialist diagnosis for SLE versus using the American College of Rheumatology (ACR) SLE criteria (22), as ACR SLE criteria are not documented systematically in notes. We previously demonstrated that requiring documentation of ACR SLE criteria excludes approximately 26% of true SLE patients (8). Researchers can, however, select a case definition based on their study's goals and available data.

Across all 3 centers, algorithms had higher PPVs but similar sensitivities in African American versus White patients. While increasing code counts and adding clinical data improved the algorithms' PPVs for White subjects somewhat, requiring rheumatology coding dramatically improved PPVs. Our results suggest a high rate of SLE over-labeling in White subjects, particularly by physicians other than rheumatologists. Thus, different algorithms may be needed for different races. Specifically, algorithms to identify White subjects accurately may require rheumatologists to use SLE billing codes. The impact of race on EHR phenotyping has not been explored in SLE or other chronic conditions. We hypothesize that the higher prevalence of SLE in African American subjects compared to White subjects (23–26) may contribute to this observation. PPVs are a function of disease prevalence, while sensitivities are a function of the algorithm.

In addition to rule-based algorithms, we performed machine learning models. These models had a similar F score to the high-performing rule-based algorithm of ≥4 counts of SLE ICD-9 or ICD-10-CM codes. The machine learning methods confirmed results from the rule-based algorithms, including variable model performance based on subject race. Machine learning methods, particularly XGB, are robust (27) and have advantages, including automatic model tuning and the ability to model complex interactions.

We developed and validated algorithms to identify SLE deliveries in the EHR. Similar methodology can identify and assemble other rare diseases or outcomes in the EHR. Researchers can choose methods based on available data and research goals (Figure 2). If the goal is to identify subjects with high certainty, one would select an algorithm with the highest PPV. In contrast, if one wants to select as many subjects as possible to increase sample size, one would select an algorithm with a high sensitivity and F score. We chose to use an algorithm with a high F score to amass the largest number of true SLE deliveries.

While we validated multiple EHR-based algorithms, our study has limitations. We started our search for possible SLE deliveries using at least 1 SLE ICD-9 or ICD-10-CM code. This search strategy in finding SLE in the EHR has an NPV of 98% (8), so we anticipate very few potential SLE deliveries were missed. Search strategies for identifying training and validation sets varied slightly due to differences in how data are stored and accessed at the 3 centers. In machine learning methods, model performance did not vary by center. Therefore, center heterogeneity had minimal impact on algorithm performance. Our algorithms were developed and validated at 3 tertiary care referral centers in the Southeastern US, which may limit generalizability to other centers.

While the EHR does not substitute for prospective cohort studies, EHRs contain longitudinal, real-word data that can dramatically increase the efficiency and sample size of a study. Using 1 of our validated, high-performing algorithms, we assembled over 400 potential SLE deliveries across 3 centers. With this large SLE delivery cohort, we will have the power in future studies to examine the impact of disease and provider factors on important outcomes such as preterm delivery and preeclampsia in SLE.

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. Barnado 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. Barnado, Eudy, Blaske, Wheless, Kirchoff, Oates, Clowse.

Acquisition of data. Barnado, Eudy, Blaske, Kirchoff, Clowse.

Analysis and interpretation of data. Barnado, Eudy, Wheless, Kirchoff, Oates, Clowse.

REFERENCES

1. Andrade RM, McGwin G Jr, Alarcón GS, Sanchez ML, Bertoli AM, Fernández M, et al. Predictors of post-partum damage accrual in systemic lupus erythematosus: data from LUMINA, a multiethnic US cohort (XXXVIII). Rheumatology (Oxford) 2006;45:1380–4.
2. Buyon JP, Kim MY, Guerra MM, Laskin CA, Petri M, Lockshin MD, et al. Predictors of pregnancy outcomes in patients with lupus: a cohort study. Ann Intern Med 2015;163:153–63.
3. Palmsten K, Simard JF, Chambers CD, Arkema E. Medication use among pregnant women with systemic lupus erythematosus and general population comparators. Rheumatology (Oxford) 2017;56: 561–9.
4. Skorpen CG, Lydersen S, Gilboe IM, Skomsvoll JF, Salvesen KA, Palm O, et al. Influence of disease activity and medications on offspring birth weight, pre-eclampsia and preterm birth in systemic lupus erythematosus: a population-based study. Ann Rheum Dis 2018;77: 264–9.
5. Zusman EZ, Sayre EC, Avina-Zubieta JA, De Vera MA. Patterns of medication use before, during and after pregnancy in women with systemic lupus erythematosus: a population-based cohort study. Lupus 2019;28:1205–13.
6. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther 2008; 84:362–9.
7. Moores KG, Sathe. A systematic review of validated methods for identifying systemic lupus erythematosus (SLE) using administrative or claims data. Vaccine 2013;31 Suppl 10:K62–73.
8. Barnado A, Casey C, Carroll RJ, Wheless L, Denny JC, Crofford LJ. Developing electronic health record algorithms that accurately identify patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69:687–93.
9. Jorge A, Castro VM, Barnado A, Gainer V, Hong C, Cai T, et al. Identifying lupus patients in electronic health records: development and validation of machine learning algorithms and application of rule-based algorithms. Semin Arthritis Rheum 2019;49: 84–90.
10. Williams A, Grantz K, Seeni I, Robledo C, Li S, Ouidir M, et al. Obstetric and neonatal complications among women with autoimmune disease. J Autoimmun 2019;103:102287.
11. Yasmeen S, Romano PS, Schembri M, Keyzer JM, Gilbert WM. Accuracy of obstetric diagnoses and procedures in hospital discharge data. Am J Obstet Gynecol 2006;194:992–1001.
12. Boulet SL, Okoroh EM, Azonobi I, Grant A, Hooper WC. Sickle cell disease in pregnancy: maternal complications in a Medicaid-enrolled population. Matern Child Health J 2013;17:200–7.
13. Zhang S, Cardarelli K, Shim R, Ye J, Booker KL, Rust G. Racial disparities in economic and clinical outcomes of pregnancy among Medicaid recipients. Matern Child Health J 2013;17:1518–25.
14. Breiman L, Friedman J, Olshen RA, Stone CJ. Classification and regression trees. In: The Wadsworth statistics/probability series. 1st ed. Belmont (CA): Wadsworth International Group; 1983.
15. Breiman L. Random forests. Machine learning 2001; 45:5–32.
16. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008; 28:1–26.
17. R Core Team. The R project for statistical computing. 2018. URL: https://www.R-project.org/.
18. Wright MN, Ziegler A. ranger: a fast implementation of random forests for high dimensional data in C++ and R. J Stat Softw 2017; 77: 1–17.
19. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. xgboost: extreme gradient boosting. 2019. URL: https://CRAN.R-project.org/package=xgboost.
20. Hardy JR, Holford TR, Hall GC, Bracken MB. Strategies for identifying pregnancies in the automated medical records of the General Practice Research Database. Pharmacoepidemiol Drug Saf 2004;13:749–59.
21. Devine S, West S, Andrews E, Tennis P, Hammad TA, Eaton S, et al. The identification of pregnancies within the general practice research database. Pharmacoepidemiol Drug Saf 2010;19:45–50.
22. tHochberg, MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997; 40:1725.
23. Lim SS, Bayakly AR, Helmick CG, Gordon C, Easley KA, Drenkard C. The incidence and prevalence of systemic lupus erythematosus, 2002-2004: The Georgia Lupus Registry. Arthritis Rheumatol 2014; 66:357–68.
24. Somers EC, Marder W, Cagnoli P, Lewis EE, DeGuire P, Gordon C, et al. Population-based incidence and prevalence of systemic lupus erythematosus: the Michigan Lupus Epidemiology and Surveillance program. Arthritis Rheumatol 2014;66:369–78.
25. Centers for Disease Control and Prevention. Trends in deaths from systemic lupus erythematosus: United States, 1979-1998. MMWR Morb Mortal Wkly Rep 2002;51:371–4.
26. Krishnan E, Hubert HB. Ethnicity and mortality from systemic lupus erythematosus in the US. Ann Rheum Dis 2006;65:1500–5.
27. Fernández-Delgado M, Sirsat MS, Cernadas E, Alawadi S, Barro S, Febrero-Bande M. An extensive experimental survey of regression methods. Neural Netw 2019;111:11–34.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 5, May 2022, pp 858–865
DOI 10.1002/acr.24713

