In this Issue

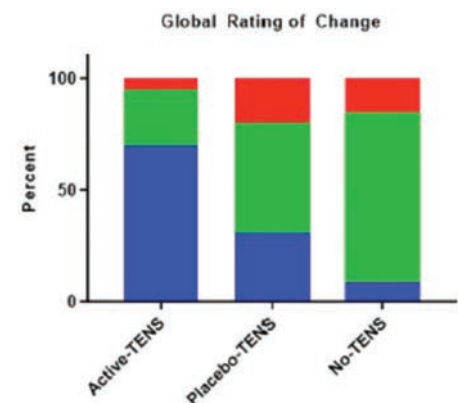## Highlights from this issue of *A&R* | By Lara C. Pullen, PhD

Transcutaneous Electrical Nerve Stimulation Effective in Patients with Fibromyalgia

Patients with fibromyalgia (FM) experience pain and fatigue, especially during physical activity. Transcutaneous electrical nerve stimulation (TENS) activates endogenous pain inhibitory mechanisms and has been proposed as a treatment option for FM. In this issue, Dailey et al (p. 824) report the results of their double-blind randomized controlled trial designed to investigate whether using TENS during activity could improve movement-evoked pain, and look at other patient-reported outcomes.

p. 824

For their study, the investigators included women with FM who were on a stable medication regimen. They randomly assigned participants to receive active TENS (m = 103), placebo TENS (n = 99), or no TENS (n = 99), and instructed the women to use the device at home, during activity, for 2 hours each day over a 4-week period. Patients applied TENS to their lumbar and cervicothoracic regions using a modulated frequency (2–125 Hz) at the highest tolerable intensity. Patients rated their movement-evoked pain and fatigue on an 11-point scale before and during application of TENS, at baseline, and at 4 weeks.

The researchers found that 4 weeks of active TENS use resulted in significant improvement in movement-evoked pain and movement-evoked fatigue compared to placebo TENS or no TENS. A greater percentage (70%) of the patients in the active TENS group reported improvement on the global impression of change (GIC) compared to the placebo TENS group (31%) and the no TENS group (9%). The authors suspect that the reductions in pain and fatigue likely contributed to the improvements in GIC reported by individuals who received active TENS. The investigators identified no serious adverse events related to TENS, and <5% of participants experienced minor adverse events from TENS.

Figure 1. The percentage of participants who reported feeling better or much better (blue), no change (green), and worse or extremely worse (red) after 4 weeks of active TENS, placebo TENS, or no TENS treatment.

Asthma and COPD Associated with Increased Risk for RA

Past research has suggested that asthma increases the risk for rheumatoid arthritis (RA). The reason for this, is it believed, is that inflamed airways contribute the RA-related autoantibody production. Moreover, rheumatologists have identified smoking as the strongest environmental risk factor for RA.

p. 704

In this issue, Ford et al (p. 704) report on the association between asthma and chronic obstructive pulmonary disease (COPD) and an elevated risk of incident RA. The investigators performed a large prospective cohort study with long follow-up and found that asthma was associated with a >50% increase in the risk of subsequent RA compared to no asthma/COPD.

The associations were independent of smoking status/intensity and other potential confounders. The results are consistent with the hypothesis that chronic airway inflammation may be crucial in RA pathogenesis.

The study utilized data derived from the Nurses' Health Study (NHS, 1988–2014) and NHSII (1991–2015) to identify 15,148 women with confirmed asthma, 3,573 women with confirmed COPD, and 1,060 incident RA cases during 4,384,471 person-years of follow-up. After adjusting for covariates, the investigators found that asthma was associated with increased RA risk (hazard ratio [HR] 1.53 [95% confidence interval (95% CI) 1.24–1.88]) compared to no asthma/COPD.

Asthma remained associated with increased RA among never-smokers only (HR 1.53 [95% CI 1.14–2.05]). The researchers found that COPD was also associated with an increased RA risk (HR 1.89 [95% CI 1.31–2.75]), with the association of COPD with RA being most pronounced in the subgroup of ever-smokers age >55 years (HR 2.20 [95% CI 1.38–3.51]).

Ustekinumab Safe and Effective in SLE Phase II Trial

Although research has delivered advances in treatment of systemic lupus erythematosus (SLE), over time patients continue to experience progressive organ damage, including pulmonary, cardiovascular, and renal disease, and to suffer high rates of morbidity. In this issue, van Vollenhoven et al (p. 761) report the results of a phase II prospective trial designed to evaluate the efficacy and safety of ustekinumab through 1 year in patients with SLE. The researchers found that ustekinumab provided sustained clinical benefit in patients.

This study included patients who had clinically active SLE despite standard background therapy. The investigators defined active disease as having an SLE Disease Activity Index 2000 (SLEDAI-2K) score of ≥6 as well as ≥1 British Isles Lupus Assessment Group (BILAG) A organ domain score and/or ≥2 BILAG B organ domain scores at screening. The investigators randomized (3:2) the 102 patients to receive either ustekinumab (~6 mg/kg of single intravenous infusion at week 0, then 90 mg of subcutaneous injections every 8 weeks beginning at week 8) or a matching placebo added to standard therapy. At week 24, the placebo group crossed over to receive a subcutaneous 90-mg dose of ustekinumab every 8 weeks, and the original ustekinumab group continued to receive therapy through week 40. The researchers measured efficacy using the SLEDAI-2K, the SLE Responder Index 4 (SRI-4), physician global assessment, and mucocutaneous and joint disease measures in a modified intent-to-treat population.

The investigators found that a greater proportion of ustekinumab-treated patients had improvements in clinical efficacy assessments at week 24, compared to patients in the placebo group. The SRI-4 response rates were 62% in the ustekinumab group and 33% in the placebo group at the week 24 primary end point analysis. These response rates were maintained at week 48. Patients in the placebo group who crossed over to the ustekinumab group experienced increased response rates across efficacy measures.

The researchers observed no deaths, malignancies, opportunistic infections, or tuberculosis cases. Among all ustekinumab-treated patients, 81.7% had ≥1 adverse event (AE) and 15.1% had ≥1 serious AE through week 56. The AEs reported through week 56 were consistent with those reported through week 24 and were consistent with the safety profile of ustekinumab's other indications.

B Cell Synovitis More Common in Established RA

Previous studies have indicated that ~40% of rheumatoid arthritis (RA) patients (both early and late-stage disease) present with synovitis characterized by a B cell–rich infiltrate. In this issue, Rivellese et al (p. 714) describe the results of their efforts to define the relationship of synovial B cells to clinical phenotypes at different stages of disease evolution and drug exposure. They did this by exploring the prevalence of B cell synovitis in a large, well-characterized cohort of 329 patients.

The authors describe in their paper a robust semiquantitative histologic B cell score that can closely replicate the quantification of B cells by digital or molecular analyses. The investigators first established that semiquantitative B cell scores correlated with digital image analysis and B cell lineage–specific transcripts. When they used the approach to analyze patients with established RA and early RA, they found an ongoing B cell–rich synovitis in a larger proportion of patients with established RA than early RA. They note in their paper that this synovitis does not seem to be captured by standard clinimetric assessment.

The researchers report that B cell–rich synovitis was present in 35% of the cohort of patients with early RA and 47.7% of the cohort of patients with established RA with an inadequate response to tumor necrosis factor inhibitors (TNFi). In addition, early RA patients, but not established RA patients, who were B cell rich showed higher levels of disease activity and seropositivity for rheumatoid factor and anti–citrullinated protein antibody. In contrast, the researchers found that both cohorts had significantly higher histologic synovitis scores in B cell–rich patients.
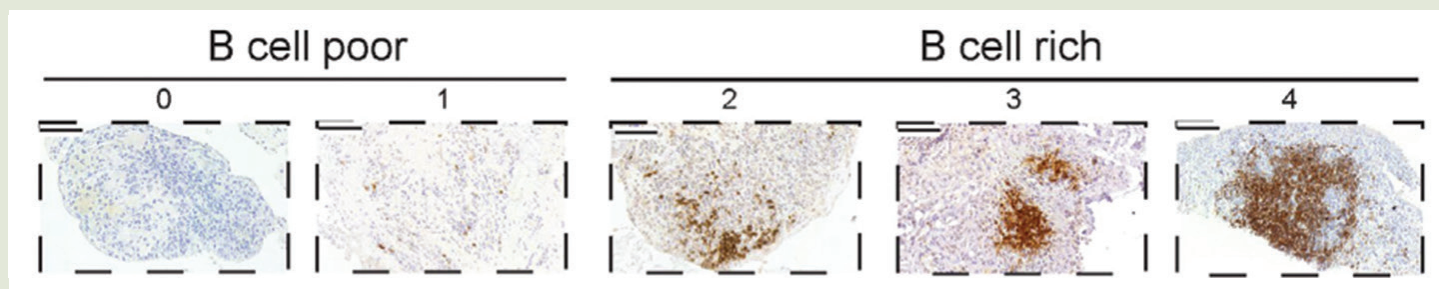
Figure 1. B cell scoring. High-magnification views of representative images for each semiquantitative CD20 immunohistochemical score in RA patients with an inadequate response to TNFi. Bars = 100 μm.

Clinical Connections

Negative Regulation of Osteoclast Commitment by Intracellular Protein Phosphatase Magnesium-Dependent 1A

Kwon et al, *Arthritis Rheumatol* 2020;76:750–760

CORRESPONDENCE
Yong-Gil Kim, MD, PhD: bestmd2000@amc.seoul.kr
Eun-Ju Chang, PhD: ejchang@amc.seoul.kr

SUMMARY
Spondyloarthritis (SpA) is characterized by new bone formation leading to bony ankylosis. However, increased bone resorption in axial SpA is also evident, with decreased bone mineral density. Protein phosphatase magnesium-dependent 1A (PPM1A) has been previously identified as a protein involved in osteoblast differentiation in patients with ankylosing spondylitis. To gain insight into whether PPM1A is also involved in osteoclast differentiation, Kwon et al used macrophage-specific PPM1A[fl/fl] mice and demonstrated enhanced osteoclast differentiation and a bone resorptive phenotype in these mice. The intracellular expression of PPM1A in macrophages from wild-type mice was decreased under inflammatory conditions, indicating that inflammation may enhance osteoclast differentiation by down-regulating PPM1A. PPM1A expression in the peripheral blood mononuclear cells (PBMCs) of patients with SpA also showed an inverse correlation with disease activity. These findings suggest a link between PPM1A and bone resorption in the setting of inflammation and that PPM1A may be a marker of bone metabolism in SpA.
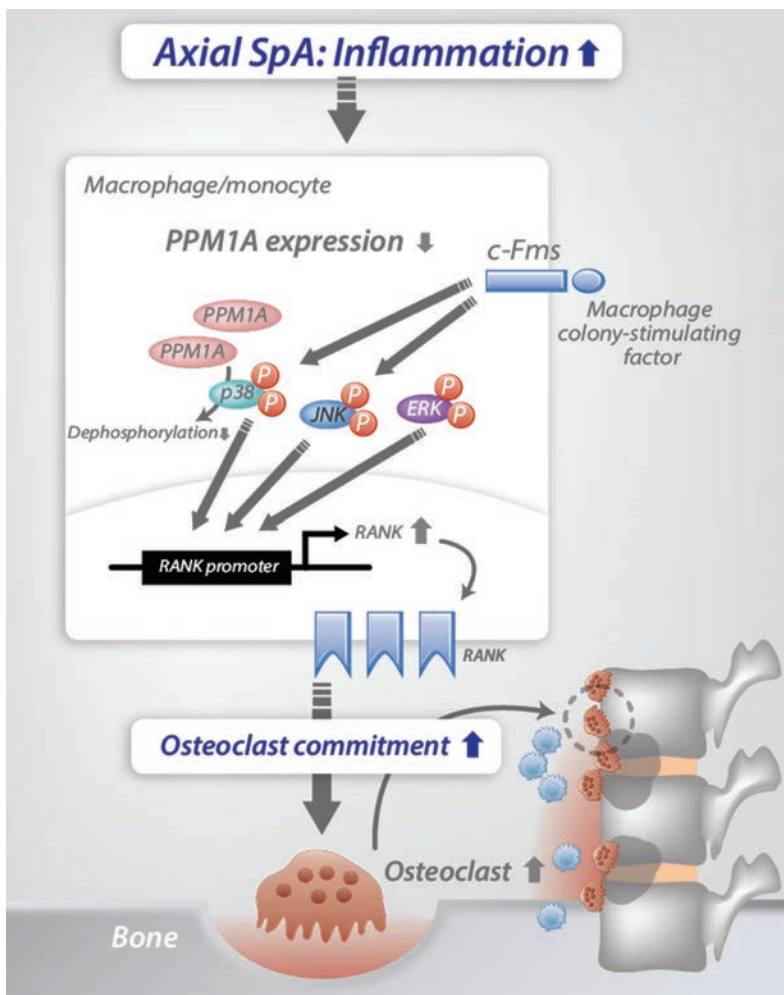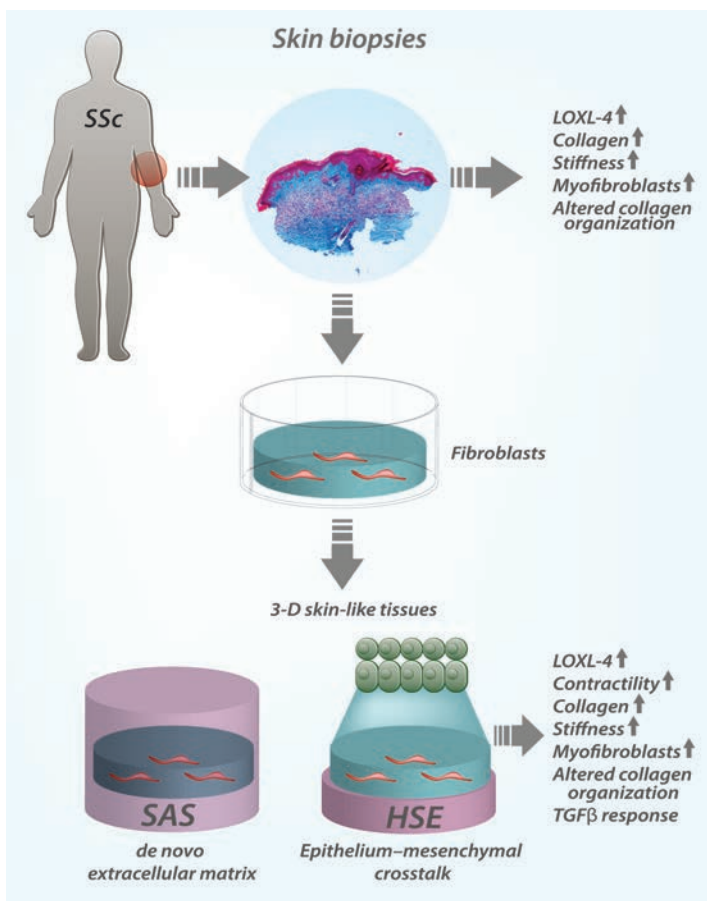
KEY POINTS

• Under inflammatory conditions, PPM1A in macrophages is down-regulated and associated with enhanced osteoclast commitment.

• Intracellular PPM1A expression in PBMCs correlates negatively with disease activity in SpA.

• Depending on inflammatory burden in SpA, the alteration of intracellular PPM1A may shift bone homeostasis toward a resorptive phase.

Systemic Sclerosis Dermal Fibroblasts Induce Cutaneous Fibrosis Through Lysyl Oxidase–like 4: New Evidence From Three-Dimensional Skin-like Tissues

CORRESPONDENCE
Jonathan Garlick, PhD: jonathan.garlick@tufts.edu

KEY POINTS

- Up-regulation of LOXL-4 is found in SSc patient skin, specifically in SSc patients with inflammatory and proliferative features as identified by an intrinsic subsets classifier.

- Skin-like 3-D tissues, fabricated and populated with SSc fibroblasts, were enriched in *LOXL4* expression and recapitulate the cutaneous fibrotic features in vitro.

- TGFβ-induced collagen deposition was blocked by suppression of *LOXL4* in in both normal and SSc fibroblasts.

SUMMARY

Lysyl oxidase–like 4 (LOXL-4) is a member of the lysyl oxidase family of extracellular enzymes that play an important role in assembling the collagen fibrils needed for mechanical stability in connective tissues. Huang et al found that LOXL-4 was up-regulated in systemic sclerosis (SSc) skin biopsies and particularly, in the inflammatory and proliferative intrinsic patient subsets identified by their previously published classifier. These findings further elucidate the heterogeneity of SSc patients at the molecular level. In order to investigate the potential role of LOXL4 in the pathogenesis of SSc, the researchers adapted 2 complimentary 3-dimensional (3-D) skin-like tissues (self-assembled stromal tissue [SAS] and human skin equivalents [HSEs]). These have both been well characterized to demonstrate hallmarks of tissue fibrosis in vitro, as indicated by increased stroma deposition and stiffness, altered collagen modification, and interleukin-6 secretion in constructs that incorporated SSc dermal fibroblasts (SScDFs; n = 8) compared to normal dermal fibroblasts (NDFs; n = 6). Conversely, loss of function of LOXL-4 was associated with decreased stromal rigidity, tissue contraction, and α–smooth actin muscle+ myofibroblasts.

Transforming growth factor β (TGFβ)–altered collagen modification was decreased by suppression of *LOXL4* in both SScDFS and NDFs. Collectively, this study suggests a role for LOXL-4 as a mediator that drives fibrosis in a biologically relevant, in vitro tissue model that can help bridge the gap between existing animal models of SSc and findings in SSc patient skin.

ACR Announcements

ACR Meetings

Annual Meeting
  November 6–11, 2020, Washington, DC
  November 5–10, 2021, San Francisco

For additional information, visit rheumatology.org/Learning-Center/Educational-Activities.

Applications Invited for *Arthritis Care & Research* Editor-in-Chief (2021–2026 Term)

The American College of Rheumatology Committee on Journal Publications announces the search for the position of Editor, *Arthritis Care & Research*. The official term of the next *Arthritis Care & Research* editorship is July 1, 2021–June 30, 2026; however, some of the duties of the new Editor will begin during a transition period starting April 1, 2021. ACR/ARP members who are considering applying for this prestigious and rewarding position should submit a nonbinding letter of intent by May 4, 2020 to the Managing Editor, Maggie Parry, at mparry@rheumatology.org, and are also encouraged to contact the current Editor-in-Chief, Dr. Marian Hannan, to discuss details. Initial contact should be made via e-mail to Hannan@hsl.harvard.edu. Applications will be due by June 15, 2020 and will be reviewed during the summer of 2020. Application materials are available on the ACR web site at https://www.rheumatology.org/Learning-Center/Publications-Communications/Journals/AC-R.

Nominations for ACR Awards of Distinction and Masters Due May 15

The ACR has many Awards of Distinction, including the Presidential Gold Medal. Members who wish to nominate a colleague or mentor for an Award of Distinction must complete the online form at www.rheumatology.org by May 15, 2020. The nomination process includes a letter of nomination, 2 additional letters of recommendation, and a copy of the nominee's curriculum vitae. Recognition as a Master of the American College of Rheumatology is one of the highest honors the ACR bestows. The designation of Master is conferred on ACR members age 65 or older who have made outstanding contributions to the field of rheumatology through scholarly achievements and/or service to their patients, students, and the profession. To nominate someone for a Master designation, members must complete the online nomination form at www.rheumatology.org and include a letter of nomination, 2 supporting letters from voting members of the ACR, and the nominee's curriculum vitae. Nominees for ACR Master must have reached the age of 65 before October 1, 2020

ACR Invites Nominations for Volunteer Positions

The ACR encourages all members to participate in forming policy and conducting activities by assuming positions of leadership in the organization. Positions are available in all areas of the work of the American College of Rheumatology and the Rheumatology Research Foundation. Please visit www.rheumatology.org for information about nominating yourself or a colleague for a volunteer position with the College. The deadline for volunteer nominations is June 1, 2020. Letters of recommendation are not required but are preferred.

Arthritis & Rheumatology

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

Arthritis & Rheumatology

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

VOLUME 72 • May 2020 • NO. 5

Cover image: The figure on the cover (from Berntsen et al, pages 837–848) shows the muscle fiber composition (types I and II muscle fibers) and relative size of the muscle fibers in 7 patients with active juvenile dermatomyositis (left) and 5 patients with inactive juvenile dermatomyositis (right) after a combined mean disease duration of almost 26 years. Type I muscle fibers are depicted in blue, and type II muscle fibers are depicted in brown.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

EDITORIAL

Form, Function, and Dysfunction: Airway Diseases Are Associated With Increased Risk for Rheumatoid Arthritis

Maor Sauler [iD]

Whether it be the sweeping eagle in his flight, or the open apple-blossom, the toiling work-horse, the blithe swan, the branching oak, the winding stream at its base, the drifting clouds, over all the coursing sun, form ever follows function, and this is the law.

Louis Sullivan

Rheumatoid arthritis (RA) risk is strongly influenced by environmental exposures, and the organ with the largest interface with our environment is the lung. To effectively exchange oxygen for CO_2, the lung must support a delicately thin mucosal border (2 μm) spread over 140 m² (1). While ideal for gas exchange, the immensity of the lung's surface area makes it challenging to mitigate threats from inhaled pathogens and toxins while simultaneously promoting immunologic tolerance to harmless epitopes. A failure to maintain homeostatic host–environment interactions along this mucosal border is implicated in the pathogenesis of many lung diseases, and growing evidence suggests it is also implicated in the pathogenesis of RA.

Pulmonary diseases occur commonly in RA. Clinically significant interstitial lung disease (ILD) occurs in 5–10% of patients with RA, and subclinical ILD can be detected in 30–50% of patients with RA (2). Other pulmonary manifestations of RA include pleural effusions, bronchiectasis, pulmonary hypertension, pulmonary nodules, and both follicular and constrictive bronchiolitis. Recently, RA was also identified as a risk factor for chronic obstructive pulmonary disease (COPD) (3). The biologic mechanisms underlying RA-associated lung disease are poorly understood but they likely involve systemic autoimmunity and inflammation. Overexpression of tumor necrosis factor in mice causes both joint inflammation and lung disease, while in an SKG murine model of RA, the development of pulmonary disease requires a "second-hit" lung injury (e.g., cigarette smoke or bleomycin) (4,5). Additionally, autoimmunity is now implicated as a cause of lung destruction in COPD (6).

However, lung disease may not simply be a consequence of RA, but also a cause of RA pathogenesis. Epidemiologic studies have identified cigarette smoking as a major risk factor for RA, and the relative risk for RA conveyed by certain HLA–DRB1 polymorphisms is synergistic with a history of smoking (7). Smokers have increased expression of peptidylarginine deiminase enzymes, which citrullinate proteins through the conversion of arginine to citrulline (8). In turn, these citrullinated proteins can trigger formation of anti–citrullinated protein autoantibodies (ACPAs) that are highly specific to RA (9). Citrullinated proteins and ACPAs can be found at higher concentrations in bronchoalveolar fluid than blood in individuals with early RA (8,10), and RA is also associated with the presence of inducible bronchus-associated lymphoid tissue, which contains ACPA-producing B cells (11). Airway inflammation can also be identified in subjects with early ACPA-positive RA as well as those who have elevated serum levels of ACPA and/or rheumatoid factor (RF) but do not have inflammatory arthritis (10,12,13). In addition, ACPAs and RF can be found in sputum from individuals at risk for future RA (14,15). The risk of RA due to cigarette smoke is likely a consequence of pulmonary injury and inflammation rather than a direct effect of cigarette smoke. Inflammation is a common cause of citrullination in many tissues, and many volatile environmental exposures (not just cigarette smoke) increase RA risk, including mineral and textile dust (9,16). In another study, ACPA titers were higher in smokers with COPD than those without COPD (17). Additionally, individuals with COPD from $α_1$-antitrypsin deficiency have an increased risk for elevated ACPA titers, but among these individuals, there is no association between ACPA titers and smoking (18).