Allopurinol and Cardiovascular Events: Time-Related Biases in Observational Studies

Samy Suissa,[1] Karine Suissa,[2] and Marie Hudson[1]

Objective. Several observational studies reported that allopurinol, an effective treatment for gout, was associated with important reductions in cardiovascular (CV) events, with calls for large, randomized trials, although some results were conflicting. The present study was undertaken to assess the extent of time-related biases in these observational studies.

Methods. We searched the literature for all observational studies reporting on allopurinol and CV events, focusing on 2 time-related biases. Time-related confounding bias results from studies using cohorts of patients all exposed to allopurinol, with comparisons based on episodes of allopurinol discontinuation, where confounding factors are not updated over follow-up time. Immortal time bias arises from the exposure misclassification of periods of cohort follow-up during which the outcome under study cannot occur.

Results. We identified 12 studies, of which 8 were affected by time-related confounding bias or immortal time bias, while the remaining 4 studies avoided these biases. The studies affected by time-related confounding bias resulted in significant reductions in the incidence of CV events with allopurinol use (pooled hazard ratio [HR] 0.88 [95% confidence interval (95% CI) 0.85–0.92]), as did the studies affected by immortal time bias (pooled HR 0.79 [95% CI 0.72–0.87]). The 4 studies that avoided these biases resulted in a pooled HR of 1.07 (95% CI 0.91–1.25).

Conclusion. Observational studies reporting significantly reduced incidence of CV events with allopurinol use were affected by time-related biases. Overall, studies that avoided these biases did not find a protective effect. The ALL-HEART randomized trial will provide important and accurate evidence on the potential effectiveness of allopurinol on CV outcomes.

INTRODUCTION

Patients with gout are at increased risks of cardiovascular (CV) events and CV-related mortality (1,2). In addition, hyperuricemia appears to be independently associated with CV disease risk, although whether this effect is causal is debated (3,4). Thus, as treatment with allopurinol is effective at decreasing urate levels and the frequency of flares in patients with gout, it has been postulated that allopurinol could also reduce the elevated CV risk in these patients (3,5).

A meta-analysis of 65 randomized controlled trials reporting on the comparison of allopurinol or oxypurinol with placebo in patients with gout and other indications found a nonsignificant odds ratio of major adverse CV events of 0.65 (95% confidence interval [95% CI] 0.41–1.05) (6). Several observational studies found significant reductions in CV events with allopurinol use compared with non-use, although some studies reported hazard ratios (HRs) above unity, resulting in uncertainty regarding the potential CV benefits of allopurinol (2,7,8). Indeed, the European Alliance of Associations for Rheumatology (EULAR) task force for the management of gout recognized that support for earlier treatment of gout based on the CV benefit of these drugs is founded on observational studies that yielded conflicting results (9). Consequently, the task force and others called for well-conducted trials of the CV effects of allopurinol, such as the ALL-HEART randomized trial (9–11).

Recently, with respect to the outcome of all-cause mortality, several observational studies reporting important reductions in mortality with allopurinol use were shown to be affected by major time-related biases (12). This examination of the mortality studies

Dr. S. Suissa is recipient of the Distinguished James McGill Professorship award.

[1]Samy Suissa, PhD, Marie Hudson, MD: Jewish General Hospital and McGill University, Montreal, Quebec, Canada; [2]Karine Suissa, PhD: Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

Dr. S. Suissa has received consulting fees and/or speaking fees from Atara, Bristol Myers Squibb, Merck, Pfizer, Seqirus, Boehringer-Ingelheim, and Novartis (less than $10,000 each). Dr. Hudson has received consulting fees from Boehringer-Ingelheim, Alexion, and Octapharma (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Samy Suissa, PhD, Centre for Clinical Epidemiology, Jewish General Hospital, 3755 Cote Ste-Catherine, Montreal, Quebec H3T 1E2, Canada. Email: samy.suissa@mcgill.ca.

Submitted for publication March 30, 2021; accepted in revised form May 20, 2021.

SIGNIFICANCE & INNOVATIONS

- Several observational studies suggest that allopurinol, an effective treatment for gout, also lowers the incidence of cardiovascular (CV) outcomes.
- This study identifies significant time-related biases in these observational studies, biases that tend to exaggerate the potential benefit of treatments.
- These observational studies cannot be used as evidence of decreased CV events with allopurinol use, especially since studies that avoided these biases did not find such a protective effect.
- Allopurinol is very effective at reducing urate levels and the frequency of flares in patients with gout, but it would be confusing at this time for clinicians to suggest that it also improves CV outcomes and mortality.

casts some doubt on the accuracy of the observational studies reporting remarkable effectiveness of allopurinol on CV outcomes. In this study, we conduct a methodologic review of the observational studies of the effect of allopurinol on CV events in patients with gout to evaluate potential sources of bias.

MATERIALS AND METHODS

We searched the literature on March 15, 2021, for all publications of observational studies investigating the effectiveness of allopurinol versus non-use on various CV outcomes. We searched PubMed for keywords "allopurinol" and ("cardiac" or "cardiovascular") and ("cohort" or "observational") with no restriction on publication date. Two independent reviewers systematically screened the titles and abstracts and carried forward potentially relevant publications for review. Subsequently, we also searched Embase with the same keywords to identify any studies that were missed. Articles that were not observational studies, such as comments, editorials, trials, meta-analyses, reviews, and studies on nonhuman subjects, were excluded. The remaining studies were screened by full-text review to exclude articles that did not report a hazard or rate ratio, allopurinol exposure, CV outcomes, or that did not provide data on patients with gout, hyperuricemia, or using urate-lowering treatment (ULT). Additionally, we scanned the references of these papers and included studies that were not identified by our primary search. The final selection was classified according to potential time-related biases induced by the study design. In particular, 2 time-related biases were identified in these studies: time-related confounding and immortal time biases.

Time-related confounding bias arises from a study design based exclusively on cohorts of patients all exposed to allopurinol, without a separate non-user comparator group (Figure 1). Thus, rather than a separate "unexposed" group of patients, this approach compares the allopurinol-exposed episodes with unexposed episodes defined by permanent or transient allopurinol discontinuations within the same group of patients, all users of allopurinol. Thus, typically, a Cox proportional hazards models with time-dependent exposure to allopurinol would be used to estimate the HR comparing "current" allopurinol use with "current" discontinued use. This design can introduce bias from 2 different sources. The first source arises from the absence of adjustment for time-dependent confounders, particularly factors associated with allopurinol discontinuation. Indeed, the effect of allopurinol is based entirely on the comparison between "current" users of allopurinol and those who discontinued this drug at different points of follow-up. Thus, the profile of patients who discontinue at a given point in time could be different from those who continue, potentially introducing confounding.

The second source of time-related confounding bias arises from the study subjects all being exposed to allopurinol during the early period of follow-up, before the first discontinuation. Thus, CV outcome events that occur during this early exposed period will be noninformative; that is, they will not contribute to the estimation of the HR. Indeed, to contribute to the HR, there must be some variation in exposure between the patient with the outcome event and the risk set formed by all members of the cohort present at the time of the event. Figure 1 shows that when patient 1 has the CV event, this patient and all 4 patients in the risk set defined by follow-up time are also exposed to allopurinol, thus not contributing to the HR. In contrast, the other 3 patients with a CV event have varying allopurinol exposure with members of their risk set and thus do contribute to the HR. A consequence of noninformative exposed early events is that early risks or early benefits of the drug will not be captured, biasing the overall HR.

Immortal time bias arises, in cohort studies, from the inclusion of periods of follow-up or exposure during which the outcome under study cannot occur (13,14). Immortal time bias is

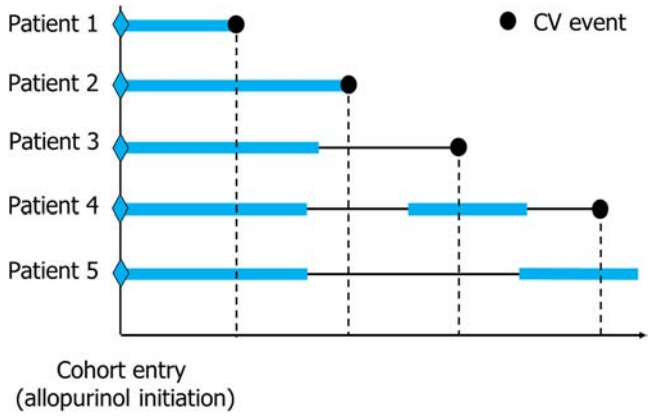

Figure 1. Illustration of time-related confounding bias with 5 typical patients from the study of Weisman et al (20) comparing allopurinol-exposed (blue line) with unexposed (black line) person-moments within risk sets generated by each cardiovascular (CV) event over time (broken lines).