If indeed pulmonary injury and inflammation are key steps in RA pathogenesis, then it may be hypothesized that chronic inflammatory airway diseases like COPD and asthma should also increase RA risk. Yet there have been no studies showing COPD as an RA risk factor, and the evidence for asthma as an RA risk factor has been limited until now. However, in this issue of *Arthritis & Rheumatology*, Ford and colleagues provide

Address correspondence to Maor Sauler, MD, 300 Cedar Street, New Haven, CT 06519. E-mail: maor.sauler@yale.edu.

compelling evidence that both asthma and COPD increase the risk of RA, independent of smoking history (19). They used Cox proportional hazard models to assess risk of incident RA in a prospective cohort study of 205,153 women in the Nurses' Health Study. Participants were categorized by exposure (asthma or COPD) as determined by self-reported physician diagnosis of disease in combination with validated questionnaires. The outcome measure was RA confirmed by medical record review by 2 rheumatologists. Seropositive RA was determined by the presence of elevated RF or ACPAs. Importantly, their multivariable model was adjusted for potential confounders including smoking status (current/former/never) and duration/intensity (pack-years), and passive smoke exposure. Among women with COPD, the multivariable-adjusted hazard ratio (HR) for developing RA was 1.89 (95% confidence interval [95% CI] 1.31–2.75). Interestingly, there was an increased risk for seropositive RA (HR 2.07 [95% CI 1.31–3.25]) but not seronegative RA (HR 1.59 [95% CI 0.83–3.05]). Because COPD commonly occurs among older individuals, Ford et al evaluated a subgroup of ever-smokers age >55 years and identified a stronger association between COPD and seropositive RA (HR 2.85 [95% CI 1.63–4.99]). Among women with a diagnosis of asthma, the multivariable-adjusted HR for developing RA was 1.53 (95% CI 1.24–1.88). Asthma was associated with both seropositive RA and seronegative RA, but interestingly, the presence of asthma in women who never smoked was associated with seronegative RA (HR 1.90 [95% CI 1.22–2.96]) but not with seropositive RA (HR 1.32 [95% CI 0.88–1.96]).

The study conducted by Ford et al was well designed, and their use of a large and thoroughly characterized cohort of subjects allowed them to demonstrate that asthma and COPD are risk factors for RA. Naturally, the study also raises important questions. Only women were included, and therefore the findings may not be generalizable to men. There was also no assessment of medication use for asthma or COPD (e.g., inhaled corticosteroids), and it would be important to understand the role of medications in decreasing or increasing subsequent RA risk. Additionally, while validated tools were used to identify the presence of asthma or COPD, both diseases are remarkably complex and heterogeneous. COPD involves multiple pathologic manifestations, including emphysema, small airway disease, mucous metaplasia, and vascular remodeling. These pathologic features occur in variable degrees among affected individuals, and therefore current preclinical and clinical assessments of COPD include measurements of airflow obstruction, percentage of radiographic emphysema, respiratory diffusion capacity, patient-reported symptoms, and emerging biomarkers (20). Similarly, asthma is a complex disease involving various subtypes (i.e., endotypes). While some individuals with asthma have classic features of atopy and Th2 inflammation, other subtypes of asthma are more closely associated with the presence of obesity, neutrophilic inflammation, and/or exposure to volatile pathogens and toxins (e.g., cigarette

smoke) (21). There is also an increasing recognition of a subgroup of patients with asthma/COPD overlap syndrome. While poorly defined, patients with this condition have features of both diseases. Future studies will require thorough phenotyping of COPD and asthma to more completely understand their relationships to RA.

The need for more extensive phenotyping may be particularly important for asthma because of the intriguing finding by Ford et al that asthma and COPD have different associations with RA seropositivity (19). The authors identified COPD as a risk factor for seropositive RA but not seronegative RA, a finding rooted in mechanistic evidence as described above. In contrast, Ford and colleagues identified asthma as a risk factor for seronegative RA but not seropositive RA among never-smokers (19), and the cause of increased seronegative RA in this population can only be speculated. What are the inflammatory profiles and mechanisms underlying this relationship? Is the relationship strongest among those with atopic asthma? Are specific exposures or genetic associations involved? Understanding how asthma subtypes relate to seronegative RA may lead to the identification of yet-undiscovered mechanisms and treatments for RA.

Despite the limitations of this study, the findings have the potential to immediately impact the care of patients with asthma and COPD. Awareness of the heightened risk for RA in individuals with asthma and COPD can lead to reduced time to diagnosis and earlier initiation of treatment. In nature, form follows function. Ford and colleagues remind us that the lung's form also exposes it to environmental factors that in turn contribute to diseases both inside and outside the lung.

AUTHOR CONTRIBUTIONS

Dr. Sauler drafted the article, revised it critically for intellectual content, and approved the final version to be published.

REFERENCES

1. Gehr P, Bachofen M, Weibel ER. The normal human lung: ultrastructure and morphometric estimation of diffusion capacity. Respir Physiol 1978;32:121–40.

2. Fischer A, Strek ME, Cottin V, Dellaripa PF, Bernstein EJ, Brown KK, et al. Proceedings of the American College of Rheumatology/Association of Physicians of Great Britain and Ireland connective tissue disease–associated interstitial lung disease summit: a multidisciplinary approach to address challenges and opportunities. Arthritis Rheumatol 2019;71:182–95.

3. Sparks JA, Lin TC, Camargo CA Jr, Barbhaiya M, Tedeschi SK, Costenbader KH, et al. Rheumatoid arthritis and risk of chronic obstructive pulmonary disease or asthma among women: a marginal structural model analysis in the Nurses' Health Study. Semin Arthritis Rheum 2018;47:639–48.

4. Wu EK, Henkes ZI, McGowan B, Bell RD, Velez MJ, Livingstone AM, et al. TNF-induced interstitial lung disease in a murine arthritis model: accumulation of activated monocytes, conventional dendritic cells, and CD21$^+$/CD23$^-$ B cell follicles is prevented with anti-TNF therapy. J Immunol 2019;203:2837–49.

5. Keith RC, Powers JL, Redente EF, Sergew A, Martin RJ, Gizinski A, et al. A novel model of rheumatoid arthritis-associated interstitial lung disease in SKG mice. Exp Lung Res 2012;38:55–66.

6. Nunez B, Sauleda J, Antó JM, Julià MR, Orozco M, Monsó E, et al. Anti-tissue antibodies are related to lung function in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2011;183:1025–31.

7. Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L. A gene–environment interaction between smoking and shared epitope genes in HLA–DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 2004;50:3085–92.

8. Klareskog L, Stolt P, Lundberg K, Källberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA–DR (shared epitope)–restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006;54:38–46.

9. Catrina AI, Ytterberg AJ, Reynisdottir G, Malmström V, Klareskog L. Lungs, joints and immunity against citrullinated proteins in rheumatoid arthritis. Nat Rev Rheumatol 2014;10:645–53.

10. Reynisdottir G, Karimi R, Joshua V, Olsen H, Hensvold AH, Harju A, et al. Structural changes and antibody enrichment in the lungs are early features of anti–citrullinated protein antibody–positive rheumatoid arthritis. Arthritis Rheumatol 2014;66:31–9.

11. Rangel-Moreno J, Hartson L, Navarro C, Gaxiola M, Selman M, Randall TD. Inducible bronchus-associated lymphoid tissue (iBALT) in patients with pulmonary complications of rheumatoid arthritis. J Clin Invest 2006;116:3183–94.

12. Demoruelle MK, Weisman MH, Simonian PL, Lynch DA, Sachs PB, Pedraza IF, et al. Airways abnormalities and rheumatoid arthritis–related autoantibodies in subjects without arthritis: early injury or initiating site of autoimmunity? Arthritis Rheum 2012;64:1756–61.

13. Fischer A, Solomon JJ, du Bois RM, Deane KD, Olson AL, Fernandez-Perez ER, et al. Lung disease with anti-CCP antibodies but not rheumatoid arthritis or connective tissue disease. Respir Med 2012;106:1040–7.

14. Demoruelle MK, Harrall KK, Ho L, Purmalek MM, Seto NL, Rothfuss HM, et al. Anti–citrullinated protein antibodies are associated with neutrophil extracellular traps in the sputum in relatives of rheumatoid arthritis patients. Arthritis Rheumatol 2017;69:1165–75.

15. Willis VC, Demoruelle MK, Derber LA, Chartier-Logan CJ, Parish MC, Pedraza IF, et al. Sputum autoantibodies in patients with established rheumatoid arthritis and subjects at risk of future clinically apparent disease. Arthritis Rheum 2013;65:2545–54.

16. Makrygiannakis D, af Klint E, Lundberg IE, Löfberg R, Ulfgren AK, Klareskog L, et al. Citrullination is an inflammation-dependent process. Ann Rheum Dis 2006;65:1219–22.

17. Lugli EB, Correia RE, Fischer R, Lundberg K, Bracke KR, Montgomery AB, et al. Expression of citrulline and homocitrulline residues in the lungs of non-smokers and smokers: implications for autoimmunity in rheumatoid arthritis. Arthritis Res Ther 2015;17:9.

18. Wood AM, de Pablo P, Buckley CD, Ahmad A, Stockley RA. Smoke exposure as a determinant of autoantibody titre in α_1-antitrypsin deficiency and COPD. Eur Respir J 2011;37:32–8.

19. Ford JA, Liu X, Chu SH, Lu B, Cho MH, Silverman EK, et al. Asthma, chronic obstructive pulmonary disease, and subsequent risk for incident rheumatoid arthritis among women: a prospective cohort study. Arthritis Rheumatol 2020;72:704–13.

20. Castaldi PJ, Benet M, Petersen H, Rafaels N, Finigan J, Paoletti M, et al. Do COPD subtypes really exist? COPD heterogeneity and clustering in 10 independent cohorts. Thorax 2017;72:998–1006.

21. Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, Li X, et al. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. Am J Respir Crit Care Med 2010;181:315–23.

EDITORIAL

Advancing Rheumatoid Arthritis Synovial Biopsy Analysis: One B Cell at a Time

Dana E. Orange[1] (ID) and Laura T. Donlin[2] (ID)

While biologic therapies have dramatically improved the management of rheumatoid arthritis (RA), not all patients respond well to treatment, and the field continues to await individualized medicine approaches that promise to determine effective medications based on a patient's distinct disease features. The cellular and molecular characterization of an individual's synovium holds promise to understand pathology, predict outcomes, and tailor treatment decisions, akin to the histologic and genetic assessment of tumor tissue that guides cancer care. Recently, several large consortia (1–5) have set out to examine whether histologic and functional genomic assays of synovial tissue samples can provide important insights into RA disease pathology and predict clinical outcomes. These large-scale efforts have also taken on the task of harmonizing methods for assessment of synovial tissues to facilitate standardized clinical use.

B cells are of particular interest in efforts to connect synovial tissue features with clinical measures. The notion that B cells impact RA pathogenesis is supported by the presence of B cells in RA synovium, the fact that autoantibodies are one of the few objective molecular markers of RA disease, and the fact that B cell–depleting therapies have been shown to be an effective treatment. However, there is considerable patient-to-patient variability in the extent of synovial B cell infiltration, and while some groups have found B cell infiltrate scores useful for predicting RA treatment response, others have not found associations with clinical assessments of RA disease activity. One possible source of this discrepancy is the lack of standardized B cell scoring systems for synovial tissue.

In this issue of *Arthritis & Rheumatology*, Rivellese et al (6) present steps to validate a semiquantitative CD20 immunohistochemistry–based B cell score. They applied their approach to synovial biopsy specimens from 2 sizable cohorts: the Pathobiology of Early Arthritis Cohort (PEAC), which consisted of 165 consecutive patients with untreated RA with a disease duration of <1 year, and a subset of patients with an inadequate response to tumor necrosis factor inhibitors (TNFi) from the Response–Resistance to Rituximab versus Tocilizumab in RA (R4RA) trial (the TNFi-IR cohort), which consisted of 164 patients who had a mean disease duration of 12.8 years. All patient synovial biopsy specimens were assessed by immunohistochemical staining for CD20 (B cell marker), CD68 (macrophage/fibroblast marker), and CD138 (plasma cell marker), while a substantial number were also analyzed by RNA sequencing. From the transcriptome analysis, the authors generated B cell modules using subsets of 91 and 127 patients from the PEAC and TNFi-IR cohorts, respectively, using FANTOM data, an online resource of complementary DNA data established in part by the RIKEN institute in Japan. Their analysis found that the pathologist-interpreted semiquantitative CD20 immunohistochemistry–based B cell scores correlated with both a digital image analysis of CD20 immunohistochemistry and B cell gene expression module scores. With validation of the semiquantitative CD20 immunohistochemistry–based B cell scores using these 2 orthogonal approaches, they applied the B cell scores to the larger cohort of 329 patients and tested for clinical associations.

There were 2 major observations from this analysis. First, the investigators found that B cell–rich synovitis was more common in the TNFi-IR cohort than in the early RA cohort. The TNFi-IR cohort was also enriched for anti–citrullinated protein antibody (ACPA)–positive patients, raising the possibility that increased B cell–rich synovitis could be attributed to increased B cell–rich synovitis in seropositive patients. It is also possible that B cell–rich synovitis is more prevalent in patients with a longer disease duration. Future

studies evaluating the frequencies of B cell–rich synovitis across a spectrum of disease durations and treatment exposures are needed to determine whether time or treatments impact synovial B cell accumulation.

A second notable finding related to B cell scores associating with numerous disease features, such as the Krenn synovitis score (a global histology severity score for RA that includes assessments of synovial lining hyperplasia, infiltrates, and stromal cell activation), the Disease Activity Score in 28 joints (DAS28), swollen joint count, erythrocyte sedimentation rate, C-reactive protein (CRP) level, and ACPA and rheumatoid factor positivity, in the PEAC but not in the TNFi-IR cohort (where B cell scores associated with only Krenn score and CRP level). The finding that B cell scores associated with disease activity scores in the PEAC confirms the findings of previous studies demonstrating that B cell–rich lymphoid aggregates are associated with disease activity and predict treatment response to disease-modifying antirheumatic drugs as well as radiographic progression (3).

Interestingly, the same B cell scores did not associate with measures of disease activity such as the DAS28 in the TNFi-IR group. One possible explanation for this discordance is that the TNFi-IR cohort had more aggressive disease overall, thereby reducing the chances of detecting variance in disease activity. This discrepancy between histology scores for synovitis and the DAS28 is consistent with the results of studies from our group that failed to find associations of inflammatory infiltrates with the DAS28 in a cohort of patients with established disease undergoing arthroplasty (7). We noted that some patients with disease in remission or with low disease activity according to the DAS28 harbored subclinical synovitis (7,8), suggesting that current clinical measures of annotating disease activity, such as the DAS28, are not sufficiently sensitive for detecting synovitis in all patients. Indeed, multiple studies have noted ongoing radiographic synovitis as well as progression of erosions in patients with sustained low disease activity or remission (9–12). Further, over time, patients with long-standing disease may accrue various levels and stages of synovial inflammation and damage in different joints. This increased joint-to-joint variability may exacerbate sampling error in patients with established disease.

The findings of Rivellese et al (6), along with those of other recent studies, are advancing our understanding of the transcriptional and cellular characteristics of the synovium in RA. Incorporation of synovial assessments into clinical management is the next step in empowering clinicians to apply advances in molecular immunology to better tailor treatment decisions. Identification of subsets of RA that will specifically respond to therapeutics will enable targeting of their respective underlying disease mechanisms. Together, this work and complementary work by other groups will set the stage for "precision therapy for RA" and thereby advance the care of our patients.

AUTHOR CONTRIBUTIONS

Drs. Orange and Donlin drafted the article, revised it critically for important intellectual content, and approved the final version to be published.

REFERENCES

1. Astorri E, Nerviani A, Bombardieri M, Pitzalis C. Towards a stratified targeted approach with biologic treatments in rheumatoid arthritis: role of synovial pathobiology. Curr Pharm Des 2015;21: 2216–24.

2. De Groof A, Ducreux J, Humby F, Nzeusseu Toukap A, Badot V, Pitzalis C, et al. Higher expression of TNFα-induced genes in the synovium of patients with early rheumatoid arthritis correlates with disease activity, and predicts absence of response to first line therapy. Arthritis Res Ther 2016;18:19.

3. Humby F, Lewis M, Ramamoorthi N, Hackney JA, Barnes MR, Bombardieri M, et al. Synovial cellular and molecular signatures stratify clinical response to csDMARD therapy and predict radiographic progression in early rheumatoid arthritis patients. Ann Rheum Dis 2019;78:761–72.

4. Najm A, Le Goff B, Orr C, Thurlings R, Cañete JD, Humby F, et al. Standardisation of synovial biopsy analyses in rheumatic diseases: a consensus of the EULAR Synovitis and OMERACT Synovial Tissue Biopsy Groups. Arthritis Res Ther 2018;20:265.

5. Zhang F, Wei K, Slowikowski K, Fonseka CY, Rao DA, Kelly S, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat Immunol 2019;20:928–42.

6. Rivellese F, Humby F, Bugatti S, Fossati-Jimack L, Rizvi H, Lucchesi D, et al, and the PEAC-R4RA Investigators. B cell synovitis and clinical phenotypes in rheumatoid arthritis: relationship to disease stages and drug exposure. Arthritis Rheumatol doi: http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract. E-pub ahead of print.

7. Orange DE, Agius P, DiCarlo EF, Robine N, Geiger H, Szymonifka J, et al. Identification of three rheumatoid arthritis disease subtypes by machine learning integration of synovial histologic features and RNA sequencing data. Arthritis Rheumatol 2018;70:690–701.

8. Orange DE, Agius P, DiCarlo EF, Mirza SZ, Pannellini T, Szymonifka J, et al. Histologic and transcriptional evidence of subclinical synovial inflammation in patients with rheumatoid arthritis in clinical remission. Arthritis Rheumatol 2019;71:1034–41.

9. Molenaar ET, Voskuyl AE, Dinant HJ, Bezemer PD, Boers M, Dijkmans BA, et al. Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission. Arthritis Rheum 2004;50:36–42.

10. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E, et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug–induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum 2006;54:3761–73.

11. Ranganath VK, Motamedi K, Haavardsholm EA, Maranian P, Elashoff D, McQueen F, et al. Comprehensive appraisal of magnetic resonance imaging findings in sustained rheumatoid arthritis remission: a substudy. Arthritis Care Res (Hoboken) 2015;67:929–39.

12. Kohler MJ, editor. Musculoskeletal ultrasound in rheumatology review. Cham: Springer International Publishing Switzerland; 2016.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Asthma, Chronic Obstructive Pulmonary Disease, and Subsequent Risk for Incident Rheumatoid Arthritis Among Women: A Prospective Cohort Study

Julia A. Ford,[1] (ID) Xinyi Liu,[2] Su H. Chu,[1] (ID) Bing Lu,[1] Michael H. Cho,[1] Edwin K. Silverman,[1] Karen H. Costenbader,[1] (ID) Carlos A. Camargo Jr.,[3] and Jeffrey A. Sparks[1] (ID)

Objective. Inflamed airways are hypothesized to contribute to rheumatoid arthritis (RA) pathogenesis due to RA-related autoantibody production, and smoking is the strongest environmental RA risk factor. However, the role of chronic airway diseases in RA development is unclear. We undertook this study to investigate whether asthma and chronic obstructive pulmonary disease (COPD) were each associated with RA.

Methods. We performed a prospective cohort study of 205,153 women in the Nurses' Health Study (NHS, 1988–2014) and NHSII (1991–2015). Exposures were self-reported physician-diagnosed asthma or COPD confirmed by validated supplemental questionnaires. The primary outcome was incident RA confirmed by medical record review by 2 rheumatologists. Covariates (including smoking pack-years/status) were assessed via biennial questionnaires. Multivariable hazard ratios (HRs) and 95% confidence intervals (CIs) for RA were estimated using Cox regression.

Results. We identified 15,148 women with confirmed asthma, 3,573 women with confirmed COPD, and 1,060 incident RA cases during 4,384,471 person-years (median 24.0 years/participant) of follow-up in the NHS and NHSII. Asthma was associated with increased RA risk (HR 1.53 [95% CI 1.24–1.88]) compared to no asthma/COPD after adjustment for covariates, including smoking pack-years/status. Asthma remained associated with increased RA risk when analyzing only never-smokers (HR 1.53 [95% CI 1.14–2.05]). COPD was also associated with increased RA risk (HR 1.89 [95% CI 1.31–2.75]). The association of COPD with RA was most pronounced in the subgroup of ever-smokers age >55 years (HR 2.20 [95% CI 1.38–3.51]).

Conclusion. Asthma and COPD were each associated with increased risk of incident RA, independent of smoking status/intensity and other potential confounders. These results provide support for the hypothesis that chronic airway inflammation may be crucial in RA pathogenesis.

INTRODUCTION

Patients with rheumatoid arthritis (RA) have increased respiratory morbidity and mortality (1,2). Pulmonary inflammation has been implicated in RA pathogenesis (3–6). Whether diseases related to chronic airway inflammation increase risk of developing RA, however, is unclear.

Asthma is a common disease characterized by chronic airway inflammation (7). Prior studies investigating an association between asthma and RA risk (8–16) were limited by small sample

size (8,9), lack of adjustment for smoking (an established RA risk factor) (8,12,13,15), and inability to measure RA phenotypes characterized by autoantibodies (9,11,13,15). Chronic obstructive pulmonary disease (COPD) is characterized by chronic inflammation and narrowing of airways, and smoking is a proven major risk factor (17). While RA has been shown to increase risk of subsequent COPD (18–21), to our knowledge no prior prospective cohort studies have examined COPD as a risk factor of incident RA.

We investigated the associations between asthma, COPD, and incident RA using 2 large prospective cohorts, the Nurses' Health Study (NHS) and NHSII. We hypothesized that asthma and COPD would each increase the risk of incident RA, independent of smoking.

SUBJECTS AND METHODS

Study population and design. We performed a prospective cohort study by pooling data from 2 prospective cohorts of female registered nurses in the US: the NHS, which began in 1976 and enrolled 121,700 nurses between the ages of 30 and 55 years, and the NHSII, which began in 1989 and enrolled 116,429 nurses between the ages of 25 and 42 years. Participants completed baseline and biennial questionnaires regarding lifestyle, health behaviors, medications, and diseases. Both cohorts had a follow-up response rate of >90%, and only 5% of person-time was lost to follow-up (22).

For this analysis, participants who had self-reported RA or other connective tissue disease (CTD) at study baseline, had missing data related to smoking pack-years at baseline, or did not return any follow-up questionnaire after study baseline were excluded. For the asthma analysis, participants with self-reported COPD at baseline were also excluded. For the COPD analysis, participants ≤35 years old with self-reported COPD were excluded, as has been done in previous studies (23), since COPD is rarely diagnosed prior to 35 years of age (24). Flow diagrams of the analyzed study populations for both the asthma and COPD analyses are presented in Supplementary Figures 1 and 2 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41194/abstract). All participants provided informed consent, and the study protocol was approved by the Institutional Review Boards of the Brigham and Women's Hospital and Harvard T. H. Chan School of Public Health.

Exposure variables (asthma and COPD). *Asthma.* Beginning with the 1988 (NHS) and 1991 (NHSII) questionnaires, participants were asked to report physician-diagnosed asthma. Positive responders were sent a previously validated supplemental respiratory questionnaire with detailed questions regarding asthma symptoms, medications, and diagnostic testing (25). The supplemental respiratory questionnaire categorized reported asthma according to the following case definitions: case definition 1 ("possible" asthma) was considered met if the

participant reiterated a physician diagnosis of asthma and reported taking an asthma medication since diagnosis; case definition 2 ("probable" asthma) was met if the participant fulfilled the criteria for case definition 1 and reported use of a long-term preventive asthma medication in the past year; and case definition 3 ("definite" asthma) was met if all preceding criteria were met and participant-reported physician diagnosis of asthma was made within 1 month of symptom onset. In 1999, Camargo and colleagues validated case definition 2 in a random sample of 100 women with high accuracy compared to the gold standard of presence of asthma by medical record review from a physician (25). We considered asthma as described in case definition 2 or higher (probable or definite) as confirmed asthma in our analyses. Participants who self-reported asthma but either failed to return the supplemental respiratory questionnaire or were disconfirmed per the respiratory questionnaire (did not meet the criteria for case definition 2 or higher) were censored at the time of initial self-report. Asthma status was time-updated during study follow-up.