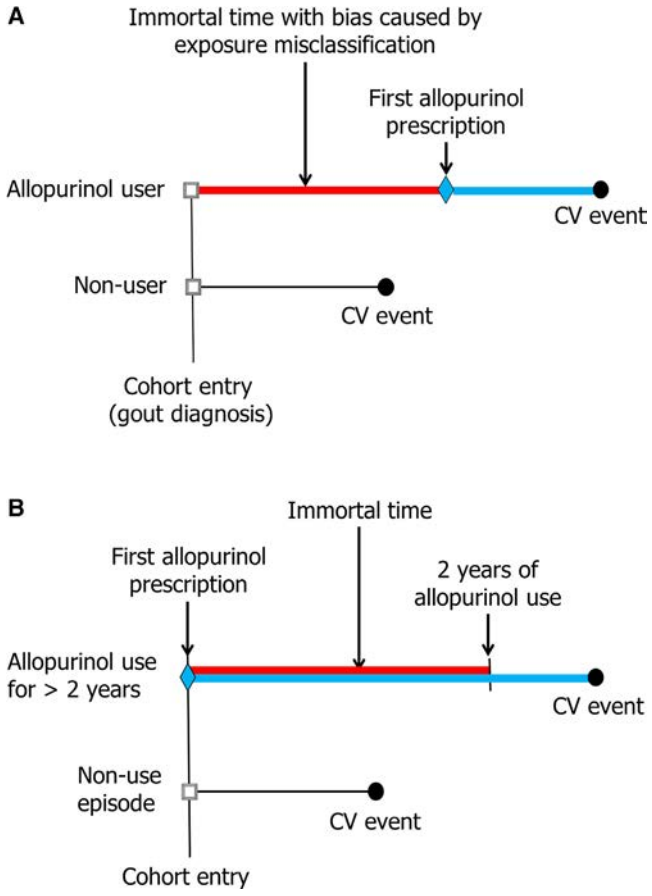

Figure 2. Illustration of immortal time bias in cohort studies according to 2 forms of exposure definition. **A**, The immortal time period between cohort entry (the diagnosis of gout) and the first allopurinol prescription is misclassified as "exposed to allopurinol" when in fact the patient is unexposed; and **B**, The immortal time period between the first allopurinol prescription and the end of >2 years of continuous use is misclassified as exposed to >2 years of allopurinol duration. CV = cardiovascular. Red lines indicate immortal time, blue lines indicate allopurinol-exposed, and black horizontal lines indicate unexposed.

introduced in a cohort study by classifying patients as exposed to allopurinol from the day of cohort entry, even if they only filled their first prescription during follow-up (Figure 2A). This bias can be intensified if the definition of exposure also includes a duration of allopurinol use, which will increase the length of the immortal time, thus also increasing the bias (Figure 2B). Study-specific HRs were pooled using random-effects models stratified according to the type of time-related bias.

RESULTS

We identified 229 potential papers, of which we found the majority to be reviews, editorials or opinion pieces, meta-analyses, observational studies comparing allopurinol against another ULT, or those that did not analyze allopurinol separately among ULTs. We reviewed the full text of 25 studies and further excluded 15 that did not report on exposure to allopurinol, the outcome, or indication of interest (Figure 3). We also identified 2 studies by reference search of published articles, resulting in 12 studies for our methodologic review (Table 1). The search of Embase identified 283 references, which included no new studies beyond the 12 identified via PubMed.

Of these 12 studies, we found 6 studies affected by time-related confounding bias, with a pooled HR of a CV event with allopurinol use of 0.88 (95% CI 0.85–0.92), and 6 by immortal time bias (pooled HR 0.79 [95% CI 0.72–0.87]), including 4 studies contributing to both sources of bias. The remaining 4 studies avoided these time-related biases (pooled HR 1.07 [95% CI 0.91–1.25]) but still reported variable results (Figure 4).

Time-related confounding bias. Time-related confounding bias affected 6 of the identified studies (15–20). For example, the study by Weisman et al used this exposed cohort design in part because "indication bias would be difficult to overcome without important confounder data" (20). This study adjusted for covariates measured at cohort entry (initiation of allopurinol treatment), but these were not updated at the time of each risk set during follow-up, a limitation recognized by the authors ("We were not able to adjust for time-varying comorbidities and medications that may have been related to treatment discontinuation") (20). A particular example of this time-related confounding is adjustment for age at cohort entry but not for age at the time of each risk set if age is categorized or fitted as a nonlinear term. Note that fitting age as a linear term will not change the relative risk estimate with time-dependent adjustment for age. Thus, in a cohort of patients ≥65 years of age at cohort entry and followed up for 20 years, aging during follow-up will be associated with important increases in the risk of CV. Indeed, the incidence rate of CV events increases from 27 to 50 to 70 per 1,000 per year for 65–74, 75–85, and ≥85 years of age, respectively (21). Time-related confounding bias is introduced if age (categorized) is also associated with allopurinol discontinuation, as are several other risk factors that can occur during follow-up (22,23).

With respect to the second source of bias, the studies do not provide the number of CV events that occurred early when all patients were exposed for the time period before the first discontinuation. Thus, the extent of the magnitude of this bias cannot be quantified. Furthermore, there is no solution to avoid this second source of bias, particularly if study subjects are all compliant with treatment for an extended early period of follow-up, thus discarding early outcome events. This study design depends on patients being noncompliant, with the estimate of effect being based on a comparison of continued users with discontinuers. For the first source, the incidence of outcomes such as CV events or mortality will increase with follow-up time, particularly with older study populations. Thus, the analysis should not only be based on a time-varying exposure approach, but confounding factors should

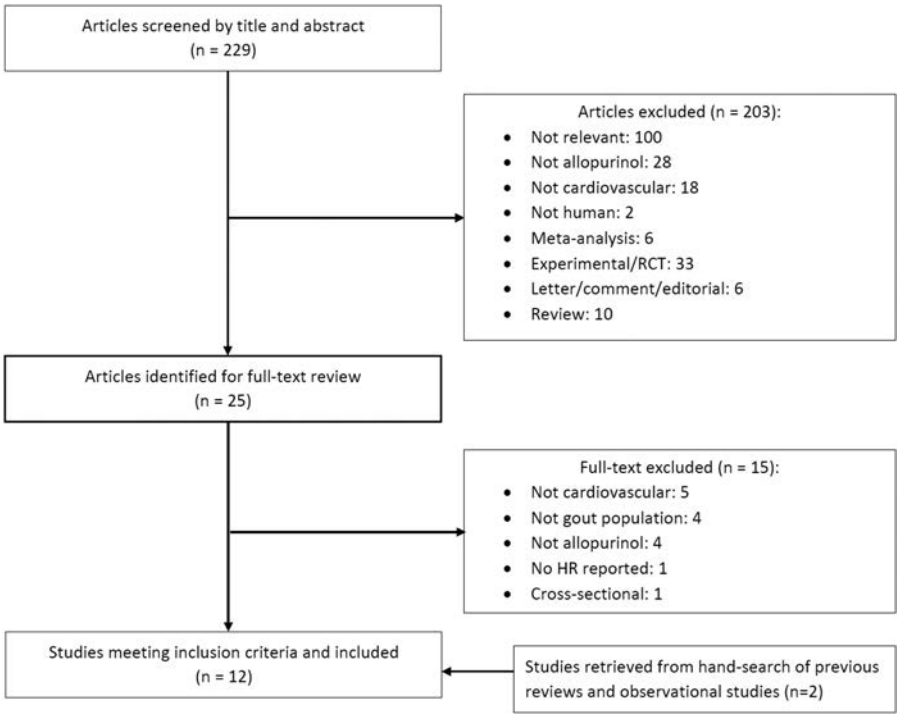

Figure 3. Flow chart for selection of reviewed observational studies. HR = hazard ratio; RCT = randomized controlled trial.

also be updated over time. In this context, some challenging methodologic considerations for studies based on treatment discontinuation should be taken into account (24).

Immortal time bias. Immortal time bias was identified in 6 of the studies (15,16,18,19,25,26). An example of immortal time bias is the study by Wei et al, based on a cohort of >7,000 patients, which reported an HR of a CV event of 0.88 (95% CI 0.73–1.05) with allopurinol use among patients with a urate level of >6 mg/dl (25). This study compared patients who received allopurinol at any time throughout the observation period of 2000–2007 with those who did not receive any ULT. Immortal time bias is introduced in this study by classifying patients as exposed to allopurinol from the day of cohort entry, even if the first prescription was filled during follow-up (Figure 2A). Indeed, the time between cohort entry and the first allopurinol prescription is immortal, as the patient must be alive to receive it, and the patient is unexposed during this period, resulting in immortal time bias (13,14).

Another distinctive example of immortal time bias is the observational study by Singh et al, which investigated the effect of duration of allopurinol treatment on the risk of myocardial infarction (MI), reporting a 30% reduction in this risk with >2 years of allopurinol use (15). The study used Medicare data to identify a cohort of >28,000 initiators of allopurinol, comparing episodes of use with episodes of non-use. The adjusted HR of

Table 1. Observational studies reporting on the risk of a cardiovascular event associated with allopurinol use*

Author, year (ref.)	Patient population	Outcome event	Data source region
Wei et al, 2011 (25)	Population based†	Composite cardiovascular	Scotland
Gotsman et al, 2012 (26)	Chronic heart failure†	Cardiac hospitalization	Israel
Kok et al, 2014 (28)	Gout	Cardiovascular hospitalization	Taiwan
Kim et al, 2015 (29)	Gout	Composite cardiovascular	US
Larsen et al, 2016 (30)	Hyperuricemia	Composite cardiovascular	Denmark
Singh and Yu, 2016 (15)	Receiving allopurinol	Acute myocardial infarction	US
Singh and Yu, 2016 (16)	Receiving allopurinol	Stroke	US
Singh and Yu, 2017 (19)	Receiving allopurinol	Atrial fibrillation	US
Singh et al, 2017 (17)	Diabetes mellitus and gout	Myocardial infarction or stroke	US
Singh and Cleveland, 2018 (18)	Receiving allopurinol	Peripheral arterial disease	US
Weisman et al, 2019 (20)	Diabetes mellitus and allopurinol use	Composite cardiovascular	Ontario, Canada
Ju et al, 2020 (31)	Gout	Composite cardiovascular	Hong Kong

* Ref. = reference.
† Studies also reporting on patients with gout or hyperuricemia.