COPD. Participants self-reported physician diagnosis of emphysema or chronic bronchitis biennially beginning in 1988 (NHS) and 1999 (NHSII), which was confirmed with a validated supplemental respiratory questionnaire (26). The supplemental respiratory questionnaire classified participants as having possible COPD if they answered affirmatively to having physician-diagnosed chronic bronchitis or emphysema or COPD; as having probable COPD if they met the criteria for a possible case, and the participant reported having a diagnostic test such as pulmonary function testing, chest radiograph, or chest computed tomography scan at the time that the criteria for a possible case were met; or as having definite COPD if the criteria for a possible case were met and the participant reported having pulmonary function testing within the past year demonstrating forced expiratory volume in 1 second (FEV$_1$) of <80% predicted or FEV$_1$/forced vital capacity of <0.7. Barr et al validated these definitions in a cohort of 422 women, reporting a positive predictive value of 88% for probable COPD against the gold standard of medical record review by a physician (26). We considered a participant who self-reported COPD to have confirmed COPD if the criteria for a probable or definite case were met. Participants who self-reported having COPD but did not return the respiratory questionnaire or were disconfirmed per the respiratory questionnaire (did not meet criteria for a probable or definite case) were censored at time of report. Participants who self-reported having asthma (not validated by the questionnaire) prior to receiving a validated COPD diagnosis were included as exposed individuals in the COPD analysis. COPD status was time-updated during study follow-up.

Nonexposed group (no asthma or COPD). For each analysis, subjects contributed person-time to the nonexposed group until they self-reported having asthma or COPD. If asthma/COPD was confirmed by validated supplemental questionnaires, person-time was contributed to that exposed group thereafter. If participants reported having asthma/COPD but did not return

the supplemental questionnaire or the diagnosis was not validated by the supplemental questionnaire, they were censored and no longer contributed person-time to that analysis. Therefore, at the beginning of each cycle considered in all analyses, the nonexposed group had never reported having asthma or COPD.

Outcome (incident RA). Participants who self-reported a new diagnosis of RA were mailed the CTD Screening Questionnaire (CSQ) (27). Medical records of participants with positive CSQ results were obtained and reviewed independently by 2 rheumatologists to identify RA cases meeting the 1987 American College of Rheumatology (ACR) criteria (28) or 2010 ACR/European League Against Rheumatism RA criteria (29). The date of RA diagnosis, and results of laboratory testing for rheumatoid factor (RF) and anti–cyclic citrullinated peptide (anti-CCP) antibodies, were collected from medical records. An RA case was determined to be seropositive if RF or CCP levels were above the upper limit of normal on the laboratory assay documented.

Covariates. We selected covariates as potential confounders associated with asthma, COPD, and RA based on prior literature (30–46), and all covariates were time-updated. Sociodemographic covariates included age, race, geographic region, and household income (categorized by quartile of US Census tract–based median household income at ZIP code level). Potential reproductive confounders were parity/total breastfeeding duration, menopausal status, and postmenopausal hormone use. We used a combined parity/total breastfeeding duration variable categorized as nulliparous, parous/0–<1 month, parous/1–11 months, or parous/≥12 months, and a combined variable for menopausal status and postmenopausal hormone use (premenopausal, postmenopausal/never, or postmenopausal/ever). We categorized body mass index as <25.0, 25.0–<30.0, or ≥30.0 kg/m^2. Physical activity was categorized as <3 metabolic equivalent (MET) hours/week versus ≥3 MET hours/week (47). Dietary intake, including alcohol consumption, was assessed by a semiquantitative food frequency questionnaire, the Alternative Healthy Eating Index (48), and categorized in quartiles. Considering health care utilization as a potential confounder on each questionnaire, we assessed whether the participant had a physical examination in the past 2 years.

Given previous associations of active and passive smoking with risk of COPD (49–51), asthma (52,53), and RA (54–57), adjustment for those was an important aspect of our analysis. On the baseline questionnaire, participants reported smoking status (never/past/current) and age at which they started smoking. Current smokers reported the number of cigarettes typically smoked per day, and past smokers provided the age at which they stopped smoking and the number of cigarettes smoked per day before quitting. On subsequent questionnaires, women reported smoking status and intensity (1–14 cigarettes/day, 15–24 cigarettes/day, ≥25 cigarettes/day). Smoking pack-years

were determined by multiplying packs of cigarettes smoked per day (20 cigarettes/pack) with the number of years smoked. We used smoking pack-years and smoking pack-years squared as continuous variables in our model, to include both a linear and quadratic term to account for possible nonlinear impact of smoking intensity on RA risk. We also adjusted for smoking status (never/past/current). All smoking variables were time-updated. To address passive smoking, participants were asked whether parents smoked in the house during childhood (yes/no) and whether they lived with a smoker for >1 year (ever/never).

Statistical analysis. We performed separate analyses for the coprimary exposures of asthma and COPD, each compared to participants without reported asthma or COPD. We pooled individual-level data from the NHS and NHSII for statistical efficiency. We reported descriptive statistics for covariates at the baseline of this analysis (NHS 1988 and NHSII 1991) in 3 groups: asthma with no COPD, COPD (with or without asthma), and no asthma or COPD.

Person-years of follow-up for each participant accrued from the date the study baseline questionnaire was returned (NHS 1988 and NHSII 1991 for asthma; NHS 1988 and NHSII 1999 for COPD) to the date of censoring, whichever came first: RA diagnosis (outcome), reported other CTD not confirmed as RA, death, loss to follow-up (date of last questionnaire submitted), or end of follow-up for this analysis (June 1, 2014 for the NHS and June 1, 2015 for the NHSII). For the asthma analysis, we also censored at date of self-reported COPD diagnosis. For the COPD analysis, we included participants who self-reported having asthma prior to a COPD diagnosis that was not confirmed by supplemental questionnaire, with the rationale that self-reported asthma prior to confirmed COPD diagnosis likely represented COPD.

We used Cox proportional hazards models to test for the association between the exposure (asthma or COPD) and RA risk. Base models were adjusted for age, cohort, and questionnaire cycle (cohorts pooled by similar calendar times; e.g., the 1988 cycle in the NHS cohort was pooled with the 1989 cycle in the NHSII cohort). The multivariable model was additionally adjusted for the covariates noted above. Given the possibility of collinearity among certain covariates (such as smoking status and pack-years), we initially considered partial models that adjusted for smoking status and continuous smoking pack-years separately. The point estimates were unchanged in the partial models compared to the final model that considered them together, with little evidence for collinearity between the categorical smoking status variable and the continuous smoking pack-years variable. Therefore, we reported the final model that included both of these terms since tightly controlling for possible confounding by smoking, especially for the COPD analysis, was a priority of the study. Since smoking is known to be strongly related to COPD, we expected relatively few women with COPD to be nonsmokers. Therefore, we performed a subgroup analysis among ever-smokers, again

using partial models for smoking status (past/current) and continuous pack-years prior to including both in the final model; all produced nearly identical point estimates for RA risk without evidence for collinearity.

We further investigated the association between asthma and RA risk by additional analysis. First, we restricted the analysis to never-smokers. Second, we only considered the asthma exposure status as being prevalent, as of study baseline, as a proxy for the presence or absence of early-onset asthma since data on the precise age at onset prior to baseline were unavailable. Third, we only considered the asthma exposure as incident

after study baseline among women without prevalent asthma to investigate adult-onset asthma. For COPD, we performed an additional analysis restricted to participants who were ever-smokers and >55 years old, since the prevalence of COPD is highest in this demographic group. Finally, we analyzed COPD and RA risk among women with confirmed COPD who never self-reported having asthma.

We tested the proportional hazards assumption by including an interaction term between time after baseline and exposure status for the RA outcome, and in all analyses, it was verified that there were no statistically significant interactions. Two-sided

Table 1. Pooled baseline characteristics of the Nurses' Health Study (1988) sample and the Nurses' Health Study II (1991) sample (n = 205,153)*

	Asthma, no COPD (n = 6,250)	COPD (n = 1,004)	No asthma or COPD (n = 197,899)
Cohort, %			
Nurses' Health Study	37.9	73.9	46.3
Nurses' Health Study II	62.1	26.1	53.7
Age, years	42.5 ± 9.5	52.7 ± 10.2	44.4 ± 10.9
White race, %	94.2	95.1	93.0
Body mass index, kg/m^2	25.8 ± 5.8	26.4 ± 6.2	25.0 ± 5.1
US geographic region, %			
West	25.7	21.7	23.0
Midwest	26.5	22.3	27.0
Mid-Atlantic	33.3	37.9	35.0
New England	9.5	11.0	9.0
Southeast	5.0	7.2	6.0
Median household income, %			
Quartile 1 (lowest income)	22.3	30.1	26.7
Quartile 2	24.5	25.6	24.3
Quartile 3	25.7	24.7	24.4
Quartile 4 (highest income)	27.5	19.6	24.6
Smoking, pack-years	5.3 ± 10.1	24.3 ± 24.5	7.5 ± 13.6
Ever-smokers only	13.6 ± 12.2	34.0 ± 22.5	17.6 ± 16.0
Smoking status, %			
Never	61.5	28.5	57.1
Past	31.2	39.5	28.0
Current	7.3	32.0	14.5
Ever lived with smoker, %	49.4	70.8	44.0
Parents smoked in house during childhood, %	64.4	66.8	51.8
Cumulative average physical activity <3 MET hours/week, %	17.6	27.1	15.1
Cumulative average Alternate Healthy Eating Index, %			
Quartile 1 (lowest-quality diet)	23.0	26.2	19.7
Quartile 2	23.7	20.8	19.8
Quartile 3	22.9	21.9	19.9
Quartile 4 (highest-quality diet)	25.1	21.8	19.9
Parity/breastfeeding, %			
Nulliparous	19.9	11.8	15.1
Parous/<1 month	22.7	38.3	23.3
Parous/1–11 months	26.9	30.5	24.0
Parous/≥12 months	27.1	14.9	20.5
Menopausal status/postmenopausal hormone use, %			
Premenopausal	69.6	29.7	65.4
Postmenopausal/never	13.3	29.3	18.1
Postmenopausal/ever	17.1	41.0	16.5

* A total of 15,148 women with confirmed asthma and a total of 3,573 women with confirmed chronic obstructive pulmonary disease (COPD) were identified by the end of study follow-up for these analyses. Data were missing for some subjects/parameters. Values shown are based on the totals with data available. Except where indicated otherwise, values are the mean ± SD. MET = metabolic equivalent.

P values less than 0.05 were considered statistically significant. Analyses were performed using SAS version 9.4.

RESULTS

Sample size, asthma/COPD exposures, and RA outcomes. After baseline exclusions, there were a total of 196,409 participants included in the asthma analysis and 205,153 participants included in the COPD analysis. We identified 15,148 women with confirmed asthma, 3,573 women with confirmed COPD, and 1,060 incident RA cases (63% seropositive) during a total of 4,384,471 person-years of follow-up (for asthma analysis, median 23.9 years/participant [interquartile range 18.3–24.5]; for COPD analysis, median 24.0 years/participant [interquartile range 20.0–24.5]).

Characteristics of participants. Pooled baseline characteristics of the NHS and NHSII study participants categorized by exposure (asthma without COPD, COPD, and no asthma or COPD) are shown in Table 1. Women in the COPD group were older (mean 52.7 years, compared to 42.5 years in the asthma group and 44.4 years in the no asthma/COPD group). Those in the COPD group were also more likely to be postmenopausal

Table 2. Hazard ratios for incident RA (overall and by serologic phenotype) by time-updated asthma compared to women without asthma or COPD in the Nurses' Health Studies (n = 196,409)*

Asthma (versus reference of no asthma or COPD)	Outcome measure
All RA	
Cases/person-years†	100/265,359
HR (95% CI)	
Age-adjusted model	1.67 (1.35–2.05)
Multivariable model	1.53 (1.24–1.88)
Seropositive RA	
Cases/person-years†	60/264,341
HR (95% CI)	
Age-adjusted model	1.51 (1.15–1.97)
Multivariable model	1.42 (1.08–1.86)
Seronegative RA	
Cases/person-years†	40/264,329
HR (95% CI)	
Age-adjusted model	1.98 (1.42–2.76)
Multivariable model	1.75 (1.25–2.45)

* Multivariable models were adjusted for age, questionnaire period, cohort, US geographic region (West, Midwest, Mid-Atlantic, New England, Southeast), median household income (quartile), smoking pack-years (continuous), smoking status (never/past/current), cumulative average physical activity (<3, ≥3 metabolic equivalent hours/week), parity/months of breastfeeding (nulliparous, parous/<1 month, parous/1–11 months, parous/≥12 months), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/never, postmenopausal/ever), cumulative average Alternate Healthy Eating Index (quartile), body mass index category (<25.0, 25.0–<30.0, ≥30.0 kg/m^2), parents smoked in house during childhood (yes/no), lived with smoker (ever/never), and physical examination in last 2 years (yes/no). HR = hazard ratio; 95% CI = 95% confidence interval.
† Among subjects with no asthma or chronic obstructive pulmonary disease (COPD) (referent), cases/person-years for all rheumatoid arthritis (all RA), seropositive RA, and seronegative RA were 874/3,841,747, 562/3,834,291, and 312/3,833,574, respectively.

Table 3. Hazard ratios for incident RA (overall and by serologic phenotype) by time-updated asthma compared to women without asthma or COPD in the Nurses' Health Studies, stratified by never smoking (n = 110,872) and ever smoking (n = 85,537)*

Asthma (versus reference of no asthma or COPD)	Outcome measure
Never-smokers only	
All RA	
Cases/person-years†	52/159,472
HR (95% CI)	
Age-adjusted model	1.69 (1.27–2.27)
Multivariable model	1.53 (1.14–2.05)
Seropositive RA	
Cases/person-years†	28/158,816
HR (95% CI)	
Age-adjusted model	1.40 (0.95–2.08)
Multivariable model	1.32 (0.88–1.96)
Seronegative RA	
Cases/person-years†	24/158,973
HR (95% CI)	
Age-adjusted model	2.23 (1.44–3.46)
Multivariable model	1.90 (1.22–2.96)
Ever-smokers only‡	
All RA	
Cases/person-years§	48/105,887
HR (95% CI)	
Age-adjusted model	1.59 (1.18–2.15)
Multivariable model	1.49 (1.10–2.02)
Seropositive RA	
Cases/person-years§	32/105,525
HR (95% CI)	
Age-adjusted model	1.57 (1.08–2.27)
Multivariable model	1.50 (1.04–2.18)
Seronegative RA	
Cases/person-years§	16/105,356
HR (95% CI)	
Age-adjusted model	1.64 (0.97–2.75)
Multivariable model	1.48 (0.87–2.50)

* Multivariable models were adjusted for age, questionnaire period, cohort, US geographic region (West, Midwest, Mid-Atlantic, New England, Southeast), median household income (quartile), cumulative average physical activity (<3, ≥3 metabolic equivalent hours/week), parity/months of breastfeeding (nulliparous, parous/<1 month, parous/1–11 months, parous/≥12 months), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/never, postmenopausal/ever), cumulative average Alternate Healthy Eating Index (quartile), body mass index category (<25.0, 25.0–<30.0, ≥30.0 kg/m^2), parents smoked in house during childhood (yes/no), lived with smoker (ever/never), and physical examination in last 2 years (yes/no). See Table 2 for definitions.
† Among subjects with no asthma or COPD (referent), cases/person-years for all RA, seropositive RA, and seronegative RA were 421/2,218,758, 268/2,214,956, and 153/2,214,780, respectively.
‡ Multivariable model for the ever-smoker analysis was additionally adjusted for smoking pack-years (continuous) and smoking status (past/current).
§ Among subjects with no asthma or COPD (referent), cases/person-years for all RA, seropositive RA, and seronegative RA were 453/1,622,988, 294/1,619,336, and 159/1,618,793, respectively.

(70.3%, compared to 30.4% in the asthma group and 34.6% in the no asthma/COPD group). Pooled baseline characteristics of the NHS and NHSII study participants included in the asthma analysis are presented in Supplementary Table 1 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41194/abstract).

Asthma and RA risk. Compared to women without asthma or COPD, the multivariable-adjusted hazard ratio (HR) for developing RA was 1.53 (95% confidence interval [95% CI] 1.24–1.88) among women with asthma (Table 2). Asthma was associated with both seropositive RA and seronegative RA, and HRs for risk of seropositive RA versus seronegative RA were not significantly different (*P* for heterogeneity = 0.45).

Table 4. Hazard ratios for incident RA (overall and by serologic phenotype) according to prevalent asthma at study baseline or incident asthma during follow-up, each compared to women without asthma or COPD in the Nurses' Health Studies (n = 196,409)*

Asthma (versus reference of no asthma or COPD)	Outcome measure
Prevalent asthma, at baseline	
All RA	
Cases/person-years†	40/115,491
HR (95% CI)	
Age-adjusted model	1.60 (1.16–2.20)
Multivariable model	1.46 (1.06–2.01)
Seropositive RA	
Cases/person-years†	23/115,007
HR (95% CI)	
Age-adjusted model	1.37 (0.90–2.09)
Multivariable model	1.29 (0.85–1.97)
Seronegative RA	
Cases/person-years†	17/115,063
HR (95% CI)	
Age-adjusted model	2.05 (1.25–3.36)
Multivariable model	1.79 (1.09–2.94)
Incident asthma during follow-up	
All RA	
Cases/person-years‡	60/148,348
HR (95% CI)	
Age-adjusted model	1.76 (1.35–2.29)
Multivariable model	1.61 (1.23–2.09)
Seropositive RA	
Cases/person-years‡	37/147,814
HR (95% CI)	
Age-adjusted model	1.65 (1.18–2.31)
Multivariable model	1.56 (1.11–2.18)
Seronegative RA	
Cases/person-years‡	23/147,767
HR (95% CI)	
Age-adjusted model	1.96 (1.28–3.01)
Multivariable model	1.73 (1.13–2.66)

* Multivariable models were adjusted for age, questionnaire period, cohort, US geographic region (West, Midwest, Mid-Atlantic, New England, Southeast), median household income (quartile), smoking pack-years (continuous), smoking status (never/past/current), cumulative average physical activity (<3, ≥3 metabolic equivalent hours/week), parity/months of breastfeeding (nulliparous, parous/<1 month, parous/1–11 months, parous/≥12 months), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/never, postmenopausal/ever), cumulative average Alternate Healthy Eating Index (quartile), body mass index category (<25.0, 25.0–<30.0, ≥30.0 kg/m²), parents smoked in house when growing up (yes/no), lived with smoker (ever/never), and physical examination in last 2 years (yes/no). See Table 2 for definitions.
† Among subjects with no asthma or COPD (referent), cases/person-years for all RA, seropositive RA, and seronegative RA were 855/3,792,384, 548/3,785,152, and 307/3,784,460, respectively.
‡ Among subjects with no asthma or COPD at baseline (referent), cases/person-years for all RA, seropositive RA, and seronegative RA were 874/3,841,695, 562/3,834,240, and 312/3,833,522, respectively.

We examined the relationship between asthma and RA risk stratified by never-smoking and ever-smoking (Table 3). Among never-smokers only, asthma was associated with all RA (HR 1.53 [95% CI 1.14–2.05]) and seronegative RA (HR 1.90 [95% CI 1.22–2.96]) but not with seropositive RA (HR 1.32 [95% CI 0.88–1.96]), compared to women without asthma or COPD. Among ever-smokers only, subjects with asthma had HRs for all RA of 1.49 (95% CI 1.10–2.02), for seropositive RA of 1.50 (95% CI 1.04–2.18), and for seronegative RA of 1.48 (95% CI 0.87–2.50).

The association of prevalent asthma at study baseline (proxy for early-onset) and incident asthma during follow-up (proxy for adult-onset) with RA, compared to the reference group of women without asthma or COPD, is shown in Table 4. Overall RA risk was significantly increased in both prevalent asthma (HR 1.46 [95% CI 1.06–2.01]) and incident asthma (HR 1.61 [95% CI 1.23–2.09]).

COPD and RA risk. Compared to women without asthma or COPD, the multivariable-adjusted HR for developing RA was 1.89 (95% CI 1.31–2.75) among women with COPD (Table 5). COPD significantly increased the risk for seropositive RA (HR 2.07 [95% CI 1.31–3.25]) but not seronegative RA (HR 1.59 [95% CI 0.83–3.05]). Among ever-smokers age >55 years, there

Table 5. Hazard ratios for incident RA (overall and by serologic phenotype) by time-updated COPD compared to women without asthma or COPD in the Nurses' Health Studies (n = 205,153)*

COPD (versus reference of no asthma or COPD)	Outcome measure
All RA	
Cases/person-years†	31/47,285
HR (95% CI)	
Age-adjusted model	2.39 (1.66–3.43)
Multivariable model	1.89 (1.31–2.75)
Seropositive RA	
Cases/person-years†	21/47,134
HR (95% CI)	
Age-adjusted model	2.69 (1.73–4.18)
Multivariable model	2.07 (1.31–3.25)
Seronegative RA	
Cases/person-years†	10/47,121
HR (95% CI)	
Age-adjusted model	1.93 (1.02–3.64)
Multivariable model	1.59 (0.83–3.05)

* Multivariable models were adjusted for age, questionnaire period, cohort, US geographic region (West, Midwest, Mid-Atlantic, New England, Southeast), median household income (quartile), smoking pack-years (continuous), smoking pack-years-squared (continuous), smoking status (never/past/current), cumulative average physical activity (<3, ≥3 metabolic equivalent hours/week), parity/months of breastfeeding (nulliparous, parous/<1 month, parous/1–11 months, parous/≥12 months), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/never, postmenopausal/ever), cumulative average Alternate Healthy Eating Index (quartile), body mass index category (<25.0, 25.0–<30.0, ≥30.0 kg/m²), parents smoked in house during childhood (yes/no), lived with smoker (ever/never), and physical examination in last 2 years (yes/no). See Table 2 for definitions.
† Among subjects with no asthma or COPD (referent), cases/person-years for all RA, seropositive RA, and seronegative RA were 1,029/4,337,186, 642/4,328,257, and 387/4,327,740, respectively.

Table 6. Hazard ratios for incident RA (overall and by serologic phenotype) by time-updated COPD compared to women without asthma or COPD in the Nurses' Health Studies among ever-smokers age >55 years (n = 21,525)*

COPD (versus reference of no asthma or COPD)	Outcome measure
All RA	
Cases/person-years†	21/29,365
HR (95% CI)	
Age-adjusted model	2.26 (1.44–3.55)
Multivariable model	2.20 (1.38–3.51)
Seropositive RA	
Cases/person-years†	15/29,271
HR (95% CI)	
Age-adjusted model	2.80 (1.64–4.80)
Multivariable model	2.85 (1.63–4.99)
Seronegative RA	
Cases/person-years†	6/29,279
HR (95% CI)	
Age-adjusted model	1.52 (0.66–3.50)
Multivariable model	1.40 (0.59–3.29)

* Multivariable models were adjusted for age, questionnaire period, cohort, US geographic region (West, Midwest, Mid-Atlantic, New England, Southeast), median household income (quartile), smoking pack-years (continuous), smoking pack-years-squared (continuous), smoking status (past/current), cumulative average physical activity (<3, ≥3 metabolic equivalent hours/week), parity/months of breastfeeding (nulliparous, parous/<1 month, parous/1–11 months, parous/≥12 months), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/never, postmenopausal/ever), cumulative average Alternate Healthy Eating Index (quartile), body mass index category (<25.0, 25.0–<30.0, ≥30.0 kg/m^2), parents smoked in house during childhood (yes/no), lived with smoker (ever/never), and physical examination in last 2 years (yes/no). See Table 2 for definitions.
† Among subjects with no asthma or COPD (referent), cases/person-years for all RA, seropositive RA, and seronegative RA were 295/928,014, 176/926,338, and 119/926,271, respectively.

Correction added after online publication 10 April 2020: In Table heading "Asthma" has been changed to COPD for Table 5 and Table 6.

was a stronger association between COPD and seropositive RA (HR 2.85 [95% CI 1.63–4.99]) (Table 6). Among women with confirmed COPD who never self-reported asthma, COPD was associated with RA (HR 2.57 [95% CI 1.51–4.39]) (see Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41194/abstract).

DISCUSSION

In this large prospective cohort study with long follow-up, asthma was associated with a >50% increase in the risk of subsequent RA compared to no asthma/COPD, independent of potential confounders, most notably smoking status and duration/intensity. COPD conferred a nearly 90% increased risk of developing RA compared to no asthma/COPD in this cohort after multivariable adjustment, including adjustment for smoking, with a >2-fold increased risk of RA among older smokers. These results demonstrate that asthma and COPD are risk factors for the development of RA, and to our knowledge, this is the first prospective study to examine asthma or COPD as RA risk factors.

These findings suggest that asthma and COPD patients may be an at-risk population for purposes of research into RA development and provide helpful information for clinicians caring for asthma/COPD patients regarding elevated RA risk in these groups.

While one retrospective study of Israeli soldiers showed an inverse association between asthma and RA (12), the majority of the published literature suggests that asthma increases RA risk (8–11,13,15,16). In a population-based case–control study in which asthma was identified via medical record review and RA outcomes using billing codes, Sheen et al (9) found an odds ratio (OR) for RA of 1.73 (95% CI 1.03–2.92) among patients with asthma compared to matched controls. While this study adjusted for several factors, including smoking status, and had high diagnostic accuracy for asthma, there was no adjustment for smoking duration/intensity, and sample size limited the ability to examine RA serologic status. Kronzer and colleagues (10) identified RA cases in a biobank population using a rules-based algorithm, finding an increased OR for RA of 1.28 (95% CI 1.04–1.67) among subjects with self-reported asthma compared to matched controls. Similar to Sheen et al (9), the authors adjusted for smoking status but not duration/intensity, and the results from this clinically sampled population mostly with prevalent RA may lack generalizability. Our study similarly showed a positive association between asthma and RA risk, and had the further strengths of prospective, detailed data collection with time-updated adjustment for multiple covariates, including smoking status as well as smoking duration/intensity. Furthermore, we utilized time-updated exposure and outcome assessment using confirmed, not self-reported, diagnoses of both asthma and RA with serologic phenotype. Our large sample size enabled us to phenotype asthma exposures (according to never/ever smoking status and incident/prevalent disease) and RA outcomes (according to serologic status).

Prior studies have demonstrated a positive association between RA and subsequent risk of developing COPD (18–21). However, we identified only 2 studies examining COPD as an RA risk factor, neither of which were prospective. A Swedish nested case–control study (58) examined COPD (according to the Global Initiative for Obstructive Lung Disease stage results from spirometry performed for research purposes) as a risk factor for RA, and ORs were not significant, likely due to small number of COPD exposures. In the case–control study by Sheen and colleagues (9), COPD was not associated with RA as an unadjusted baseline variable. Our findings are the first to demonstrate COPD as a risk factor for seropositive RA after adjustment for smoking. While our primary analysis for COPD included some cases who previously self-reported asthma, the secondary analysis that did not include women who reported both asthma and COPD showed an even stronger association of COPD with increased RA risk. Therefore, it is unlikely that the COPD results were explained by inclusion of participants who may have also had preexisting or concurrent asthma.

There is increasing interest in the lung and airway mucosa as an initiating site for RA pathogenesis. RA-specific autoantibodies

are increased in the sputum of unaffected first-degree relatives of RA patients prior to detection of elevation in the serum (59). In newly diagnosed RA patients, lymphoid aggregates are present near airways and interstitium (5,6). These findings provided the biologic basis of our hypotheses that asthma and COPD would increase RA risk. We might have also expected stronger associations between asthma/COPD for seropositive RA compared to seronegative RA. While we did observe this in COPD, there was not a statistically significant difference in the HRs for seropositive RA and seronegative RA in our primary asthma analysis. Prevalent asthma at study baseline was significantly positively associated with overall and seronegative RA, but not seropositive RA. We were unable to investigate RF-positive RA and CCP-positive RA separately, as we relied on clinical testing and the CCP test was not available in the early years of follow-up. It is unclear whether these discrepancies in risk according to RA serologic status have biologic and clinical meaning, or whether the smaller numbers of exposed cases in certain analyses affected statistical power.

Our study has several key strengths. We used a validated method to identify women with asthma and COPD throughout follow-up based on self-report and then confirmed the diagnosis with a supplemental respiratory questionnaire. The reference group had never reported asthma or COPD on any main questionnaire. Since many women reported having asthma/COPD during the study, manual medical record review to confirm the presence or absence of these diseases would have been infeasible. Therefore, our methods allowed us to compare confirmed asthma/COPD cases to women who never reported having these diseases throughout the lengthy follow-up in these large prospective cohorts. We have a high level of confidence in RA outcomes and date of RA diagnosis, which were agreed upon by 2 rheumatologists according to classification criteria. The NHS and NHSII are large cohorts with detailed prospectively collected data, which permitted adjustment for multiple time-updated covariates, most importantly smoking (both smoking status and duration/intensity). Our findings were robust to sensitivity analyses regarding smoking, particularly the positive association between asthma and RA among never-smokers.

A limitation of this study was that the Nurses' Health Studies included only women, the majority of whom were white, limiting generalizability. While we used validated methods to identify confirmed cases of asthma/COPD, the supplemental questionnaire was sent only once for the purpose of establishing the presence of these diseases. The presence of these diseases has acceptable validity when compared to medical records, but it is still possible that there was some misclassification since both of the separate questionnaires were by self-report. We mitigated the possibility of misclassification by not analyzing ambiguous cases that were reported once but not again or were not deemed to have high likelihood of truly having asthma or COPD.

Furthermore, inclusion of women without confirmed asthma or COPD in these analyses would be expected to bias results to the null, and among ever-smokers age >55 years (a group at higher risk for COPD), we observed the strongest association with RA risk. Further, reports have quantified similar proportions of nonsmokers with COPD as were observed in our study, perhaps from air pollution, occupational exposures, or other inhalants (60). Since data on clinical characteristics of asthma/COPD were only collected once, we were unable to analyze how the disease course of asthma/COPD may impact RA risk. We also had limited ability to analyze asthma and COPD according to disease severity, imaging/pulmonary function test abnormalities, medication use, or (in the case of asthma) atopy. Future studies should investigate whether characteristics of asthma/COPD based on subtype, severity, medication use, and flares are associated with RA risk and RA-related autoantibody production.

As expected, women with confirmed COPD were older and thus were more likely to be postmenopausal. It is possible that they may have had other differences related to hormonal/reproductive factors that could also impact RA risk. Our multivariable analyses therefore included the following hormonal/reproductive factors: parity, breastfeeding duration, menopausal status, and postmenopausal hormone use. We further mitigated the possibility that COPD results may have been explained by hormonal differences by performing a secondary analysis for COPD among only older women. This showed that COPD remained strongly associated with seropositive RA, so it is unlikely that differences in menopause explained our results. Regarding the strong association between COPD and subsequent RA risk among ever-smokers age >55 years, it is possible that hormonal factors may particularly impact RA risk in this subgroup; we mitigated this possibility by adjusting for menopausal status/postmenopausal hormone use as well as parity/breastfeeding. We chose the cutoff point of age 55 years to classify women as younger or older in that particular analysis based on prior literature (61) as well as to reflect the menopausal transition and because COPD is known to be an older-onset disease. However, we acknowledge the possibility that the selection of a different age cutoff may have yielded different results. Finally, inherent to any observational study is the possibility of unmeasured confounding, particularly related to health care utilization.

In this large prospective cohort study of women, asthma and COPD were risk factors for subsequent RA after adjustment for multiple potential confounders, including smoking status and duration/intensity. These novel findings further implicate chronic airway inflammation in the pathogenesis of RA, and they identify populations at risk for RA for the purposes of research as well as informing clinical care. Providers caring for patients with asthma or COPD should be aware of increased RA risk in these populations and have a low threshold to evaluate for RA in patients with asthma or COPD with inflammatory joint symptoms.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

REFERENCES

1. Sparks JA, Chang SC, Liao KP, Lu B, Fine AR, Solomon DH, et al. Rheumatoid arthritis and mortality among women during 36 years of prospective follow-up: results from the Nurses' Health Study. Arthritis Care Res (Hoboken) 2016;68:753–62.

2. England BR, Sayles H, Michaud K, Caplan L, Davis LA, Cannon GW, et al. Cause-specific mortality in male US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:36–45.

3. Willis VC, Demoruelle MK, Derber LA, Chartier-Logan CJ, Parish MC, Pedraza IF, et al. Sputum autoantibodies in patients with established rheumatoid arthritis and subjects at risk of future clinically apparent disease. Arthritis Rheum 2013;65:2545–54.

4. Reynisdottir G, Karimi R, Joshua V, Olsen H, Hensvold AH, Harju A, et al. Structural changes and antibody enrichment in the lungs are early features of anti–citrullinated protein antibody–positive rheumatoid arthritis. Arthritis Rheumatol 2014;66:31–9.

5. Demoruelle MK, Weisman MH, Simonian PL, Lynch DA, Sachs PB, Pedraza IF, et al. Airways abnormalities and rheumatoid arthritis–related autoantibodies in subjects without arthritis: early injury or initiating site of autoimmunity? Arthritis Rheum 2012;64:1756–61.

6. Reynisdottir G, Olsen H, Joshua V, Engström M, Forsslund H, Karimi R, et al. Signs of immune activation and local inflammation are present in the bronchial tissue of patients with untreated early rheumatoid arthritis. Ann Rheum Dis 2016;75:1722–7.

7. Sears MR. Trends in the prevalence of asthma. Chest 2014;145: 219–25.

8. De Roos AJ, Cooper GS, Alavanja MC, Sandler DP. Personal and family medical history correlates of rheumatoid arthritis. Ann Epidemiol 2008;18:433–9.

9. Sheen YH, Rolfes MC, Wi CI, Crowson CS, Pendegraft RS, King KS, et al. Association of asthma with rheumatoid arthritis: a population-based case-control study. J Allergy Clin Immunol Pract 2018;6:219–26.

10. Kronzer VL, Crowson CS, Sparks JA, Vassallo R, Davis JM III. Investigating asthma, allergic disease, passive smoke exposure, and risk of rheumatoid arthritis. Arthritis Rheumatol 2019;71:1217–24.

11. Jeong HE, Jung SM, Cho SI. Association between rheumatoid arthritis and respiratory allergic diseases in Korean adults: a propensity score matched case-control study. Int J Rheumatol 2018;2018:3798124.

12. Tirosh A, Mandel D, Mimouni FB, Zimlichman E, Shochat T, Kochba I. Autoimmune diseases in asthma. Ann Intern Med 2006;144: 877–83.

13. Hemminki K, Li X, Sundquist J, Sundquist K. Subsequent autoimmune or related disease in asthma patients: clustering of diseases or medical care? Ann Epidemiol 2010;20:217–22.

14. Yun HD, Knoebel E, Fenta Y, Gabriel SE, Leibson CL, Loftus EV Jr, et al. Asthma and proinflammatory conditions: a population-based retrospective matched cohort study. Mayo Clin Proc 2012;87:953–60.

15. Lai NS, Tsai TY, Koo M, Lu MC. Association of rheumatoid arthritis with allergic diseases: a nationwide population-based cohort study. Allergy Asthma Proc 2015;36:99–103.

16. Hou YC, Hu HY, Liu IL, Chang YT, Wu CY. The risk of autoimmune connective tissue diseases in patients with atopy: a nationwide population-based cohort study. Allergy Asthma Proc 2017;38: 383–9.

17. Landis SH, Muellerova H, Mannino DM, Menezes AM, Han MK, van der Molen T, et al. Continuing to Confront COPD International Patient Survey: methods, COPD prevalence, and disease burden in 2012–2013. Int J Chron Obstruct Pulmon Dis 2014;9: 597–611.

18. Ursum J, Nielen MM, Twisk JW, Peters MJ, Schellevis FG, Nurmohamed MT, et al. Increased risk for chronic comorbid disorders in patients with inflammatory arthritis: a population based study. BMC Fam Pract 2013;14:199.

19. Nannini C, Medina-Velasquez YF, Achenbach SJ, Crowson CS, Ryu JH, Vassallo R, et al. Incidence and mortality of obstructive lung disease in rheumatoid arthritis: a population-based study. Arthritis Care Res (Hoboken) 2013;65:1243–50.

20. Ungprasert P, Srivali N, Cheungpasitporn W, Davis JM III. Risk of incident chronic obstructive pulmonary disease in patients with rheumatoid arthritis: a systematic review and meta-analysis. Joint Bone Spine 2016;83:290–4.

21. Sparks JA, Lin TC, Camargo CA Jr, Barbhaiya M, Tedeschi SK, Costenbader KH, et al. Rheumatoid arthritis and risk of chronic obstructive pulmonary disease or asthma among women: a marginal structural model analysis in the Nurses' Health Study. Semin Arthritis Rheum 2018;47:639–48.

22. Bao Y, Bertoia ML, Lenart EB, Stampfer MJ, Willett WC, Speizer FE, et al. Origin, methods, and evolution of the Three Nurses' health studies. Am J Public Health 2016;106:1573–81.

23. Rana JS, Mittleman MA, Sheikh J, Hu FB, Manson JE, Colditz GA, et al. Chronic obstructive pulmonary disease, asthma, and risk of type 2 diabetes in women. Diabetes Care 2004;27:2478–84.

24. Pauwels RA, Buist AS, Ma P, Jenkins CR, Hurd SS, GOLD Scientific Committee. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: National Heart, Lung, and Blood Institute and World Health Organization Global Initiative for Chronic Obstructive Lung Disease (GOLD)—executive summary. Respir Care 2001;46:798–825.

25. Camargo CA Jr, Weiss ST, Zhang S, Willett WC, Speizer FE. Prospective study of body mass index, weight change, and risk of adult-onset asthma in women. Arch Intern Med 1999;159:2582–8.

26. Barr RG, Herbstman J, Speizer FE, Camargo CA Jr. Validation of self-reported chronic obstructive pulmonary disease in a cohort study of nurses. Am J Epidemiol 2002;155:965–71.

27. Karlson EW, Sanchez-Guerrero J, Wright EA, Lew RA, Daltroy LH, Katz JN, et al. A connective tissue disease screening questionnaire for population studies. Ann Epidemiol 1995;5:297–302.

28. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

29. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.

30. Akinbami LJ, Moorman JE, Bailey C, Zahran HS, King M, Johnson CA, et al. Trends in asthma prevalence, health care use, and mortality in the United States, 2001–2010. NCHS Data Brief 2012;1–8.

31. Sahni S, Talwar A, Khanijo S, Talwar A. Socioeconomic status and its relationship to chronic respiratory disease. Adv Respir Med 2017;85:97–108.

32. Bergström U, Jacobsson LT, Nilsson JÅ, Berglund G, Turesson C. Pulmonary dysfunction, smoking, socioeconomic status and the risk of developing rheumatoid arthritis. Rheumatology (Oxford) 2011;50:2005–13.

33. Shah R, Newcomb DC. Sex bias in asthma prevalence and pathogenesis [review]. Front Immunol 2018;9:2997.

34. Romieu I, Fabre A, Fournier A, Kauffmann F, Varraso R, Mesrine S, et al. Postmenopausal hormone therapy and asthma onset in the E3N cohort. Thorax 2010;65:292–7.

35. Matulonga-Diakiese B, Courbon D, Fournier A, Sanchez M, Bédard A, Mesrine S, et al. Risk of asthma onset after natural and surgical menopause: results from the French E3N cohort. Maturitas 2018;118:44–50.

36. Kamil F, Pinzon I, Foreman MG. Sex and race factors in early-onset COPD. Curr Opin Pulm Med 2013;19:140–4.

37. Karlson EW, Mandl LA, Hankinson SE, Grodstein F. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses' Health Study. Arthritis Rheum 2004;50:3458–67.

38. Orellana C, Saevarsdottir S, Klareskog L, Karlson EW, Alfredsson L, Bengtsson C. Postmenopausal hormone therapy and the risk of rheumatoid arthritis: results from the Swedish EIRA population-based case-control study. Eur J Epidemiol 2015;30:449–57.

39. Beuther DA, Sutherland ER. Overweight, obesity, and incident asthma: a meta-analysis of prospective epidemiologic studies. Am J Respir Crit Care Med 2007;175:661–6.

40. Hanson C, Rutten E, Wouters EF, Rennard S. Influence of diet and obesity on COPD development and outcomes. Int J Chron Obstruct Pulmon Dis 2014;9:723–33.

41. Lu B, Hiraki LT, Sparks JA, Malspeis S, Chen CY, Awosogba JA, et al. Being overweight or obese and risk of developing rheumatoid arthritis among women: a prospective cohort study. Ann Rheum Dis 2014;73:1914–22.

42. Liu X, Tedeschi SK, Lu B, Zaccardelli A, Speyer CB, Costenbader KH, et al. Long-term physical activity and subsequent risk for rheumatoid arthritis among women: a prospective cohort study. Arthritis Rheumatol 2019;71:1460–71.

43. Garcia-Aymerich J, Lange P, Benet M, Schnohr P, Antó JM. Regular physical activity modifies smoking-related lung function decline and reduces risk of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2007;175:458–63.

44. Li Z, Kesse-Guyot E, Dumas O, Garcia-Aymerich J, Leynaert B, Pison C, et al. Longitudinal study of diet quality and change in asthma symptoms in adults, according to smoking status. Br J Nutr 2017;117:562–71.

45. Hu Y, Sparks JA, Malspeis S, Costenbader KH, Hu FB, Karlson EW, et al. Long-term dietary quality and risk of developing rheumatoid arthritis in women. Ann Rheum Dis 2017;76:1357–64.

46. Bengtsson C, Malspeis S, Orellana C, Sparks JA, Costenbader KH, Karlson EW. Association between menopausal factors and the risk of seronegative and seropositive rheumatoid arthritis: results from the Nurses' Health Studies. Arthritis Care Res (Hoboken) 2017;69:1676–84.

47. Ainsworth BE, Haskell WL, Leon AS, Jacobs DR Jr, Montoye HJ, Sallis JF, et al. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sports Exerc 1993;25:71–80.

48. Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. Am J Clin Nutr 1999;69:243–9.

49. Tager IB, Speizer FE. Risk estimates for chronic bronchitis in smokers: a study of male-female differences. Am Rev Respir Dis 1976;113:619–25.

50. Løkke A, Lange P, Scharling H, Fabricius P, Vestbo J. Developing COPD: a 25 year follow up study of the general population. Thorax 2006;61:935–9.

51. Hagstad S, Bjerg A, Ekerljung L, Backman H, Lindberg A, Rönmark E, et al. Passive smoking exposure is associated with increased risk of COPD in never smokers. Chest 2014;145:1298–304.

52. Coogan PF, Castro-Webb N, Yu J, O'Connor GT, Palmer JR, Rosenberg L. Active and passive smoking and the incidence of asthma in the Black Women's Health Study. Am J Respir Crit Care Med 2015;191:168–76.

53. Strachan DP, Butland BK, Anderson HR. Incidence and prognosis of asthma and wheezing illness from early childhood to age 33 in a national British cohort. BMJ 1996;312:1195–9.

54. Liao KP, Alfredsson L, Karlson EW. Environmental influences on risk for rheumatoid arthritis. Curr Opin Rheumatol 2009;21:279–83.

55. Karlson EW, Lee IM, Cook NR, Manson JE, Buring JE, Hennekens CH. A retrospective cohort study of cigarette smoking and risk of rheumatoid arthritis in female health professionals. Arthritis Rheum 1999;42:910–7.

56. Sparks JA, Karlson EW. The roles of cigarette smoking and the lung in the transitions between phases of preclinical rheumatoid arthritis. Curr Rheumatol Rep 2016;18:15.

57. Seror R, Henry J, Gusto G, Aubin HJ, Boutron-Ruault MC, Mariette X. Passive smoking in childhood increases the risk of developing rheumatoid arthritis. Rheumatology (Oxford) 2018;58:1154–62.

58. Bergström U, Jacobsson LT, Nilsson JÅ, Berglund G, Turesson C. Pulmonary dysfunction, smoking, socioeconomic status and the risk of developing rheumatoid arthritis. Rheumatology (Oxford) 2011;50:2005–13.

59. Ramos-Remus C, Castillo-Ortiz JD, Aguilar-Lozano L, Padilla-Ibarra J, Sandoval-Castro C, Vargas-Serafin CO, et al. Autoantibodies in prediction of the development of rheumatoid arthritis among healthy relatives of patients with the disease. Arthritis Rheumatol 2015;67:2837–44.

60. Syamlal G, Doney B, Mazurek JM. Chronic obstructive pulmonary disease prevalence among adults who have never smoked, by industry and occupation—United States, 2013–2017. MMWR Morb Mortal Wkly Rep 2019;68:303–7.

61. Sparks JA, Barbhaiya M, Tedeschi SK, Leatherwood CL, Tabung FK, Speyer CB, et al. Inflammatory dietary pattern and risk of developing rheumatoid arthritis in women. Clin Rheumatol 2019;38:243–50.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

B Cell Synovitis and Clinical Phenotypes in Rheumatoid Arthritis: Relationship to Disease Stages and Drug Exposure

F. Rivellese,[1] [ID] F. Humby,[1] [ID] S. Bugatti,[2] L. Fossati-Jimack,[1] [ID] H. Rizvi,[3] D. Lucchesi,[1] [ID] G. Lliso-Ribera,[1]
A. Nerviani,[1] R. E. Hands,[1] G. Giorli,[1] B. Frias,[1] G. Thorborn,[1] E. Jaworska,[1] [ID] C. John,[1] K. Goldmann,[1]
M. J. Lewis,[1] [ID] A. Manzo,[2] M. Bombardieri,[1] C. Pitzalis,[1] and the PEAC-R4RA Investigators

Objective. To define the relationship of synovial B cells to clinical phenotypes at different stages of disease evolution and drug exposure in rheumatoid arthritis (RA).

Methods. Synovial biopsy specimens and demographic and clinical data were collected from 2 RA cohorts (n = 329), one of patients with untreated early RA (n = 165) and one of patients with established RA with an inadequate response to tumor necrosis factor inhibitors (TNFi-IR; n = 164). Synovial tissue was subjected to hematoxylin and eosin and immunohistochemical staining and semiquantitative assessment for the degree of synovitis (on a scale of 0–9) and of CD20+ B cell infiltrate (on a scale of 0–4). B cell scores were validated by digital image analysis and B cell lineage–specific transcript analysis (RNA-Seq) in the early RA (n = 91) and TNFi-IR (n = 127) cohorts. Semiquantitative CD20 scores were used to classify patients as B cell rich (≥2) or B cell poor (<2).

Results. Semiquantitative B cell scores correlated with digital image analysis quantitative measurements and B cell lineage–specific transcripts. B cell–rich synovitis was present in 35% of patients in the early RA cohort and 47.7% of patients in the TNFi-IR cohort (*P* = 0.025). B cell–rich patients showed higher levels of disease activity and seropositivity for rheumatoid factor and anti–citrullinated protein antibody in early RA but not in established RA, while significantly higher histologic synovitis scores in B cell–rich patients were demonstrated in both cohorts.