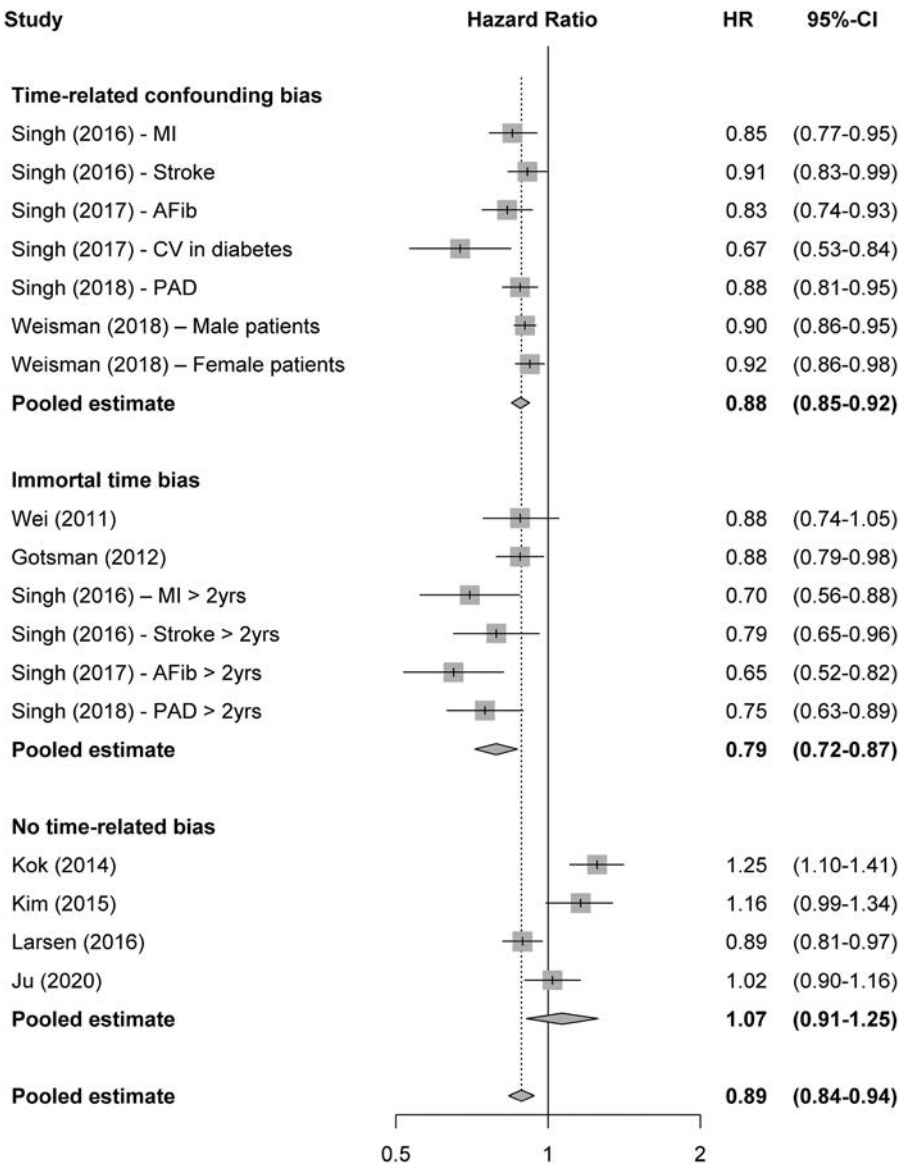

Figure 4. Forest plot of hazard ratios (HRs) and rate ratios of cardiovascular (CV) events associated with allopurinol use from observational studies listed in Table 1, with pooled estimates by a random-effects approach, according to studies affected by time-related confounding time bias, immortal time bias, and no time-related bias. 95% CI = 95% confidence interval; AFib = atrial fibrillation; MI = myocardial infarction; PAD = peripheral arterial disease. Shaded boxes represent individual HRs; diamonds represent pooled HRs and 95% confidence intervals. The broken line indicates the overall HR; horizontal solid lines indicate 95% confidence intervals.

MI associated with >2 years of allopurinol use was 0.70 (95% CI 0.56–0.88). Immortal time bias is introduced by first classifying patients according to duration of use (>2 years) before analyzing the time to MI event. Clearly, the first 2 years of allopurinol use are immortal because the exposed patient must be free of MI during at least 2 years of allopurinol use to be included in the analysis. In contrast, the comparator, unexposed patients, has no such restriction. Figure 2B depicts this bias by comparing the survival times between 2 patients, a user of allopurinol for >2 years, and a non-user, who both incur an MI during follow-up. Clearly, the allopurinol user will have a longer time to MI, artificially created by imposing the >2 years duration of allopurinol use. Counting these first 2 years of allopurinol use as "exposed" leads to immortal time bias, which results in an exaggerated protective effect of allopurinol use compared with non-use (14).

To illustrate the impact of the bias in this study, we used some data reported in the paper on the number of person-days of follow-up according to duration of allopurinol use and simulated others that were not provided (15). The unexposed comparator had 8,327,742 person-days of follow-up and 607 MI cases, for a rate of 73 per 1 million days. On the other hand, the >2 years duration group had 1,847,218 person-days of follow-up and 97 MI cases, for a rate of 53 per 1 million days. However, these person-days include the first 2 years necessary to identify a

Table 2. Comparison between immortal time-biased analysis and corrected analysis for the cohort study by Singh and Yu (15) of >28,000 initiators of allopurinol to assess the effect of >2 years of allopurinol on the risk of acute myocardial infarction (MI), corrected for 2-year immortal time, as illustrated in Figure 2B*

	Acute MI, no.	Person-days	Rate per 1,000,000 person-days	Crude rate ratio (95% CI)
Immortal time-biased analysis				
Allopurinol use >2 years	97	1,847,218	52.5	0.72 (0.58–0.89)
Non-use (ref.)	607	8,327,742	72.9	1.0 (ref.)
Corrected analysis (assuming 1,000 users >2 years)				
Immortal person-time†	0	730,000		
At-risk person-time >2 years	97	1,117,218	86.8	1.19 (0.96–1.48)
Non-use (ref.)	607	8,327,742	72.9	1.0 (ref.)
Corrected analysis (assuming 750 users >2 years)				
Immortal person-time†	0	547,500		
At-risk person-time >2 years	97	1,299,718	74.6	1.02 (0.83–1.27)
Non-use (ref.)	607	8,327,742	72.9	1.0 (ref.)
Corrected analysis (assuming 500 users >2 years)				
Immortal person-time†	0	365,000		
At-risk person-time >2 years	97	1,482,218	65.4	0.90 (0.72–1.11)
Non-use (ref.)	607	8,327,742	72.9	1.0 (ref.)

* 95% CI = 95% confidence interval; ref. = reference.
† Immortal person-time corresponds to the first 2 years of allopurinol use.

duration of allopurinol use of >2 years. Since the paper does not report the number of patients for each duration of use, we assumed 3 values: 1,000, 750, and 500 patients. Table 2 shows that if 1,000 patients had used allopurinol for >2 years, 730,000 person-days of the 1,847,218 person-days of follow-up are immortal and should be subtracted from this denominator (Figure 4B). Consequently, the biased rate ratio of acute MI with >2 years of allopurinol use of 0.72 (95% CI 0.58–0.89) becomes 1.19 (95% CI 0.96–1.48) after correcting for the 2 years of immortal time. With 500 users for >2 years, the corrected rate ratio becomes 0.90 (95% CI 0.72–1.11).

Immortal time bias can be avoided by using a time-dependent definition of exposure that properly classifies exposure in the data analysis, as illustrated in our example using simple rates and proper reclassification of exposed and unexposed person-time (14). Similarly, when assessing the effects of treatment duration, proper classification should be given to the time before which the patient reaches the duration of interest. To address this bias, one can also use statistical models for time-varying exposures, such as the Cox proportional hazards model with time-dependent factors (14). Alternatively, one could use approaches based on study design, such as the prevalent new-user design that matches allopurinol initiators with non-users at the same time point in the disease course, thus avoiding immortal time bias (27).

Studies avoiding time-related biases. We found 4 studies that addressed these time-related biases, all using a separate comparator group (28–31). Two of the studies used a design that started follow-up at initiation of allopurinol exposure, with a comparable time point for the non-users (28,29). For example, the study by Kok et al identified a cohort of patients diagnosed with gout to compare 2,483 users of allopurinol who were matched on several factors to 2,483 non-users, including on the index date of allopurinol initiation, thus avoiding immortal time bias (28). The other 2 studies avoided immortal time bias by using a time-dependent definition for allopurinol use in the regression model (30,31). Indeed, both studies properly classified the immortal time prior to allopurinol initiation as unexposed, after which point patients were classified as exposed. Nonetheless, the current use time-dependent exposure definition to allow for varying adherence to allopurinol treatment, used in one of these studies, is limiting, as it inherently assumes that the risk of the CV events changes immediately to the unexposed level once a patient stops treatment temporarily, and that there is no long-term effect of having been on treatment (30).

DISCUSSION

Our methodologic review of the 12 observational studies of allopurinol effectiveness on major CV events in patients with gout or hyperuricemia noted wide variations in results. We found that 8 of the 12 studies were affected by time-related biases that resulted in significant reductions in the incidence of CV events with allopurinol use. Indeed, the studies affected by time-related confounding bias had a pooled HR of CV events with allopurinol use of 0.88 (95% CI 0.85–0.92), while for those affected by immortal time bias, the HR was 0.79 (95% CI 0.72–0.87). The 4 studies that avoided these biases found variable effects of allopurinol on CV events (pooled HR 1.07 [95% CI 0.91–1.25]).

Time-related confounding bias is rather unique, as it results from cohort studies in which all subjects are exposed to

allopurinol, an unusual design used in several of the reviewed studies. This study design thus relies on patients' tendency to discontinue treatment to permit a comparison of use versus non-use. This design would not be feasible if patients were fully adherent to the medication for the entire cohort follow-up. Nonetheless, even with sufficient discontinuations, this study design presents major challenges for 2 reasons. First, even with long follow-up, there could be sufficient treatment adherence early after initiation such that CV events occurring during this early period would be omitted in the analysis, as all cohort subjects would be exposed during this early period. Thus, early risks or benefits of the study drug would be missed, depending on the duration of early adherence. Second, the data analysis for this design needs to account for time-varying confounders that are associated with both the CV events and the discontinuation of allopurinol treatment, which forms the unexposed comparator in these studies. Nonetheless, this is particularly challenging in this single cohort design, for which determining the comparator of allopurinol discontinuers involves complex design issues (24). Moreover, this approach inherently assumes that the risk of a CV event instantly reverts back to the unexposed risk level once a patient stops the initial allopurinol treatment course and immediately goes back to the previous exposed risk level when the patient restarts, with no long-term effect of having taken allopurinol.

Immortal time bias arising from the misclassification of exposure to allopurinol, which was present in several studies, systematically results in an exaggerated protective effect of allopurinol use compared with non-use (14). This bias is easily avoidable with a proper study design or a data analysis that classifies exposure correctly over time. Interestingly, the immortal time bias that affected the study by Wei et al (25) had been corrected for by the same authors in a prior study of the effect of allopurinol on mortality (32). Immortal time bias was particularly prevalent in several studies investigating the effect of duration of allopurinol on the risk of different CV outcomes (15,16,18,19). These studies, which reported an HR of CV events with >2 years of allopurinol use ranging from 0.65 to 0.79, did not consider the duration of allopurinol use as a time-varying factor. Consequently, as our data illustration showed, an immortal time-biased rate ratio of acute MI with >2 years allopurinol use of 0.72 became 1.19 after correcting for the 2 years of immortal time.

Immortal time bias is most often found in studies comparing users with non-users, which, unlike studies with an active comparator drug, present the challenge of where to start follow-up for the non-users (33). Thus, the presence of immortal time bias should be suspected in studies with non-user comparators, such as the studies of allopurinol use versus non-use that we reviewed, especially if the results suggest a noteworthy benefit for the drug under study. Study designs such as the prevalent new-user design or a marginal structural approach that emulates randomized trials will avoid immortal time bias (27,34).

The 4 studies that avoided these time-related biases all used a separate comparator group, thus circumventing time-related confounding issues. In addition, 2 of the studies avoided immortal time bias using a time-dependent definition for allopurinol use in the data analysis, classifying the immortal time prior to allopurinol initiation as unexposed and exposed subsequently (30,31). The other 2 studies avoided this bias in their study design, matching subjects on the date of initiation of allopurinol and a comparable time point for the non-users (28,29). Nonetheless, 2 of the studies used an intent-to-treat analysis, which estimates the effect of allopurinol irrespective of whether patients continued their treatment over up to 10 years of follow-up (28,31). On-treatment analyses should also be included, as were presented in the other 2 studies, although these studies also found varying effects (29,30).

Observational studies are important to evaluate the real-world effectiveness of medications but must aim to minimize bias. Time-related biases, prevalent in the observational studies of allopurinol on CV events, could be avoided with suitable techniques of study design and data analysis. Along with the biases identified in the observational studies also showing remarkable reductions in mortality with allopurinol use (12), considerable uncertainty surrounds the effectiveness of allopurinol on lowering the incidence of these major outcomes. Thus, our analysis provides methodologic support for the caution of the EULAR task force on the management of gout, particularly the suggestions of earlier treatment because of the conflicting results of these observational studies on CV benefit (9). The currently underway ALL-HEART randomized controlled trial comparing allopurinol with placebo will provide important and accurate evidence on these outcomes (11).

ACKNOWLEDGMENT

The authors thank Dr. Alanna Weisman for responding promptly to queries regarding her study.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. S. Suissa 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. S. Suissa, K. Suissa, Hudson.
Acquisition of data. S. Suissa, K. Suissa.
Analysis and interpretation of data. S. Suissa, K. Suissa, Hudson.

REFERENCES

1. Clarson LE, Chandratre P, Hider SL, Belcher J, Heneghan C, Roddy E, et al. Increased cardiovascular mortality associated with gout: a systematic review and meta-analysis. Eur J Prev Cardiol 2015;22:335–43.
2. Gupta MK, Singh JA. Cardiovascular disease in gout and the protective effect of treatments including urate-lowering therapy. Drugs 2019;79:531–41.