Conclusion. We describe a robust semiquantitative histologic B cell score that closely replicates the quantification of B cells by digital or molecular analyses. Our findings indicate an ongoing B cell–rich synovitis, which does not seem to be captured by standard clinimetric assessment, in a larger proportion of patients with established RA than early RA.

INTRODUCTION

The role of B cells in the pathogenesis of rheumatoid arthritis (RA) is well recognized and has been reinforced by the established efficacy of B cell–depleting treatments (1,2). B cells and B cell effector mechanisms are recognized as a central component of RA synovitis, through local autoantibody production (3,4), osteoclastogenesis/osteoclast activation (5,6), and immune complex–mediated inflammatory responses (7,8). However, the degree of synovial B cell infiltrate is a highly variable phenomenon, ranging from complete absence to a dense distribution within organized infiltrates in up to 40% of patients (9–11) and as such has been examined as a potential source of predictive and prognostic biomarkers in RA. Indeed, recent data from a large cohort of patients with untreated early RA have suggested that a B cell–rich lymphoid synovitis is associated with highly active disease and predictive of radiographic progression (12). However, comparison of these data with other cohorts that either support (13,14) or

refute (15–17) this notion is challenging due to a lack of consistency in quantitative and qualitative assessment of B cell synovitis, the effects of concomitant diverse therapy, and examination of patients at variable disease stages and levels of disease activity.

Importantly, it remains unclear whether an association between B cell synovitis and disease severity is modulated during disease progression, and moreover, whether the prevalence of B cell synovitis remains stable or is enriched through disease evolution and/or cycles of nonresponse to therapy. Since 2008, the Centre for Experimental Medicine and Rheumatology at Queen Mary University of London (QMUL) has been collecting pretreatment synovial tissue as part of a pathobiology-driven patient stratification program in RA. This program includes patients enrolled in 2 multicenter precision-medicine clinical studies collecting pretreatment synovial tissue at specific disease stages: untreated early RA (the Pathobiology of Early Arthritis Cohort [PEAC; http://www.peac-mrc.mds.qmul.ac.uk]) and established RA in patients with an inadequate response to tumor necrosis factor inhibitors (TNFi-IR) (Response–Resistance to Rituximab versus Tocilizumab in RA [R4RA; http://www.r4ra-nihr.whri.qmul.ac.uk]).

In this study, we developed and validated a semiquantitative scoring system focused on the measurement of B cell synovitis in RA. By analyzing the prevalence of B cell synovitis using a robust histologic score, we examined whether the association between B cell synovitis and clinical phenotype is a stable phenomenon during disease evolution or is enriched in a cohort of patients with treatment-resistant established RA.

PATIENTS AND METHODS

Patients. A total of 329 patients fulfilling the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for RA (18) were evaluated at 2 disease stages: early RA and established RA with an inadequate response to TNFi. The early RA cohort consisted of 165 consecutive patients with untreated early RA (disease duration <1 year) recruited as part of the Medical Research Council–funded multicenter PEAC. The established RA cohort consisted of 164 patients with an inadequate response to TNFi from the multicenter R4RA trial. The PEAC-R4RA investigators are listed in Appendix A.

Synovial tissue specimens were obtained from all patients at study entry by ultrasound (US)–guided synovial biopsy (19) or arthroscopic approaches depending on the expertise of the local recruiting center. A minimum of 6 samples for subsequent histologic analysis and 6 samples for RNA extraction were retained from each procedure. Patient demographic characteristics and clinical parameters collected at baseline included sex, disease duration, anti–citrullinated protein antibody (ACPA) and rheumatoid factor (RF) status, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) level, swollen and tender joint counts, global health score (measured on a visual analog scale), and concomitant medications. All patients provided written informed consent, and each study received local ethics approval (PEAC LREC: 05/Q0703/198; R4RA: LREC 12/WA/0307).

Semiquantitative histopathologic scoring and digital image analysis. Formalin-fixed and paraffin-embedded synovial tissue sections (3 μm) were stained with hematoxylin and eosin (H&E) and semiquantitatively assessed for the degree of synovitis (on a scale of 0–9), according to a previously validated score (20). Consecutive sections were subjected to immunohistochemical staining for CD20 (mouse IgG2a anti-human CD20) (clone L26; Dako), CD68 (mouse IgG1 anti-human CD68) (clone KP1; Dako), and CD138 (mouse IgG1 anti-human CD138) (clone MI15; Dako), as previously described (13). Slides were counterstained with hematoxylin and mounted. Images were acquired on an Olympus microscope using cellSens software.

First, the nature and quality of the tissue biopsy specimens were assessed by evaluating tissue morphology by H&E staining and macrophage distribution (CD68 staining). To be considered suitable for further histologic analysis, synovial samples had to present either a clear synovial lining (CD68+ cells in a linear arrangement) or sublining (characteristic vessels and stroma). Samples were rejected and classified as ungraded if there was no recognizable synovial tissue or there were artifactual changes (i.e., tissue folds, cutting and staining artifacts) in immunohistochemical stainings. Samples with synovium were scored semiquantitatively (on a scale of 0–4) for the degree of CD20+ B cell infiltration, adapting a previously described B cell aggregation score (21) (Figures 1A and B). Patients were classified as B cell poor (semiquantitative score 0–1) or B cell rich (score 2–4), as shown in the flow chart in Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract. In B cell–poor samples, the presence of plasma cells was assessed, and samples with a semiquantitative CD138 score of ≥2 underwent an additional staining for CD20 at a deeper cutting level for a final classification as B cell poor or B cell rich. An overview of the cutting protocol at different levels is shown in Supplementary Figure 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract. All samples were assessed by 2 independent observers.

For quantitative analysis using digital image analysis, whole CD20-stained slides were reacquired using a Panoramic 250 High-Throughput Scanner (3DHISTECH). Images were visualized and analyzed on Fiji software (22) using a pipeline developed in-house (by DL) to obtain the following measurements: total stained area, total tissue area, and area fraction (percent of stained area, calculated as total stained area/total tissue area × 100). In the digital image analysis pipeline, the operator first manually selected region(s) of interest corresponding to the synovial tissue detectable on the slide (the "total tissue area"). The script then 1) removed the areas outside the region(s) of interest, 2) performed a color deconvolution to isolate the diaminobenzidine (DAB) color vector (method "H-DAB"),
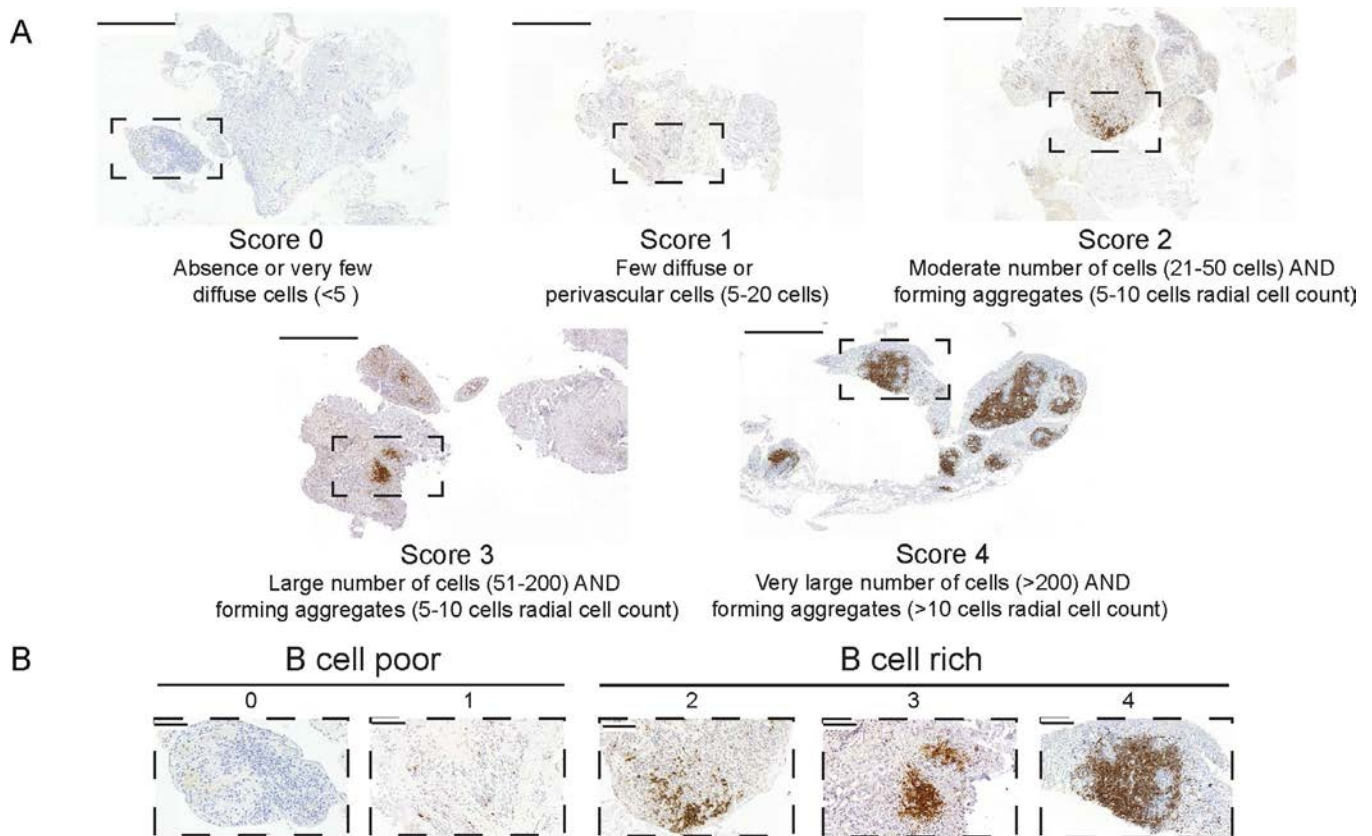
Figure 1. B cell scoring. **A**, Representative images for each semiquantitative CD20 immunohistochemical score in patients with rheumatoid arthritis with an inadequate response to tumor necrosis factor inhibitors. Bars = 500 μm. **B**, Higher-magnification views of the boxed areas in **A**. Bars = 100 μm.

and 3) calculated the intensity threshold (method "Default," manually adjusted if necessary) on the DAB channel. Finally, the software calculated the area included in the threshold ("total stained area"). An example of digital image analysis is shown in Supplementary Figure 3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract.

In a pilot assessment exercise, digital image analysis was performed in parallel in a blinded manner by 2 observers from 2 different centers (QMUL and University of Pavia). The analysis was performed on 15 random samples from the early RA PEAC cohort and replicated on the first 100 consecutive samples from the TNFi-IR R4RA cohort in order to obtain independent measurements. Intraclass correlation coefficients for the level of agreement between the 2 observers and Bland-Altman plots were generated. Following this validation exercise, digital image analysis was performed in the subset of patients from the early RA cohort who underwent RNA sequencing (n = 91) and in the patients in the TNFi-IR cohort who had graded synovial tissue samples (n = 155).

RNA extraction and RNA sequencing. RNA extraction and RNA sequencing were performed as previously described (12) on synovial tissue samples from 94 patients in the early RA PEAC cohort and 128 patients in the TNFi-IR R4RA cohort

for whom RNA was available. For the PEAC cohort, RNA was extracted from synovial tissue homogenized at 4°C in TRIzol reagent, according to the recommendations of the manufacturer (ThermoFisher Scientific, Invitrogen Division), followed by a phenol–chloroform extraction. Briefly, the tissue lysate was mixed vigorously with chloroform before centrifugation. The aqueous phase was removed and mixed with ice-cold isopropanol for 30 minutes. After further centrifugation, the RNA pellet was washed in 70% ethanol before air-drying and resuspension in RNase-free water.

The concentration/purity of RNA samples was measured using a NanoDrop 2000c spectrophotometer (LabTech) and Qubit assay (ThermoFisher Scientific, Invitrogen Division). RNA quality (RNA integrity number) was assessed using an Agilent 2100 Bioanalyzer or Agilent TapeStation system. One microgram of total RNA was used as input material for library preparation using a TruSeq RNA Sample Preparation Kit v2 (Illumina). Generated libraries were amplified with 10 cycles of polymerase chain reaction (PCR). The size of the libraries was confirmed using a 2200 TapeStation system and High-Sensitivity D1K ScreenTape (Agilent Technologies), and their concentration was determined by a quantitative PCR (qPCR)–based method using a Library Quantification kit (KAPA). The libraries were first multiplexed (5 per lane)

and then sequenced on an Illumina HiSeq 2500 system to generate 50 million paired ends with 75 base pair reads.

For the R4RA cohort, the RNA extraction method was the same as for the PEAC, but subsequent to homogenization in TRIzol, a column-based extraction (Zymo Direct-zol RNA MicroPrep) was used according to the recommendations of the manufacturer for the majority of samples (n = 124), while phenol–chloroform extraction as described above was used on a small number of samples (n = 4). Principal components analysis confirmed that there were no major differences between the 2 extraction methods (data not shown). One hundred fifty to five hundred nanograms of total RNA was used as input material for library preparation using NEBNext Ultra RNA Library Prep Kit for Illumina according to the recommendations of the manufacturer (New England Biolabs). Generated libraries were amplified with 13 cycles of PCR. Library quality control was performed with MiSeq Nano QC, and the final concentration was determined by a qPCR-based method using a Library Quantification kit. The libraries were first multiplexed and then sequenced on an Illumina HiSeq 4000 system to generate 50 million paired ends with 150 base pair reads (Genewiz). A summary of the RNA extraction and sequencing methods for the 2 cohorts is included in Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract.

RNA sequencing analysis. For the PEAC cohort, transcript abundance was derived from fastq files over Gencode version24/GRCh38 transcripts using Kallisto version 0.43.0 (23). Transcript abundances and average transcript lengths were imported into R using tximport 1.10.0. Imported abundances were normalized in R, including a correction for average transcript length and incorporating batch, sex, and synovial histology patterns (pathotypes) (12) as model covariates, using DESeq2 1.22.0 (24). Transcript abundances underwent regularized log expression (RLE) transformation. The RNA-Seq data have been deposited in the ArrayExpress database (online at https://www.ebi.ac.uk/arrayexpress; accession no. E-MTAB-6141). Outliers were identified and removed using principal components analysis, leaving 91 samples for further analysis.

For the R4RA cohort, transcript abundance was quantified using Salmon version 0.12.0 (25) with the Gencode version 29/GRCh38 annotation, and quantifications were aggregated to genes using tximport version 1.10.1. Data were subjected to variance-stabilizing transformation to remove the dependency of the variance on the count mean using DESeq2 version 2.16.1 (24). Outliers were identified and removed using principal components analysis, leaving 127 samples for further analysis.

Definition of the B cell–specific gene module. Cell-specific gene modules were obtained as previously described (26). Briefly, we downloaded RLE-normalized FANTOM5 data from http://fantom.gsc.riken.jp/5/data/. Upon selecting primary

cells and tissue and excluding derived cells, stimulated cells and cell line data were Z score–normalized, and the expression of each gene was ranked across all tissues and cells. A specificity score was determined by counting the number of tissues and cells showing increased gene expression (Z score >3 [>3 SD above the mean across all tissues]), so that the most tissue-specific genes would have the lowest specificity scores. Genes were considered specific to a tissue type or cell type when: 1) the level of gene expression was in the top 3 tissues (i.e., rank 1–3); 2) the Z score was >5 (i.e., >5 SD above the mean expression across all tissues); and 3) the specificity score was <10 tissues. Gene modules for different cell types were consistent with lists of genes previously published by the FANTOM5 consortium for several cell types (27,28). The list of B cell–specific genes (B cell module; including Z scores, specificity scores, and rank) is shown in Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract.

Synovial RNA-Seq B cell module scores. The B cell module was used to calculate synovial RNA-Seq B cell module scores, by applying singular value decomposition to synovial RNA-Seq data (29). Briefly, the normalized gene expression matrix for the patients was filtered to include only the B cell module genes. Next, the matrix was subjected to singular value decomposition, and the first principal component score was taken as the module expression or "module score" for plotting. B cell module scores were analyzed for correlation against histologic markers in synovial tissue using Spearman's correlation.

Statistical analysis. Measures of central tendency and dispersion and statistical tests used are indicated in each figure legend. Generally, the following statistical tests were used: Mann-Whitney test for comparison between 2 groups, one-way analysis of variance with Bonferroni post hoc adjustment for comparison between multiple groups, chi-square test for proportions, and Spearman's test for correlations. Statistical analyses were performed using IBM SPSS version 20 and RStudio version 0.99.486. *P* values less than 0.05 were considered significant.

RESULTS

Demographic and clinical features of the patients with early RA and those with established RA. A total of 329 patients with RA were included in this study: 165 with untreated early RA (PEAC) and 164 with an inadequate response to TNFi (R4RA trial). Table 1 summarizes the patient characteristics. Both cohorts included patients with highly active RA (mean Disease Activity Score in 28 joints [DAS28] >5.1), with no differences in terms of disease activity (DAS28 and its components) and other clinical features, except for a higher prevalence of female patients, higher CRP levels, and higher proportion of ACPA-positive

Table 1. Characteristics of the patients in the early RA and TNFi-IR cohorts*

	Early RA (n = 165)	TNFi-IR (n = 164)	P†
Age, years	53.2 ± 15.3	54.6 ± 13.3	NS
Sex, % female	66.1	79.9	0.005
Disease duration, years	0.5 ± 0.6	12.8 ± 10.8	<0.001
DAS28	5.7 ± 1.4	5.63 ± 1.27	NS
Tender joint count	11.6 ± 7.6	11.85 ± 7.8	NS
Swollen joint count	7.5 ± 5.6	6.92 ± 5.10	NS
Global health score (100-mm VAS)	61.5 ± 26.9	66.33 ± 24.99	NS
ESR, mm/hour	37.8 ± 28.4	34.47 ± 25.95	NS
CRP, mg/liter	17.6 ± 27.1	21.13 ± 27.97	0.01
ACPA positive, %‡	64.2	75.3	0.03
RF positive, %	65.5	71.4	NS
No. of conventional synthetic DMARDs received, %			NA
0	100	2.4	
1	0	67.7	
2	0	21.3	
3	0	8.4	
Receiving steroids at the time of biopsy, %	0	42.2	NA
Joint biopsied, %			NS
Wrist	64.8	60.4	
Knee	18.2	23.8	
MCP	13.9	11	
Other	3.1	4.8	
Joint size, %			NS
Large	25.2	27.4	
Small	74.8	72.6	
Synovial sampling technique, %			NA
US-guided	100	86	
Arthroscopy	0	14	

* Except where indicated otherwise, values are the mean ± SD. RA = rheumatoid arthritis; TNFi-IR = inadequate response to tumor necrosis factor inhibitors; NS = not significant; DAS28 = Disease Activity Score in 28 joints; VAS = visual analog scale; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; RF = rheumatoid factor; DMARDs = disease-modifying antirheumatic drugs; NA = not applicable; MCP = metacarpophalangeal; US = ultrasound.
† By chi-square test or Mann-Whitney test.
‡ Anti–citrullinated protein antibodies (ACPAs) were measured using a clinically available standard pathology laboratory anti–cyclic citrullinated peptide antibody 2 assay.

patients in the cohort with established RA. There were no differences between the cohorts with regard to the type of joint biopsied. Importantly, >70% of the samples were obtained from small joints (i.e., wrists or metacarpophalangeal joints), and >90% of the procedures were performed under US guidance. The procedures had an excellent success rate in terms of retrieval of gradable tissue, ranging from 86.7% (143 of 165) in the early RA cohort to 94.5% (155 of 164) in the TNFi-IR cohort. These results demonstrate that both the early RA and established RA cohorts included patients with active RA, although the patients with established RA showed more aggressive features, such as a greater proportion of

patients with autoantibody positivity and higher levels of markers of inflammation, which is compatible with their recruitment after anti-TNF failure.

Validation of semiquantitative synovial B cell scoring by digital image analysis. To quantify the presence of B cells in synovia, we adapted a semiquantitative scoring method (scale of 0–4) previously described by our group (21). Representative examples of tissue samples for each score are shown in Figures 1A and B, and a flow chart of the evaluation of the biopsy samples is shown in Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract. In order to validate the semiquantitative B cell scores against an objective measurement, we used computer-aided automated digital image analysis. To this end, whole slide images were acquired, and digital image analysis was used to obtain the following measurements: total tissue area (in μm²), total stained area, and area fraction (percent of stained area, calculated as total stained area/total tissue area × 100). An example of the digital image analysis approach is shown in Supplementary Figure 3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract. First, we aimed to evaluate the reproducibility of digital image analysis, by performing the analyses in 2 different centers (QMUL and University of Pavia) in a blinded manner on 15 randomly selected synovial samples from the early RA cohort and 100 consecutive samples from the TNFi-IR cohort. These analyses showed an excellent agreement rate, with an intraclass correlation coefficient of 0.97 in the early RA cohort (Figure 2A) and 0.96 in the TNFi-IR cohort (Figure 2B).