3. Feig DI, Kang DH, Johnson RJ. Uric acid and cardiovascular risk. N Engl J Med 2008;359:1811–21.
4. Gaffo AL, Edwards NL, Saag KG. Gout. Hyperuricemia and cardiovascular disease: how strong is the evidence for a causal link? Arthritis Res Ther 2009;11:240.
5. FitzGerald JD, Dalbeth N, Mikuls T, Brignardello-Petersen R, Guyatt G, Abeles AM, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Care Res (Hoboken) 2020; 72:744–60.
6. Bredemeier M, Lopes LM, Eisenreich MA, Hickmann S, Bongiorno GK, d'Avila R, et al. Xanthine oxidase inhibitors for prevention of cardiovascular events: a systematic review and meta-analysis of randomized controlled trials. BMC Cardiovasc Disord 2018;18:24.
7. Rosas G, Gaffo A, Rahn EJ, Saag KG. Cardiovascular safety risks associated with gout treatments. Expert Opin Drug Saf 2020;19: 1143–54.
8. Schlesinger N, Brunetti L. Beyond urate lowering: analgesic and anti-inflammatory properties of allopurinol. Semin Arthritis Rheum 2020; 50:444–0.
9. Richette P, Doherty M, Pascual E, Barskova V, Becce F, Castañeda-Sanabria J, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis 2017;76:29–42.
10. Grainger R, Taylor WJ. Allopurinol and peripheral vascular disease: enough observational data to warrant interventional studies. Rheumatology (Oxford) 2017;57:408–9.
11. Mackenzie IS, Ford I, Walker A, Hawkey C, Begg A, Avery A, et al, on behalf of the ALL-HEART study group. Multicentre, prospective, randomised, open-label, blinded end point trial of the efficacy of allopurinol therapy in improving cardiovascular outcomes in patients with ischaemic heart disease: protocol of the ALL-HEART study. BMJ Open 2016;6:e013774.
12. Suissa S, Suissa K, Hudson M. Effectiveness of allopurinol in reducing mortality: time-related biases in observational studies. Arthritis Rheumatol 2021;73:1749–57.
13. Suissa S. Immortal time bias in observational studies of drug effects. Pharmacoepidemiol Drug Saf 2007;16:241–9.
14. Suissa S. Immortal time bias in pharmacoepidemiology. Am J Epidemiol 2008;167:492–9.
15. Singh JA, Yu S. Allopurinol reduces the risk of myocardial infarction (MI) in the elderly: a study of Medicare claims. Arthritis Res Ther 2016;18:209.
16. Singh JA, Yu S. Allopurinol and the risk of stroke in older adults receiving Medicare. BMC Neurology 2016;16:164.
17. Singh JA, Ramachandaran R, Yu S, Curtis JR. Allopurinol use and the risk of acute cardiovascular events in patients with gout and diabetes. BMC Cardiovasc Disord 2017;17:76–6.
18. Singh JA, Cleveland J. Allopurinol and the risk of incident peripheral arterial disease in the elderly: a US Medicare claims data study. Rheumatology (Oxford) 2018;57:451–61.
19. Singh JA, Yu S. Allopurinol and the risk of atrial fibrillation in the elderly: a study using Medicare data. Ann Rheum Dis 2017;76:72–8.
20. Weisman A, Tomlinson GA, Lipscombe LL, Perkins BA, Hawker GA. Association between allopurinol and cardiovascular outcomes and all-cause mortality in diabetes: a retrospective, population-based cohort study. Diabetes Obes Metab 2019;21:1322–9.
21. Yazdanyar A, Newman AB. The burden of cardiovascular disease in the elderly: morbidity, mortality, and costs. Clin Geriatr Med 2009; 25:563–77, vii.
22. Solomon DH, Avorn J, Levin R, Brookhart MA. Uric acid lowering therapy: prescribing patterns in a large cohort of older adults. Ann Rheum Dis 2008;67:609–13.
23. Scheepers LE, van Onna M, Stehouwer CD, Singh JA, Arts IC, Boonen A. Medication adherence among patients with gout: a systematic review and meta-analysis. Semin Arthritis Rheum 2018;47: 689–702.
24. Krumme AA, Glynn RJ, Schneeweiss S, Choudhry NK, Tong AY, Gagne JJ. Defining exposure in observational studies comparing outcomes of treatment discontinuation. Circ Cardiovasc Qual Outcomes 2018;11:e004684.
25. Wei L, Mackenzie IS, Chen Y, Struthers AD, MacDonald TM. Impact of allopurinol use on urate concentration and cardiovascular outcome. Br J Clin Pharmacol 2011;71:600–7.
26. Gotsman I, Keren A, Lotan C, Zwas DR. Changes in uric acid levels and allopurinol use in chronic heart failure: association with improved survival. J Card Fail 2012;18:694–701.
27. Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. Pharmacoepidemiol Drug Saf 2017;26:459–68.
28. Kok VC, Horng JT, Chang WS, Hong YF, Chang TH. Allopurinol therapy in gout patients does not associate with beneficial cardiovascular outcomes: a population-based matched-cohort study. PLoS One 2014;9:e99102.
29. Kim SC, Schneeweiss S, Choudhry N, Liu J, Glynn RJ, Solomon DH. Effects of xanthine oxidase inhibitors on cardiovascular disease in patients with gout: a cohort study. Am J Med 2015;128:653.e7–16.
30. Larsen KS, Pottegård A, Lindegaard HM, Hallas J. Effect of allopurinol on cardiovascular outcomes in hyperuricemic patients: a cohort study. Am J Med 2016;129:299–306.e2.
31. Ju C, Lai RW, Li KH, Hung JK, Lai JC, Ho J, et al. Comparative cardiovascular risk in users versus non-users of xanthine oxidase inhibitors and febuxostat versus allopurinol users. Rheumatology (Oxford) 2020;59:2340–9.
32. Wei L, Fahey T, Struthers AD, MacDonald TM. Association between allopurinol and mortality in heart failure patients: a long-term follow-up study. Int J Clin Pract 2009;63:1327–33.
33. Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology. Pharmacoepidemiol Drug Saf 2020;29:1101–10.
34. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol 2016;79:70–5.