Having confirmed the reproducibility of digital image analysis, we next aimed to evaluate the correlation between the semiquantitative and digital image analysis (CD20+ area fraction) histologic B cell scores in 91 patients in the early RA cohort and 164 patients in the TNFi-IR cohort. Our results demonstrated a strong correlation between the semiquantitative and digital image analysis CD20 scores (Spearman's r = 0.93 in the early RA cohort and 0.88 in the TNFi-IR cohort; P < 0.0001) (Figures 2C and D). In addition, we determined the mean CD20+ area fraction in patients classified as B cell rich and those classified as B cell poor according to the semiquantitative scores and demonstrated a significantly higher CD20+ area fraction in B cell–rich patients, as shown in Figures 2E and F (mean ± SD area fraction 1.4 ± 1.6 in B cell–rich patients and 0.02 ± 0.05 in B cell–poor patients in the early RA cohort [P < 0.0001] and 1.3 ± 1.9 in B cell–rich patients and 0.04 ± 0.1 in B cell–poor patients in the TNFi-IR cohort [P < 0.0001]). Overall, our results demonstrate that the semiquantitative CD20 score presented both as a raw score and when applied to stratify patients into B cell–rich and B cell–poor cohorts reliably reflects quantitative measures of CD20 B cell infiltration assessed using computer-aided digital image analysis.
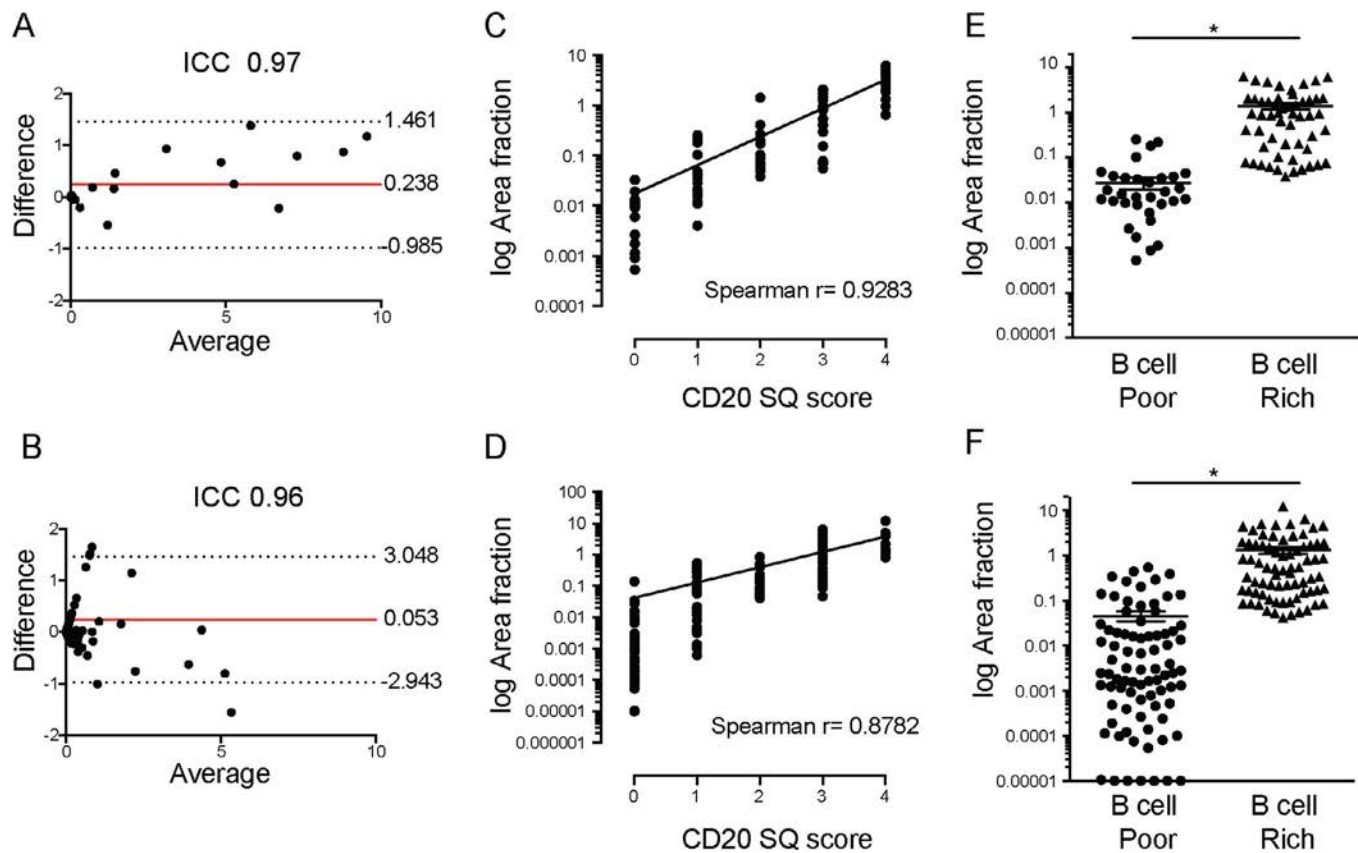
Figure 2. Semiquantitative (SQ) B cell scores and digital image analysis. **A** and **B**, Bland-Altman plots showing the difference between and average of 2 measurements of CD20+ B cell area fraction (percent of stained area, calculated as total stained area/total tissue area × 100) in the cohort of patients with early rheumatoid arthritis (RA; n = 15) (**A**) and in the cohort of patients with established RA with an inadequate response to tumor necrosis factor inhibitors (TNFi-IR; n = 100) (**B**), obtained by 2 independent observers in 2 different centers (Queen Mary University of London and University of Pavia). Solid lines indicate the mean; dotted lines indicate the 95% confidence interval. ICC = intraclass correlation coefficient. **C** and **D**, Correlation between semiquantitative CD20+ B cell score and CD20+ B cell area fraction in the early RA cohort (n = 91) (**C**) and the TNFi-IR cohort (n = 155) (**D**). **E** and **F**, CD20+ B cell area fraction in the early RA cohort (n = 91) (**E**) and the TNFi-IR cohort (n = 155) (**F**) classified as B cell poor or B cell rich. Symbols represent individual patients; horizontal lines and error bars show the mean ± SD. * = $P < 0.05$ by Mann-Whitney test. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract.

Correlation between semiquantitative and quantitative histologic scores and B cell–lineage genes. Next, we evaluated the relationship between CD20+ B cell histologic scores (semiquantitative and digital image analysis) and B cell–specific gene expression levels. To this end, we correlated histologic scores for B cells with CD20 gene expression levels and RNA-Seq B cell module scores obtained by RNA sequencing of synovial tissue from 91 patients in the early RA cohort and 127 patients in the TNFi-IR cohort. Our results demonstrated a strong correlation between the semiquantitative CD20 score and B cell module score and between the semiquantitative CD20 score and CD20 gene expression, both in the early RA cohort (Figure 3A) and the TNFi-IR cohort (Figure 3B). Similarly, we observed a positive correlation between the digital image analysis CD20 area fraction and the B cell module score and between the digital image analysis CD20 area fraction and CD20 gene expression in the early RA cohort (Figure 3C) and

the TNFi-IR cohort (Figure 3D). Finally, when patients were segregated into B cell–rich and B cell–poor subgroups, we demonstrated significantly higher levels of B cell module scores and CD20 gene expression in B cell–rich patients in both cohorts (Figures 3E and F). Overall, our results demonstrate that the semiquantitative CD20 score presented both as a raw score and when applied to stratify patients into B cell–rich and B cell–poor cohorts accurately reflects the levels of synovial B cell gene expression.

Clinical features of B cell–rich patients and B cell–poor patients at different disease stages. Next, we applied the semiquantitative CD20 score and algorithm shown in Figure 1 and Supplementary Figure 1 to stratify patients from the 2 cohorts into B cell–rich and B cell–poor categories. We assessed the prevalence of B cell–rich synovitis at different stages of disease evolution and drug exposure, i.e., in untreated early RA versus established RA with
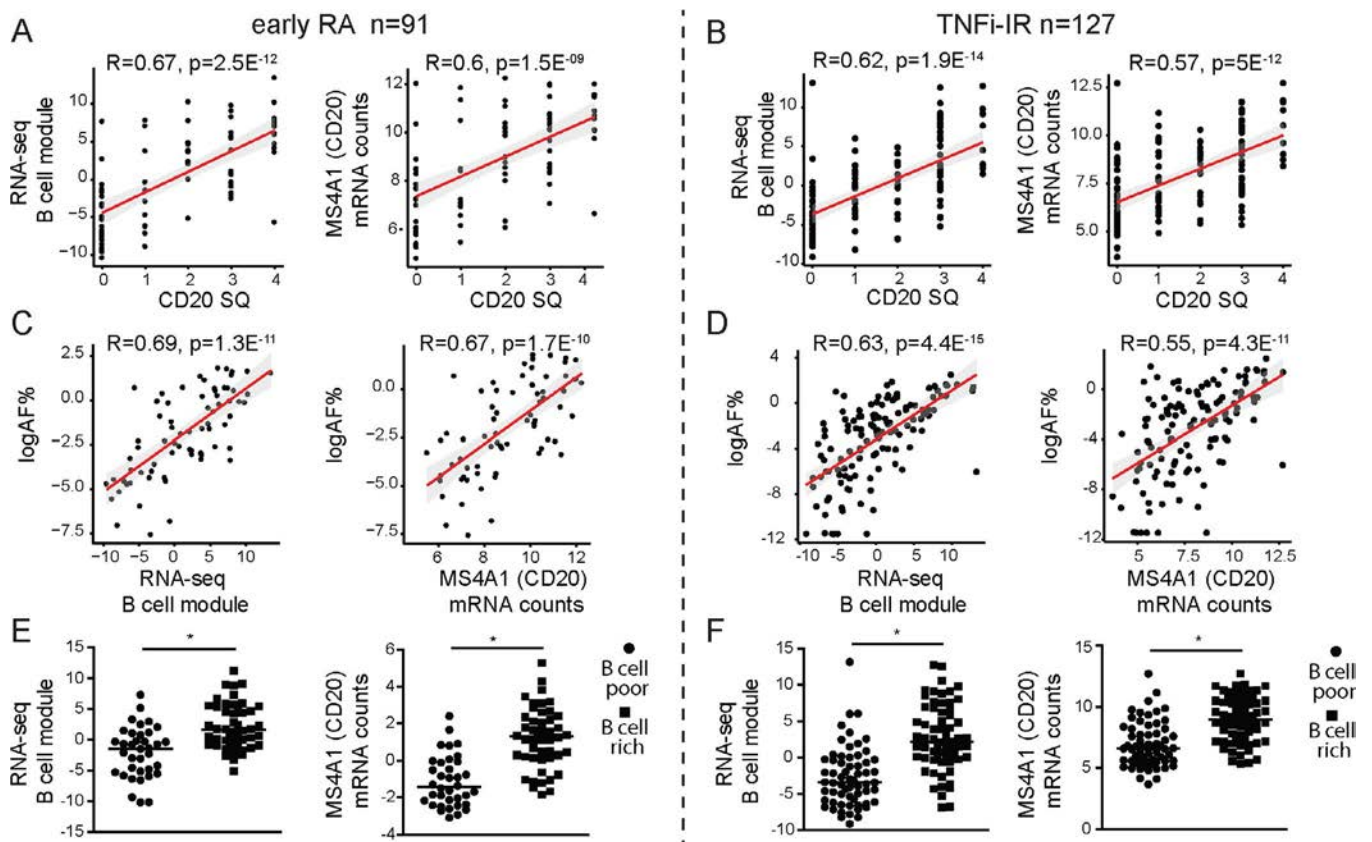
Figure 3. Correlation of semiquantitative (SQ) B cell scores with RNA-Seq B cell module scores. **A** and **B**, Correlations between the semiquantitative CD20+ B cell score and the RNA-Seq B cell module and between the semiquantitative CD20+ B cell score and CD20/MS4A1 gene expression levels in the cohort of patients with early rheumatoid arthritis (RA) (**A**) and the cohort of patients with established RA with an inadequate response to tumor necrosis factor inhibitors (TNFi-IR) (**B**). **C** and **D**, Correlations between the CD20+ B cell area fraction (AF) (percent of stained area, calculated as total stained area/total tissue area × 100) and the RNA-Seq B cell module and between the CD20+ B cell area fraction and CD20/MS4A1 gene expression levels in the early RA cohort (**C**) and the TNFi-IR cohort (**D**). In **A**–**D**, symbols represent individual patients (n = 91 for early RA and 127 for TNFi-IR); Lines and shading indicate the regression line and 95% confidence interval. **E** and **F**, RNA-Seq B cell module and CD20/MS4A1 gene expression levels in B cell–poor and B cell–rich patients in the early RA cohort (n = 91) (**E**) and TNFi-IR cohort (n = 127) (**F**). Symbols represent individual patients; horizontal lines show the mean. * = $P < 0.05$ by Mann-Whitney test. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41184/abstract.

an inadequate response to TNFi. Interestingly, we demonstrated a significantly higher prevalence of B cell–rich synovitis in patients with established RA with an inadequate response to TNFi (47.7% in the TNFi-IR cohort and 35% in the untreated early RA cohort; $P = 0.025$) (Figure 4A), suggesting an increase in the proportion of B cell–rich samples in longstanding treatment-resistant RA.

Furthermore, we examined the relationship between clinical phenotype and B cell–rich synovitis at each disease stage (Table 2). Consistent with previously published data (12), compared to B cell–poor patients, patients with B cell–rich synovitis in the early RA cohort showed higher levels of synovial inflammation, as measured by the Krenn synovitis score ($P < 0.0001$), which is a well-validated composite score measuring the enlargement of the lining cell layer, the cellular density of synovial stroma, and leukocyte infiltrate. Accordingly, B cell–rich patients had significantly higher levels of other immune cells (T cells, macrophages, and plasma cells). Clinically, B cell–rich patients showed higher disease activ-

ity ($P = 0.005$ for DAS28 and $P = 0.048$ for swollen joint count), markers of inflammation ($P = 0.001$ for ESR and $P = 0.012$ for CRP level), and a higher prevalence of autoantibody positivity ($P = 0.024$ for ACPA and $P = 0.023$ for RF) than B cell–poor patients.

Conversely, in the cohort of patients with established disease (TNFi-IR), while the significant association of B cell–rich synovitis with histologic synovitis and immune cell infiltration was maintained (Table 2 and Figure 4B), the only significant difference in terms of systemic inflammation or disease activity was a higher CRP level in B cell–rich patients than in B cell–poor patients (Table 2 and Figure 4C) ($P = 0.001$). There were no differences in terms of disease activity as measured by DAS28 (Figure 4D) and its components (Table 2). Importantly, there were also no differences in the use of synthetic disease-modifying antirheumatic drugs (DMARDs) or steroids between B cell–rich and B cell–poor patients. These data suggest that the clinical phenotype of highly active, aggressive RA associated with a B cell–rich synovitis at disease initiation
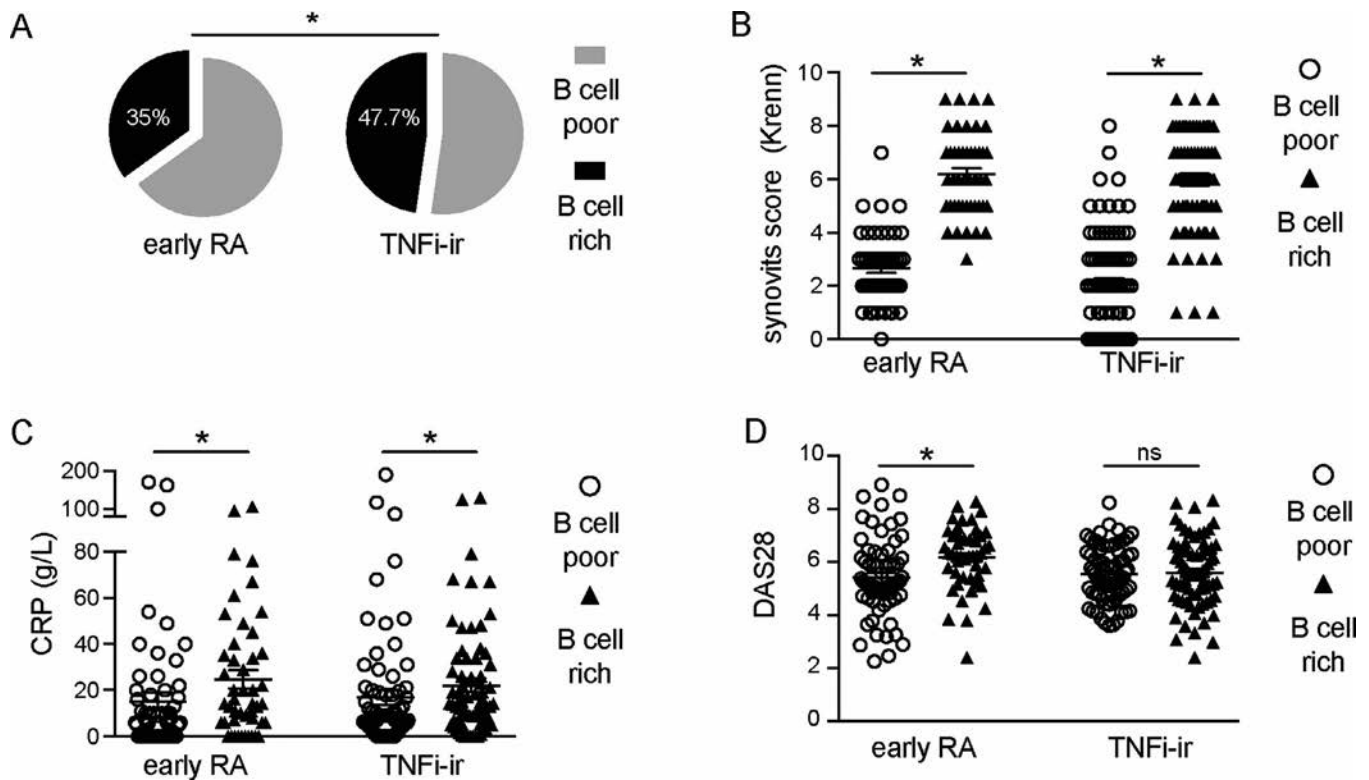
Figure 4. B cell–rich synovitis at different rheumatoid arthritis (RA) disease stages. **A**, Prevalence of B cell–rich synovitis in the cohort of patients with early RA and in the cohort of patients with established RA with an inadequate response to tumor necrosis factor inhibitors (TNFi-IR). **B**–**D**, Krenn synovitis score (**B**), C-reactive protein (CRP) level (**C**), and Disease Activity Score in 28 joints (DAS28) using the erythrocyte sedimentation rate (**D**) in B cell–poor and B cell–rich patients in the early RA cohort and the TNFi-IR cohort. Symbols represent individual patients (n = 143 for early RA and 155 for TNFi-IR); horizontal lines and error bars show the median and interquartile range in **B** and the mean ± SD in **C** and **D**. * = $P < 0.05$ by chi-square test in **A**; by Mann-Whitney test in **B**–**D**. NS = not significant.

is subsequently lost with disease progression and/or modulation of synovial pathobiology by concomitant therapy.

To further explore the association of synovial histopathology with clinical phenotypes, we compared clinical features in patients classified according to semiquantitative scores for T cells, macrophages, and plasma cells (Supplementary Tables 3–6, available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.41184/abstract). In the early RA cohort, higher levels of markers of inflammation were observed in patients with higher infiltration of all immune cells. On the contrary, a higher prevalence of autoantibody positivity was observed only in patients who, in addition to B cells, were rich in synovial T cells and plasma cells, but no relationship was observed with CD68L or CD68SL macrophage populations. Notably, on the other hand, higher disease activity measured by DAS28 was present only in patients with higher sublining macrophages and plasma cell infiltration. In the cohort of patients with established RA, patients with synovitis rich in T cells, sublining macrophages, and plasma cells showed significantly higher CRP levels, while there were no differences in any other clinical parameter. These data add new insights into the association of immune cells with specific disease features, while confirming a lack of association between synovitis

and clinical phenotypes in the cohort of patients with established RA.

DISCUSSION

In this study, we explored the prevalence and relationship to clinical phenotypes of B cell synovitis in a large, well-characterized cohort of 329 patients at 2 different stages of disease evolution and drug exposure. We validated and applied a semiquantitative B cell synovitis score that accurately reflects quantitative measures of both cellular and molecular B cell infiltrate, and demonstrated that the prevalence of B cell synovitis was enriched in patients with treatment-resistant established disease. Additionally, although patients with established RA display a more severe clinical phenotype, which could explain the enrichment in B cells, we observed that the significant association of a B cell–rich synovitis with a clinical phenotype of highly active aggressive RA in untreated early disease was diluted at later stages.

Previous data have suggested that ~40% of RA patients present synovitis characterized by a B cell–rich infiltrate, both in early (9–11) and late-stage disease (6). However, the clinical relevance of B cell synovitis has been a subject of controversy, with

Table 2. Clinical phenotype of B cell–poor versus B cell–rich patients at different RA disease stages*

	Early RA (n = 143)			TNFi-IR (n = 155)		
	B cell poor (n = 93 [65.0%])	B cell rich (n = 50 [35.0%])	P†	B cell poor (n = 81 [52.3%])	B cell rich (n = 74 [47.7%])	P†
DAS28	5.6 ± 1.38	6.1 ± 1.2	0.005	5.6 ± 1.3	5.7 ± 1.2	NS
Tender joint count	11.3 ± 7.3	13 ± 7.7	NS	12.5 ± 7.9	11.2 ± 7.8	NS
Swollen joint count	7.1 ± 5.4	8.8 ± 5.7	0.048	6.5 ± 5.1	7.4 ± 5.1	NS
Global health score (100-mm VAS)	60.6 ± 28.6	64.8 ± 25.2	NS	65.2 ± 23.2	66.8 ± 27.3	NS
ESR, mm/hour	34.6 ± 28.9	49.5 ± 28.9	0.001	31.7 ± 24.4	38.3 ± 27.2	NS
CRP, mg/liter	17.5 ± 32.7	21.5 ± 24.2	0.012	19.9 ± 35.9	26.4 ± 26.8	0.001
ACPA positive, %‡	59.8	78.4	0.024	73.1	76.4	NS
RF positive, %	62	80.4	0.023	70	72.2	NS
No. of conventional synthetic DMARDs received, %			NA			NS
0	100	100		1.2	4.1	
1	0	0		70.4	67.6	
2	0	0		21	23	
3	0	0		7.4	5.4	
Receiving steroids at the time of biopsy, %	0	0	NA	39.7	41.9	NS
Synovitis score	3 ± 1	6 ± 1	<0.0001	2 ± 2	6 ± 2	<0.0001
Semiquantitative CD3 score ≥2, %	6.5	70.8	<0.0001	9.1	90.9	<0.0001
Semiquantitative CD68L score ≥2, %	12.2	46	<0.0001	6.2	93.8	<0.0001
Semiquantitative CD68SL score ≥2, %	24.2	84	<0.0001	17.1	82.9	<0.0001
Semiquantitative CD138 score ≥2, %	4.4	78	<0.0001	9.3	90.7	<0.0001

* Patients with ungraded synovial biopsy samples were excluded. Except where indicated otherwise, values are the mean ± SD. CD68L = CD68 lining; CD68SL = CD68 sublining (see Table 1 for other definitions).
† By Mann-Whitney test or Fisher's exact test.
‡ ACPAs were measured using a clinically available standard pathology laboratory anti–cyclic citrullinated peptide antibody 2 assay.

previous reports presenting conflicting data on the relationship of B cell synovitis with clinical phenotype (15–17). We attempted to overcome the limitations presented by previous analyses (30,31) by implementing a consistent and validated semiquantitative B cell score. This score was shown to be a reliable measure of total B cell content by validation against objective histologic and molecular quantitative measurements (digital image analysis and RNA-Seq transcript lineage analysis, respectively) in large, well-defined cohorts of patients. Although the validation of semiquantitative histologic synovial T cell and macrophage scores using digital image analysis has been explored previously (32,33), validation of semiquantitative CD20 scores against both digital image analysis and B cell–lineage transcript analysis has not. In particular, the evaluation against B cell transcripts represents an invaluable confirmation that the semiquantitative histologic score used to grade limited tissue sections accurately reflects the B cell content of the whole specimen (6 additional biopsy specimens pooled for RNA extraction), therefore representing a reliable method for the assessment of B cells in synovia.

The significant association of B cell–rich synovitis with highly active seropositive RA in untreated early disease supports previous observations (12) and suggests that synovial tissue cellular infiltration could help define histologic RA subsets, similarly to what

has been described for seronegative and seropositive RA (34). Consistent with these observations, the presence of B cell–rich lymphoid aggregates in early RA has recently been shown to predict treatment response to conventional synthetic DMARDs and radiographic progression (12) and to enhance clinical classification and prognostic/treatment response algorithms (35). In established RA, the associations of B cell synovitis with local inflammation (histologic synovitis) and CRP persisted, and patients showed a more aggressive clinical phenotype overall, but there was a lack of association of B cell–rich synovitis with clinical markers of disease activity such as the DAS28. This lack of association suggests that current standard clinimetric assessment of disease activity may be too insensitive to detect ongoing histologic inflammation, as it might be confounded by factors such as concomitant therapy or comorbidities and their impact on patient-reported outcomes and disease activity scores. To overcome such limitations, future analyses may include alternative measures of disease activity, including, for example, the reweighted 2-component imaging-derived disease activity score (36), and determine their relationship with histologic synovitis. To date, in fact, very little is known about the association of histologic synovitis with clinical phenotypes. Importantly, our observations could provide an explanation for the variable associations observed between clinical phenotype

and synovial pathobiology in other disease cohorts where patients with varying disease duration and/or therapies were analyzed together (15,17).

Such disconnect between ongoing synovial inflammation and clinical disease activity could explain the reported radiographic progression in RA patients with low disease activity treated with synthetic and/or biologic DMARDs (37). However, future analyses, including the outcome of the R4RA trial, will help us to understand whether B cell synovitis continues to drive radiographic progression in late-stage disease, as recently demonstrated in early disease (12), and to further dissect the association of clinical parameters with histologic synovitis. An alternative explanation for such disconnect could be that different pathogenic mechanisms drive local synovial inflammation in early RA and in established RA and manifest different levels of clinical synovitis. For example, T cell cytokine signatures described in early RA do not appear to be present in later stages (38), while different immune cell infiltration has been found in synovial samples from patients with active RA compared with samples from patients with end-stage destructive RA obtained at joint replacement (39). Importantly, however, our cohort of patients with established RA did not include patients with end-stage RA, and the enriched B cell infiltrate in this cohort suggests that adaptive immunity continues to play a pivotal role in active established RA. Nonetheless, alternative non–B cell–driven synovitis may be present in different patients.

This study has some limitations. First, while we have analyzed 2 large cohorts of patients characterized by specific disease stages and drug exposure, our results do not include data on sequential biopsy or longitudinal clinical outcomes. Such data will emerge following completion of the R4RA trial and will be critical to determine whether our observation of an enrichment of B cell synovitis in patients with an inadequate response to TNFi is a treatment effect or is related to a propensity for B cell–rich patients to fail to respond to TNFi therapy.

Second, synovial sampling was performed by both US-guided methods and arthroscopy, which could be considered a confounding factor due to the heterogeneity of sample retrieval methods and/or patient selection. However, recent data have shown an equivalence in quality outcomes of arthroscopy and US-guided approaches (40). Moreover, the use of US-guided biopsy for the majority of the procedures ensured that patients were included in the studies irrespective of joint distribution, a factor that is likely to account for the difference in results compared to previous studies that used only large joint arthroscopy (15), which is known to bias recruitment to patients with more severe disease (41), particularly in early RA.

Finally, we acknowledge that many other immune cells play a role in RA synovitis, driving both disease activity and progression and treatment response (42–44). We found that B cell infiltration was accompanied by T cells, sublining macrophages, and plasma cells, which is consistent with previously reported data (12). However, when comparing the clinical parameters in the

early RA cohort, we observed that higher levels of markers of inflammation were associated with higher scores for all immune cells, but that a higher prevalence of autoantibody positivity was associated exclusively with higher B cell, T cell, and plasma cell scores. Higher disease activity was associated with higher B cell, sublining macrophage, and plasma cell scores. In the cohort of patients with established RA, these additional analyses confirmed the exclusive association of CRP levels with synovitis, without differences in any other parameters, once again suggesting that standard clinimetrics seems unable to pick up ongoing synovitis in established RA.

Overall, these data provide insights into the specific associations of immune cells with clinical phenotypes and highlight the relevance of B cells as markers of ongoing synovitis associated with a clinical phenotype of aggressive disease in early RA. However, alternative histologic scoring systems that aim to more comprehensively assess immune cell infiltration are available, including an integrated histologic scoring system (10,12) that we have recently developed and that has been verified by combined histologic and -omics approaches (11,45). Although it is likely that in the near future integrated histologic and molecular analyses will identify novel or previously unrecognized pathways mediating treatment response or resistance, in order to translate those studies to routine clinical care there is a need for well-validated histopathologic scores that can be easily applied for patient classification, such as the B cell synovitis score described herein, particularly if it were demonstrated to have clinical utility for stratifying patients to receive rituximab.

In conclusion, we described the application of a robust, validated B cell synovitis score that closely replicates the quantification of B cells using either digital image analysis or RNA-Seq analysis and identifies a significant enrichment of B cell synovitis in end-stage treatment-resistant RA. Furthermore, we demonstrated a variable association between B cell synovitis and clinical disease activity measurements at different stages of disease progression and drug exposure. In particular, the presence of synovial B cells in established treatment-resistant RA helps identify patients with ongoing synovial inflammation that is not detected by standard clinimetric assessment. Overall, our study confirms the relevance of synovial B cells in RA and suggests that the classification of patients into B cell rich or B cell poor can contribute to patient stratification.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

REFERENCES

1. Edwards JC, Szczepański L, Szechiński J, Filipowicz-Sosnowska A, Emery P, Close DR, et al. Efficacy of B-cell–targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med 2004;350:2572–81.

2. Kramm H, Hansen KE, Gowing E, Bridges A. Successful therapy of rheumatoid arthritis with rituximab: renewed interest in the role of B cells in the pathogenesis of rheumatoid arthritis. J Clin Rheumatol 2004;10:28–32.

3. Humby F, Bombardieri M, Manzo A, Kelly S, Blades MC, Kirkham B, et al. Ectopic lymphoid structures support ongoing production of class-switched autoantibodies in rheumatoid synovium. PLoS Med 2009;6:e1.

4. Rosengren S, Wei N, Kalunian KC, Zvaifler NJ, Kavanaugh A, Boyle DL. Elevated autoantibody content in rheumatoid arthritis synovia with lymphoid aggregates and the effect of rituximab. Arthritis Res Ther 2008;10:R105.

5. Yeo L, Toellner KM, Salmon M, Filer A, Buckley CD, Raza K, et al. Cytokine mRNA profiling identifies B cells as a major source of RANKL in rheumatoid arthritis. Ann Rheum Dis 2011;70:2022–8.

6. Harre U, Georgess D, Bang H, Bozec A, Axmann R, Ossipova E, et al. Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin. J Clin Invest 2012;122:1791–802.

7. Laurent L, Anquetil F, Clavel C, Ndongo-Thiam N, Offer G, Miossec P, et al. IgM rheumatoid factor amplifies the inflammatory response of macrophages induced by the rheumatoid arthritis-specific immune complexes containing anticitrullinated protein antibodies. Ann Rheum Dis 2015;74:1425–31.

8. Clavel C, Nogueira L, Laurent L, Iobagiu C, Vincent C, Sebbag M, et al. Induction of macrophage secretion of tumor necrosis factor α through Fcγ receptor IIa engagement by rheumatoid arthritis–specific autoantibodies to citrullinated proteins complexed with fibrinogen. Arthritis Rheum 2008;58:678–88.

9. Bombardieri M, Lewis M, Pitzalis C. Ectopic lymphoid neogenesis in rheumatic autoimmune diseases. Nat Rev Rheumatol 2017;13:141–54.

10. Pitzalis C, Kelly S, Humby F. New learnings on the pathophysiology of RA from synovial biopsies. Curr Opin Rheumatol 2013;25:334–44.

11. Dennis G Jr, Holweg CT, Kummerfeld SK, Choy DF, Setiadi AF, Hackney JA, et al. Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics. Arthritis Res Ther 2014;16:R90.

12. Humby F, Lewis M, Ramamoorthi N, Hackney JA, Barnes MR, Bombardieri M, et al. Synovial cellular and molecular signatures stratify clinical response to csDMARD therapy and predict radiographic progression in early rheumatoid arthritis patients. Ann Rheum Dis 2019;78:761–72.

13. Bugatti S, Manzo A, Vitolo B, Benaglio F, Binda E, Scarabelli M, et al. High expression levels of the B cell chemoattractant CXCL13

14. Orr C, Najm A, Biniecka M, McGarry T, Ng CT, Young F, et al. Synovial immunophenotype and anti–citrullinated peptide antibodies in rheumatoid arthritis patients. Arthritis Rheumatol 2017;69:2114–23.

15. Cantaert T, Kolln J, Timmer T, van der Pouw Kraan TC, Vandooren B, Thurlings RM, et al. B lymphocyte autoimmunity in rheumatoid synovitis is independent of ectopic lymphoid neogenesis. J Immunol 2008;181:785–94.

16. Van de Sande MG, Thurlings RM, Boumans MJ, Wijbrandts CA, Modesti MG, Gerlag DM, et al. Presence of lymphocyte aggregates in the synovium of patients with early arthritis in relationship to diagnosis and outcome: is it a constant feature over time? Ann Rheum Dis 2011;70:700–3.

17. Thurlings RM, Wijbrandts CA, Mebius RE, Cantaert T, Dinant HJ, van der Pouw-Kraan TC, et al. Synovial lymphoid neogenesis does not define a specific clinical rheumatoid arthritis phenotype. Arthritis Rheum 2008;58:1582–9.

18. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.

19. Kelly S, Humby F, Filer A, Ng N, Di Cicco M, Hands RE, et al. Ultrasound-guided synovial biopsy: a safe, well-tolerated and reliable technique for obtaining high-quality synovial tissue from both large and small joints in early arthritis patients. Ann Rheum Dis 2015;74:611–7.

20. Krenn V, Morawietz L, Burmester GR, Kinne RW, Mueller-Ladner U, Muller B, et al. Synovitis score: discrimination between chronic low-grade and high-grade synovitis. Histopathology 2006;49:358–64.

21. Manzo A, Paoletti S, Carulli M, Blades MC, Barone F, Yanni G, et al. Systematic microanatomical analysis of CXCL13 and CCL21 in situ production and progressive lymphoid organization in rheumatoid synovitis. Eur J Immunol 2005;35:1347–59.

22. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;9:676–82.

23. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 2016;34:525–7.

24. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

25. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017;14:417–9.

26. Lewis MJ, Barnes MR, Blighe K, Goldmann K, Rana S, Hackney JA, et al. Molecular portraits of early rheumatoid arthritis identify clinical and treatment response phenotypes. Cell Rep 2019;28:2455–70.

27. Motakis E, Guhl S, Ishizu Y, Itoh M, Kawaji H, de Hoon M, et al. Redefinition of the human mast cell transcriptome by deep-CAGE sequencing. Blood 2014;123:e58–67.

28. The FANTOM Consortium and the RIKEN PMI and CLST (DGT). A promoter-level mammalian expression atlas. Nature 2014;507:462–70.

29. Langfelder P, Horvath S. Eigengene networks for studying the relationships between co-expression modules. BMC Syst Biol 2007;1:54.

30. Humby F, Kelly S, Bugatti S, Manzo A, Filer A, Mahto A, et al. Evaluation of minimally invasive, ultrasound-guided synovial biopsy techniques by the OMERACT filter: determining validation requirements. J Rheumatol 2016;43:208–13.

31. Smith MD, Baeten D, Ulfgren AK, McInnes IB, Fitzgerald O, Bresnihan B, et al. Standardisation of synovial tissue infiltrate analysis: how far

in rheumatoid synovium are a marker of severe disease. Rheumatology (Oxford) 2014;53:1886–95.

have we come? How much further do we need to go? Ann Rheum Dis 2006;65:93–100.

32. Kraan MC, Haringman JJ, Ahern MJ, Breedveld FC, Smith MD, Tak PP. Quantification of the cell infiltrate in synovial tissue by digital image analysis. Rheumatology (Oxford) 2000;39:43–49.

33. Cunnane G, Bjork L, Ulfgren AK, Lindblad S, FitzGerald O, Bresnihan B, et al. Quantitative analysis of synovial membrane inflammation: a comparison between automated and conventional microscopic measurements. Ann Rheum Dis 1999;58:493–9.

34. Ajeganova S, Huizinga TW. Seronegative and seropositive RA: alike but different? Nat Rev Rheumatol 2015;11:8–9.

35. Lliso-Ribera G, Humby F, Lewis M, Nerviani A, Mauro D, Rivellese F, et al. Synovial tissue signatures enhance clinical classification and prognostic/treatment response algorithms in early inflammatory arthritis and predict requirement for subsequent biologic therapy: results from the pathobiology of early arthritis cohort (PEAC). Ann Rheum Dis 2019;78:1642–52.

36. Hensor EM, McKeigue P, Ling SF, Colombo M, Barrett JH, Nam JL, et al. Validity of a two-component imaging-derived disease activity score for improved assessment of synovitis in early rheumatoid arthritis. Rheumatology (Oxford) 2019. E-pub ahead of print.

37. Landewé R, van der Heijde D, Klareskog L, van Vollenhoven R, Fatenejad S. Disconnect between inflammation and joint destruction after treatment with etanercept plus methotrexate: results from the trial of etanercept and methotrexate with radiographic and patient outcomes. Arthritis Rheum 2006;54:3119–25.

38. Raza K, Falciani F, Curnow SJ, Ross EJ, Lee CY, Akbar AN, et al. Early rheumatoid arthritis is characterized by a distinct and transient synovial fluid cytokine profile of T cell and stromal cell origin. Arthritis Res Ther 2005;7:R784–95.

39. Smeets TJ, Barg EC, Kraan MC, Smith MD, Breedveld FC, Tak PP. Analysis of the cell infiltrate and expression of proinflammatory cytokines and matrix metalloproteinases in arthroscopic synovial biopsies: comparison with synovial samples from patients with end stage, destructive rheumatoid arthritis. Ann Rheum Dis 2003;62:635–8.

40. Humby F, Romão VC, Manzo A, Filer A, Bugatti S, Vieira-Sousa E, et al. A multicenter retrospective analysis evaluating performance of synovial biopsy techniques in patients with inflammatory arthritis: arthroscopic versus ultrasound-guided versus blind needle biopsy. Arthritis Rheumatol 2018;70:702–10.

41. Linn-Rasker SP, van der Helm-van Mil AH, Breedveld FC, Huizinga TW. Arthritis of the large joints—in particular, the knee—at first presentation is predictive for a high level of radiological destruction of the small joints in rheumatoid arthritis. Ann Rheum Dis 2007;66:646–50.

42. Rivellese F, Mauro D, Nerviani A, Pagani S, Fossati-Jimack L, Messemaker T, et al. Mast cells in early rheumatoid arthritis associate with disease severity and support B cell autoantibody production. Ann Rheum Dis 2018;77:1773–81.

43. Humby F, Kelly S, Hands R, Rocher V, DiCicco M, Ng N, et al. Use of ultrasound-guided small joint biopsy to evaluate the histopathologic response to rheumatoid arthritis therapy: recommendations for application to clinical trials. Arthritis Rheumatol 2015;67:2601–10.

44. Haringman JJ. Synovial tissue macrophages: a sensitive biomarker for response to treatment in patients with rheumatoid arthritis. Ann Rheum Dis 2005;64:834–8.

45. Orange DE, Agius P, DiCarlo EF, Robine N, Geiger H, Szymonifka J, et al. Identification of three rheumatoid arthritis disease subtypes by machine learning integration of synovial histologic features and RNA sequencing data. Arthritis Rheumatol 2018;70:690–701.

APPENDIX A: THE PEAC-R4RA INVESTIGATORS

The PEAC-R4RA investigators are as follows: Iain B. McInnes, Chris Buckley, Peter C. Taylor, Ernest Choy, Arthur Pratt, Christopher Edwards, Maya Buch, Nagui Gendi, Pauline Ho, Bhaskar Dasgupta, Patrick Durez, João Eurico Fonseca, Pier Paolo Sainaghi, Mattia Bellan, John Isaacs, Juan D. Cañete, Alberto Cauli, Mattia Congia, Piero Reynolds, Robert Moots, Nora Ng, Carlomaurizio Montecucco, and Patrick Verschueren.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Chondrocalcinosis of the Knee and the Risk of Osteoarthritis Progression: Data From the Knee and Hip Osteoarthritis Long-term Assessment Cohort

Augustin Latourte,[1] [iD] Anne-Christine Rat,[2] [iD] Willy Ngueyon Sime,[3] Hang-Korng Ea,[1] Thomas Bardin,[1] [iD] Bernard Mazières,[4] Christian Roux,[5] Francis Guillemin,[3] and Pascal Richette[1]

Objective. To assess the impact of knee chondrocalcinosis (CC) on the 5-year risk of joint replacement and disease progression in patients with knee osteoarthritis (OA).

Methods. Patients with symptomatic knee OA without previous total joint (knee or hip) replacement (TJR) were recruited from the Knee and Hip Osteoarthritis Long-term Assessment cohort. Cox proportional hazards regression and generalized estimating equation models were used to compare the time from inclusion or OA diagnosis to total knee replacement (TKR) or TJR between patients with and those without knee CC at inclusion. In patients without incident TKR, logistic regression was performed to examine the association between CC and radiographic progression (Kellgren/Lawrence [K/L] grade) or worsening of Western Ontario and McMaster Universities Arthritis Index (WOMAC) subscores for OA pain or function between years 0 and 5. Hazard ratios (HRs) and 95% confidence intervals (95% CIs) were estimated. Analyses were adjusted for age, sex, body mass index, WOMAC subscores, and K/L grade.

Results. Among the 656 patients included, 93 (14.2%) had knee CC, and 91 (13.9%) underwent TKR during the follow-up. Risk of TKR was not affected by the presence of knee CC (HR 1.26 [95% CI 0.74–2.17]). Similar results were obtained for the risk of incident TJR. For patients without incident TKR, knee CC did not affect the risk of worsening of K/L grade (odds ratio [OR] 0.9 [95% CI 0.4–1.7]), WOMAC pain subscore (OR 1.1 [95% CI 0.7–1.4]), or WOMAC function subscore (OR 0.9 [95% CI 0.4–2.0]).

Conclusion. In patients with symptomatic knee OA, the presence of knee CC did not affect the risk of arthroplasty or disease progression at 5 years.

INTRODUCTION

Calcium crystal deposition, including calcium pyrophosphate dehydrate (CPP) and basic calcium phosphate (BCP) crystals, is common in joints and frequently observed in cartilage and synovial fluid in osteoarthritis (OA) (1). The presence of these crystals is usually identified on radiographs as linear calcifications called chondrocalcinosis (CC), a condition strongly associated with OA, particularly in the knee (2,3). The role of calcium-containing crystals in the OA process has been largely debated, and it has been suggested that crystal-induced cartilage stress or synovial

inflammation contributes to cartilage damage (4). However, whether those crystals play a specific role in OA pathogenesis or simply result from aging or the progression of OA is unclear (5–8). Several cross-sectional studies have repeatedly demonstrated an association between knee CC and increased radiographic severity (9), osteophytes (3), or greater disease burden (10,11). However, a prospective magnetic resonance imaging (MRI) study showed that knee CC was not associated with greater cartilage damage over time (12).

Joint arthroplasty is the last-resort treatment for knee OA and is therefore proposed at the most advanced stages of the disease

(13). Few studies have investigated the association between knee CC and risk of joint replacement. One cross-sectional study showed that knee CC was not associated with younger age at the time of arthroplasty (14), and one retrospective study with a small sample size demonstrated a nonsignificant trend toward an increased risk of joint replacement in patients with knee OA and incident CC (15). A large study from a US veterans database suggested that the risk of arthroplasty was higher in patients with CPP deposition disease (CPPD) than in those without CPPD (16).

The aim of our study was to examine the relationship between knee CC and OA progression in a population-based cohort of patients with symptomatic knee OA, using total joint replacement (TJR) as the primary outcome measure. We also investigated other indicators of disease progression, such as radiographic outcome measures and clinical scores for pain or function, in order to provide a global overview of the impact of knee CC on knee OA prognosis.

PATIENTS AND METHODS

Study design and patients. The Knee and Hip Osteoarthritis Long-term Assessment (KHOALA) cohort is a French multicenter population-based cohort of 878 patients ages 40–75 years with symptomatic knee and/or hip OA according to the American College of Rheumatology criteria (17,18) and with a Kellgren/Lawrence (K/L) grade of ≥2 on radiographs (19,20). Patients were recruited using data from an OA national prevalence survey performed in France from 2007 to 2009 (21). Briefly, this survey included a random sample of households located within a 1-hour transportation distance of each investigating center in 6 regions of France, using random-digit phone dialing and the next-birthday method in each household. All 1,010 patients with hip or knee OA who were identified in this survey were invited to participate in the KHOALA study. Patients were not included if they had arthroplasty for the symptomatic joint, previous osteotomy, severe comorbidity leading to significant deterioration of quality of life, isolated patellofemoral OA, or other joint disease. The ethics committee Comité de Protection des Personnes Est III gave approval for the cohort study, which was registered at ClinicalTrials.gov (NCT00481338). For the purpose of this study, we included patients with symptomatic knee OA at baseline who had completed the 5-year follow-up.

Data collection. The detailed protocol for the KHOALA cohort has been previously described (19). Briefly, patients were followed up annually using self-reporting questionnaires. They underwent clinical examination and received knee and hip radiographs at baseline (year 0) and at years 3 and 5. Data collected included usual sociodemographic features, clinical data, and treatments received (including local treatments such as glucocorticoid or hyaluronic acid [HA] injections). Patient-reported outcome measures included pain score on a visual analog scale

(VAS; 0–100 mm) and pain, stiffness, and function subscores on the normalized Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC; 0–100 mm) (22). TJR (knee or hip) procedures were self-reported by participants, and data were cross-checked using radiographs or hospitalization data when feasible.

Knee radiography. All radiographs were obtained by routine analog or digital technology and were transferred to a centralized radiology center in Toulouse for reading and storage. Weight-bearing anteroposterior, posteroanterior semiflexed, and axial/sky views of both knees were obtained. All radiographs were read by 2 trained readers who were blinded with regard to clinical data and scored according to K/L grade (0 = no OA; 1 = doubtful OA; 2 = minimal OA; 3 = moderate OA; 4 = severe OA) (20). Radiographs were read independently by both observers, who trained together at the beginning of the study to adjust their K/L scoring, with a moderate level of interobserver agreement (weighted κ = 0.58) and good intraobserver reliability for each reader (weighted κ = 0.78 and 0.73). The training was repeated periodically during follow-up. CC was defined by the presence, in ≥1 knee, of calcium deposits within hyaline cartilage or fibrocartilage seen on knee radiographs (23).

Statistical analysis. Baseline characteristics of patients with and those without knee CC were compared by chi-square test or Fisher's exact test for categorical variables and by Student's t-test for quantitative variables in a preliminary analysis. The primary outcome measure was the time to total knee replacement (TKR). Secondary outcome measures were structural progression (K/L grade) and clinical worsening (WOMAC subscores) at 5 years.

With survival analyses, we estimated the association between knee CC and 1) the time from inclusion to the first TKR and 2) the time from inclusion to the first TJR. Because CPPD is likely a systemic condition (24), we hypothesized that the presence of CC in 1 knee might affect outcomes in other joints (i.e., contralateral knee or both hips), regardless of the presence or absence of CC on radiographs of those distant joints. We also examined the time from OA diagnosis to TKR or TJR: Kaplan-Meier survival curves were constructed for incidence of joint arthroplasty in the presence or absence of knee CC and were then compared by log rank test.

We calculated hazard ratios (HRs) and 95% confidence intervals (95% CIs). Data were censored at the date of last follow-up visit or death. Adjusted regression analysis with a Cox proportional hazards model included all tested variables (sex, age, body mass index [BMI], WOMAC subscores for pain and function, K/L grade, and previous joint replacement surgery [i.e., before inclusion]) with P values of <0.2 in the unadjusted model. Backward stepwise variable selection was used, with a significance level of 0.1 for entry in the model and 0.05 for staying

in the model. Since knee CC might also specifically impact the prognosis of the affected knee, we assessed at the knee level with a generalized estimating equation (GEE) model, which accounts for within-patient correlation (25).

We also examined the association between CC and the risk of radiographic progression (defined as an increase in K/L grade of ≥1 at 5-year follow-up versus baseline in ≥1 joint) and between CC and WOMAC pain or function subscore worsening (defined as an increase of at least the minimum clinically relevant change threshold of 17 [on 100-mm scale] for pain subscore and 12 [on 100-mm scale] for function subscore [26] at 5 years versus baseline). For these analyses, we retained only patients without TKR during the 5-year follow-up. We used adjusted logistic regression to calculate odds ratios (ORs) and 95% CIs.

To address potential confounding, we conducted sensitivity analyses by removing variables related to knee OA or CC disease severity (K/L grade, VAS pain score, WOMAC pain and function subscores) from the Cox regression and GEE models. All analyses were performed with SAS version 9.3, and 2-tailed P values less than 0.05 were considered significant.

RESULTS

Study sample. Of the 656 patients with symptomatic knee OA who were included (mean ± SD age 62.2 ± 8.5 years; 70.3% female), 93 (14.2%) had prevalent knee CC at baseline, and of these, 50 (53.8%) had bilateral knee CC. Patients with CC were older (mean ± SD age 64.3 ± 9.6 years versus 61.9 ± 8.2 years; P = 0.009), had longer disease duration (mean ± SD time between inclusion and first symptoms 16.4 ± 10.5 years versus 13.0 ± 7.6 years; P = 0.0002), and had a slightly lower BMI (mean ± SD 29.1 ± 5.3 kg/m² versus 30.5 ± 6.3 kg/m²; P = 0.047) (Table 1). Patients with and those without CC at baseline had similar pain levels and function impairment as assessed by VAS pain score and WOMAC pain or function subscales, and K/L grades were also similar. The presence or absence of CC did not affect the use of local treatments (i.e., glucocorticoids or HA injections) (Table 1).

Knee CC and incident TJR (primary outcome measure). During the 5-year follow-up, 105 patients (16.0%) and 91 patients (13.9%) underwent TJR and TKR, respectively. The risk of TKR at 5 years was not associated with the presence of knee CC at

Table 1. Baseline characteristics of OA patients in the study according to the presence or absence of knee chondrocalcinosis at baseline*

		Chondrocalcinosis at baseline	
	Total (n = 656)	No (n = 563)	Yes (n = 93)
Age, mean ± SD years	62.2 ± 8.5	61.9 ± 8.2	64.3 ± 9.6†
Female sex	461 (70.3)	397 (70.5)	64 (68.8)
BMI, mean ± SD kg/m²	30.3 ± 6.2	30.5 ± 6.3	29.1 ± 5.3‡
Associated hip osteoarthritis	49 (7.5)		
K/L grade			
2	288 (43.9)	249 (44.2)	39 (41.9)
3	201 (30.6)	169 (30.0)	32 (34.4)
4	167 (25.5)	145 (25.8)	22 (23.7)
VAS pain score (range 0–100 mm), mean ± SD	38.8 ± 25.4	38.7 ± 25.3	39.0 ± 25.8
WOMAC subscore (range 0–100 mm), mean ± SD			
Function subscore	34.2 ± 22.4	33.9 ± 22.1	36.4 ± 24.0
Pain subscore	33.2 ± 19.4	32.8 ± 19.0	35.3 ± 21.5
No. of painful joints§			
0	43 (6.6)	37 (6.6)	6 (6.5)
1	163 (24.8)	141 (25.0)	22 (23.7)
≥2	450 (68.6)	385 (68.4)	65 (69.9)
Previous local treatments			
Glucocorticoid injection	35 (5.3)	30 (5.3)	5 (5.4)
No. of injections, mean ± SD	1.5 ± 0.8	1.5 ± 0.8	1.6 ± 0.9
Hyaluronic acid injection	76 (11.6)	65 (11.5)	11 (11.8)
No. of injections, mean ± SD	3.2 ± 1.2	3.2 ± 1.1	3.5 ± 1.7
Previous joint replacement surgery	38 (5.8)	33 (5.9)	5 (5.4)
Time from inclusion to first symptoms, mean ± SD years	13.5 ± 8.2	13.0 ± 7.6	16.4 ± 10.5¶
Time from inclusion to OA diagnosis, mean ± SD years	8.2 ± 5.8	8.1 ± 5.7	9.0 ± 6.3
Bilateral chondrocalcinosis	–	–	50 (53.8)

* Except where indicated otherwise, values are the number (%) of patients. BMI = body mass index; K/L = Kellgren/Lawrence scale; VAS = visual analog scale; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index; OA = osteoarthritis.
† P = 0.0092 versus no chondrocalcinosis.
‡ P = 0.047 versus no chondrocalcinosis.
§ Among hip and knee joints (0–4).
¶ P = 0.0002 versus no chondrocalcinosis.

Table 2. Factors associated with 5-year risk of TKR: Cox proportional hazards regression model*

	TKR during follow-up		Unadjusted analysis		Adjusted analysis†	
	No (n = 565)	Yes (n = 91)	HR (95% CI)	P	HR (95% CI)	P
Age, mean ± SD years‡	61.9 ± 8.6	64.2 ± 7.4	1.39 (1.07–1.80)	0.01	–	–
Female sex	389 (61.9)	72 (79.1)	1.71 (1.03–2.84)	0.04	1.82 (1.07–3.08)	0.03
BMI, mean ± SD kg/m²‡	30.2 ± 6.3	31.0 ± 5.5	1.26 (0.91–1.73)	0.16	–	–
K/L grade				<0.0001	–	<0.0001
2	276 (48.8)	12 (13.2)	1.00 (referent)		1.00 (referent)	–
3	174 (30.8)	27 (29.7)	3.47 (1.76–6.84)		3.58 (1.76–7.27)	–
4	115 (20.4)	52 (57.1)	8.98 (4.79–16.83)		8.26 (4.23–16.11)	–
VAS pain score (range 0–100 mm), mean ± SD‡	36.3 ± 24.8	53.2 ± 23.8	1.29 (1.19–1.41)	<0.0001	1.19 (1.09–1.30)	<0.0001
WOMAC subscore (range 0–100 mm), mean ± SD‡						
Function subscore	32.8 ± 22.3	42.8 ± 21.3	1.21 (1.11–1.32)	<0.0001	–	–
Pain subscore	31.5 ± 18.9	43.5 ± 18.8	1.35 (1.22–1.50)	<0.0001	–	–
Knee chondrocalcinosis	77 (13.6)	16 (17.6)	1.26 (0.74–2.17)	0.39	–	–

* Except where indicated otherwise, values are the number (%) of patients. TKR = total knee replacement; 95% CI = 95% confidence interval (see Table 1 for other definitions).
† Factors with P values of <0.2 in the unadjusted model were entered into multivariable regression. Backward stepwise variable selection was used, with a significance level of 0.1 for entry in the model and 0.05 for staying in the model.
‡ For quantitative variables, the hazard ratio (HR) was calculated for a 10-point increase of the variable.

baseline (HR 1.26 [95% CI 0.74–2.17]; P = 0.39) (Table 2). Similarly, knee CC did not affect the time between OA diagnosis and the first TKR (HR 1.01 [95% CI 0.58–1.77]; P = 0.97) (Figure 1). In an adjusted analysis, female sex, increased baseline K/L grade, and VAS pain score were the only factors associated with increased risk of TKR. Finally, analyses performed at the knee level (GEE model) suggested that CC may protect against TKR, but the association remained nonsignificant (OR 0.41 [95% CI 0.14–1.23]; P = 0.11) (Table 3). Similar findings, with the exception of female sex as a

factor, were obtained for the risk of TJR in this sample (Supplementary Table 1 and Supplementary Figure 1, on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41186/abstract).

Knee CC and OA structural and clinical progression (secondary outcome measures). The presence of knee CC at baseline was not associated with increased risk of worsened K/L grade at 5 years (P = 0.71) (Table 4). In addition, baseline
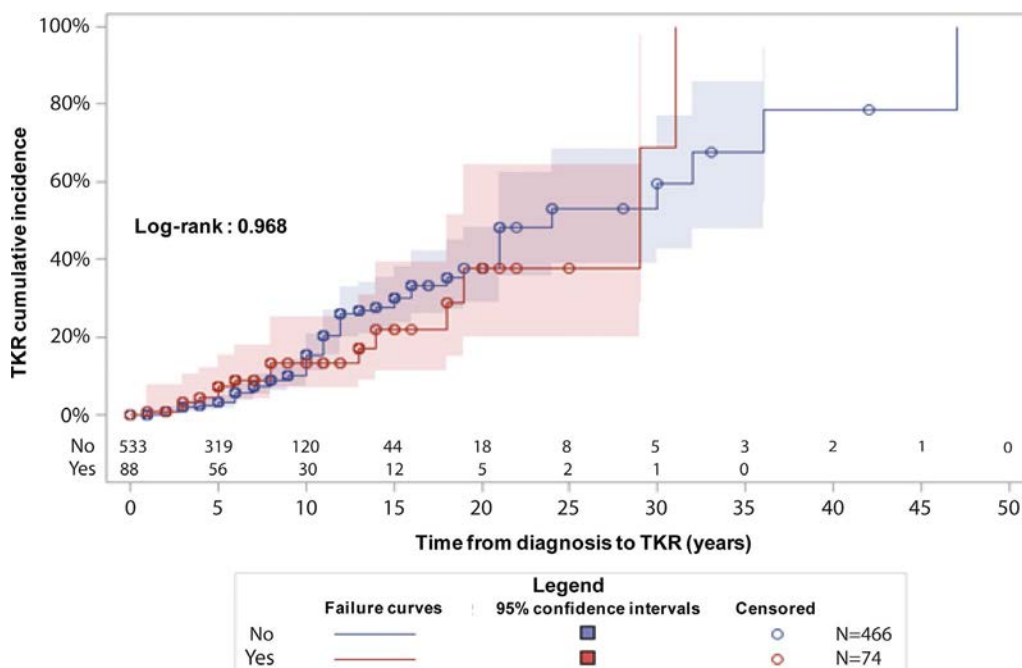
Figure 1. Risk of total knee replacement (TKR) according to the presence (Yes) or absence (No) of knee chondrocalcinosis at baseline in patients with symptomatic knee osteoarthritis.

Table 3. Factors associated with 5-year risk of TKR: generalized estimating equation model*

	OR (95% CI)	P
Age†	0.98 (0.94–1.03)	0.60
Female sex	2.06 (0.97–4.41)	0.06
BMI†	0.97 (0.92–1.01)	0.13
K/L grade		
2	1.00 (referent)	
3	5.01 (1.86–13.47)	0.001
4	14.01 (5.58–35.14)	<0.001
VAS pain score†	1.01 (0.99–1.03)	0.12
WOMAC subscore†		
Function subscore	0.98 (0.96–1.01)	0.12
Pain subscore	1.04 (1.00–1.07)	0.03
Knee chondrocalcinosis	0.41 (0.14–1.23)	0.11

* TKR = total knee replacement; 95% CI = 95% confidence interval (see Table 1 for other definitions).
† For quantitative variables, the odds ratio (OR) was calculated for a 10-point increase of the variable.

knee CC did not increase the risk of clinically relevant worsening of WOMAC function subscores (P = 0.78) or pain subscores (P = 0.89) at 5 years.

Sensitivity analyses. When variables related to OA or CC disease severity (K/L grade, VAS pain score, WOMAC pain and function subscores) were removed from the Cox regression model, we found no association between knee CC and the 5-year risk of TKR (P = 0.88), K/L grade worsening (P = 0.56), WOMAC function subscore worsening (P = 0.31), or WOMAC pain subscore worsening (P = 0.20) (Supplementary Table 2, http://onlinelibrary.wiley.com/doi/10.1002/art.41186/abstract). Additionally, knee CC did not affect the 5-year risk of TKR when the same variables were removed from the GEE model (HR 0.74 [95% CI 0.31–1.77]; P = 0.50).

DISCUSSION

CPPD is a common finding in knee OA, but its impact on the evolution of OA is poorly understood. Knee CC has been associated with synovitis in some cases, and therefore might facilitate joint destruction (27). In this population-based observational cohort of patients with symptomatic knee OA, we found a knee CC prevalence of 14.2% in a population with a mean age of 62 years, which is consistent with data reported in previous studies (3,11,12,28). We did not find an association between knee CC and the incidence of joint replacement surgery at 5 years. In addition, the presence of knee CC did not increase the risk of radiographic progression or worsening of WOMAC subscores for pain or function. These findings suggest that CPPD does not modify knee OA evolution.

In our study, patients with and those without knee CC had similar clinical and radiographic disease severity at baseline, although the disease duration of those with knee CC was significantly longer. This finding was heavily confounded by age, which is a major risk factor for CC (29). We also found a cross-sectional association between knee CC and reduced BMI, which is consistent with findings from a previous study (30). We have no clear explanation for this association, which deserves further research.

The presence of knee CC did not modify the risk of incident TKR in our cohort of knee OA patients. One previous study demonstrated that in patients undergoing TKR for knee OA (n = 102), age at the time of the surgery did not differ according to the presence or absence of CPPD crystals, and their preoperative disease burden was similar (14). Another study with a small sample size showed that among patients with knee OA, the need for surgery was higher in those with incident knee CC (n = 15, with 40% requiring surgery) than those without (n = 44, with 27.2% requiring surgery) but the difference was not significant (P = 0.27) (15). Conversely, a large cross-sectional study based on a US veterans database showed that knee arthroplasties were more frequent in patients with CPPD than in age-matched controls (n = 25,157 per group), even after controlling for the presence of OA (adjusted OR 1.64 [95% CI 1.53–1.76]) (16). This difference was not observed with hip or shoulder arthroplasties. Of note, the population of that study was highly specific because patients were mostly male (95%), and the site of OA was not specifically addressed. This precludes direct comparisons with our study.

CPPD is a systemic disease: knee CC is often associated with CC in other joints (24,31) but also with vascular or soft tissue calcification (30), which suggests a predisposition to calcium-containing crystal formation. Similarly, OA often affects multiple

Table 4. Association between baseline knee chondrocalcinosis and structural or clinical progression of knee OA at 5 years: logistic regression model*

Worsening score at 5 years		Chondrocalcinosis at baseline		OR (95% CI)
		No	Yes	
K/L grade	88/363 (24.2)	76/309 (24.6)	12/54 (22.2)	0.9 (0.4–1.7)
WOMAC function subscore (≥12 on 100-mm scale)	75/388 (19.3)	64/335 (19.1)	11/53 (20.8)	1.1 (0.7–1.4)
WOMAC pain subscore (≥17 on 100-mm scale)	68/393 (17.3)	59/339 (17.4)	9/54 (16.7)	0.9 (0.4–2.0)

* Values are the number/total tested (%) of patients. OR = odds ratio; 95% CI = 95% confidence interval (see Table 1 for other definitions).

joints and can be considered a systemic disease (32). Thus, we examined the risk of TJR (knee and/or hip arthroplasty) in patients with knee OA and did not observe any difference between those with and those without CC. However, the association between hip CC and hip OA is uncertain (24,33).

Several studies have shown an association between CPPD and radiographic features of OA. Some cross-sectional studies have suggested that CPPD was associated with a more severe radiographic OA phenotype (9,33). We found the presence of knee CC not to be associated with the severity of structural damage as assessed by the K/L scoring system at baseline or with worsening K/L grades at 5 years. This discrepancy could be explained by different definitions of knee CC. The diagnosis of knee CC was based on radiography, a standard imaging method that is easily performed but might be less revealing than ultrasonography or synovial fluid analysis, which were not available in our cohort (34). Furthermore, only knee and hip radiographs were available in the KHOALA cohort. Given the poor sensitivity of radiographs to detect hip CC (23), we chose to examine only knee radiographs to detect knee CC. Therefore, we might have missed some cases because we could not assess the presence of CC at distant joints, such as the wrists or symphysis pubis (28).

Previous studies have demonstrated a cross-sectional association between knee CC and knee pain or functional impairment (10,11). Those studies investigated older participants (≥65 years) regardless of their health status (Progetto Veneto Anziani cohort) or patients with or at risk for knee OA regardless of their symptoms (Osteoarthritis Initiative [OAI] cohort). Unlike those cohorts, which included patients with knee CC without knee OA or with asymptomatic knee OA, the KHOALA cohort included only patients with symptomatic OA. This difference might explain the lack of association between knee CC and pain or physical function in our study, both at baseline and over 5 years of follow-up.

The present study has several limitations. First, only a small proportion of patients experienced clinical or radiographic worsening of knee OA, and few underwent arthroplasty. This result is consistent with previous findings that show only a small proportion of knee OA patients with significant progression over time (35), which might affect the statistical power of our study. Second, if CC increases the risk of developing knee OA, then studying the effect of CC among patients who already have knee OA may introduce collider bias, a particular case of selection bias caused by the overlap between risk factors for incidence and for progression (36,37). In that case, the null association between CC and OA progression observed in our study may be an underestimation. This question should be addressed by future research.

The severity of OA was assessed by K/L grade alone and not by evaluating osteophytes, which have been found to be associated with knee CC (3). As discussed above, the use of radiography as the modality to detect knee CC might have underestimated its prevalence (34). The absence of double reading of radiographs is also a limitation of the present study. In the same way, radiography

might not be the best imaging method to address OA structural progression. For instance, MRI is widely used to more precisely examine cartilage loss over time in OA (38). In the OAI cohort, a cross-sectional association between MRI-detected calcium-containing crystals and greater cartilage defects was reported (39). However, data from longitudinal studies conversely showed knee CC not to be associated with cartilage loss over time (12), which is consistent with our findings. This representative real-life observational cohort was assessed using simple and standard tools close to what is typically done in routine care for patients with OA. This represents a strength of the KHOALA study, but we acknowledge that further studies are warranted to better understand the relationship between calcium-containing crystals and OA. For instance, current routine imaging procedures are unable to discriminate between CPPD and BCP crystals, and their role in cartilage damage in OA might significantly differ (1). Finally, the role of calcium-containing crystals in other joints is unknown.

In conclusion, we confirmed that knee CC is common in patients with symptomatic knee OA, associated with both age and disease duration but not disease severity at baseline. The risk of arthroplasty at 5 years was similar for patients with knee OA and CC and those with OA alone, and the progression pattern was similar in terms of pain, function, and radiographic severity. Overall, these findings suggest that knee CC is not a relevant prognostic factor in symptomatic knee OA.

AUTHOR CONTRIBUTIONS

REFERENCES

1. McCarthy GM, Dunne A. Calcium crystal deposition diseases—beyond gout [review]. Nat Rev Rheumatol 2018;14:592–602.

2. Sanmarti R, Kanterewicz E, Pladevall M, Pañella D, Tarradellas JB, Gomez JM. Analysis of the association between chondrocalcinosis and osteoarthritis: a community based study. Ann Rheum Dis 1996;55:30–3.

3. Neame RL, Carr AJ, Muir K, Doherty M. UK community prevalence of knee chondrocalcinosis: evidence that correlation with osteoarthritis is through a shared association with osteophyte. Ann Rheum Dis 2003;62:513–8.

4. Lioté F, Ea HK. Clinical implications of pathogenic calcium crystals. Curr Opin Rheumatol 2014;26:192–6.

5. Fuerst M, Bertrand J, Lammers L, Dreier R, Echtermeyer F, Nitschke Y, et al. Calcification of articular cartilage in human osteoarthritis. Arthritis Rheum 2009;60:2694–703.

6. Hawellek T, Hubert J, Hischke S, Krause M, Bertrand J, Pap T, et al. Articular cartilage calcification of the hip and knee is highly

prevalent, independent of age but associated with histological osteoarthritis: evidence for a systemic disorder. Osteoarthritis Cartilage 2016;24:2092–9.

7. Mitsuyama H, Healey RM, Terkeltaub RA, Coutts RD, Amiel D. Calcification of human articular knee cartilage is primarily an effect of aging rather than osteoarthritis. Osteoarthritis Cartilage 2007;15:559–65.

8. Nalbant S, Martinez JA, Kitumnuaypong T, Clayburne G, Sieck M, Schumacher HR Jr. Synovial fluid features and their relations to osteoarthritis severity: new findings from sequential studies. Osteoarthritis Cartilage 2003;11:50–4.

9. Ledingham J, Regan M, Jones A, Doherty M. Radiographic patterns and associations of osteoarthritis of the knee in patients referred to hospital. Ann Rheum Dis 1993;52:520–6.

10. Musacchio E, Ramonda R, Perissinotto E, Sartori L, Hirsch R, Punzi L, et al. The impact of knee and hip chondrocalcinosis on disability in older people: the ProVA Study from northeastern Italy. Ann Rheum Dis 2011;70:1937–43.

11. Han BK, Kim W, Niu J, Basnyat S, Barshay V, Gaughan JP, et al. Association of chondrocalcinosis in knee joints with pain and synovitis: data from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken) 2017;69:1651–8.

12. Neogi T, Nevitt M, Niu J, LaValley MP, Hunter DJ, Terkeltaub R, et al. Lack of association between chondrocalcinosis and increased risk of cartilage loss in knees with osteoarthritis: results of two prospective longitudinal magnetic resonance imaging studies. Arthritis Rheum 2006;54:1822–8.

13. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008;16:137–62.

14. Viriyavejkul P, Wilairatana V, Tanavalee A, Jaovisidha K. Comparison of characteristics of patients with and without calcium pyrophosphate dihydrate crystal deposition disease who underwent total knee replacement surgery for osteoarthritis. Osteoarthritis Cartilage 2007;15:232–5.

15. Reuge L, Van Linthoudt D, Gerster JC. Local deposition of calcium pyrophosphate crystals in evolution of knee osteoarthritis. Clin Rheumatol 2001;20:428–31.

16. Kleiber Balderrama C, Rosenthal AK, Lans D, Singh JA, Bartels CM. Calcium pyrophosphate deposition disease and associated medical comorbidities: a national cross-sectional study of US veterans. Arthritis Care Res (Hoboken) 2017;69:1400–6.

17. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis Rheum 1986;29:1039–49.

18. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum 1991;34:505–14.

19. Guillemin F, Rat AC, Roux CH, Fautrel B, Mazieres B, Chevalier X, et al. The KHOALA cohort of knee and hip osteoarthritis in France. Joint Bone Spine 2012;79:597–603.

20. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthrosis. Ann Rheum Dis 1957;16:494–502.

21. Guillemin F, Rat AC, Mazieres B, Pouchot J, Fautrel B, Euller-Ziegler L, et al. Prevalence of symptomatic hip and knee osteoarthritis: a two-phase population-based survey. Osteoarthritis Cartilage 2011;19: 1314–22.

22. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.

23. Zhang W, Doherty M, Bardin T, Barskova V, Guerne PA, Jansen TL, et al. European League Against Rheumatism recommendations for calcium pyrophosphate deposition. Part I: terminology and diagnosis. Ann Rheum Dis 2011;70:563–70.

24. Abhishek A, Doherty S, Maciewicz R, Muir K, Zhang W, Doherty M. Evidence of a systemic predisposition to chondrocalcinosis and association between chondrocalcinosis and osteoarthritis at distant joints: a cross-sectional study. Arthritis Care Res (Hoboken) 2013;65: 1052–8.

25. Zhang Y, Glynn RJ, Felson DT. Musculoskeletal disease research: should we analyze the joint or the person? J Rheumatol 1996;23:1130–4.

26. Tubach F, Ravaud P, Martin-Mola E, Awada H, Bellamy N, Bombardier C, et al. Minimum clinically important improvement and patient acceptable symptom state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic back pain, hand osteoarthritis, and hip and knee osteoarthritis: results from a prospective multinational study. Arthritis Care Res (Hoboken) 2012;64:1699–707.

27. Schlesinger N, Hassett AL, Neustadter L, Schumacher HR Jr. Does acute synovitis (pseudogout) occur in patients with chronic pyrophosphate arthropathy (pseudo-osteoarthritis)? Clin Exp Rheumatol 2009;27:940–4.

28. Abhishek A, Doherty S, Maciewicz R, Muir K, Zhang W, Doherty M. Chondrocalcinosis is common in the absence of knee involvement. Arthritis Res Ther 2012;14:R205.

29. Abhishek A. Calcium pyrophosphate deposition disease: a review of epidemiologic findings. Curr Opin Rheumatol 2016;28:133–9.

30. Abhishek A, Doherty S, Maciewicz R, Muir K, Zhang W, Doherty M. Association between low cortical bone mineral density, soft-tissue calcification, vascular calcification and chondrocalcinosis: a case-control study. Ann Rheum Dis 2014;73:1997–2002.

31. Filippou G, Filippucci E, Tardella M, Bertoldi I, Di Carlo M, Adinolfi A, et al. Extent and distribution of CPP deposits in patients affected by calcium pyrophosphate dihydrate deposition disease: an ultrasonographic study. Ann Rheum Dis 2013;72:1836–9.

32. Prieto-Alhambra D, Judge A, Javaid MK, Cooper C, Diez-Perez A, Arden NK. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints. Ann Rheum Dis 2014;73:1659–64.

33. Abhishek A, Doherty S, Maciewicz RA, Muir K, Zhang W, Doherty M. Does chondrocalcinosis associate with a distinct radiographic phenotype of osteoarthritis in knees and hips? A case–control study. Arthritis Care Res (Hoboken) 2016;68:211–6.

34. Filippou G, Adinolfi A, Cimmino MA, Scirè CA, Carta S, Lorenzini S, et al. Diagnostic accuracy of ultrasound, conventional radiography and synovial fluid analysis in the diagnosis of calcium pyrophosphate dihydrate crystal deposition disease. Clin Exp Rheumatol 2016;34:254–60.

35. Felson D, Niu J, Sack B, Aliabadi P, McCullough C, Nevitt MC. Progression of osteoarthritis as a state of inertia. Ann Rheum Dis 2013;72:924–9.

36. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. Int J Epidemiol 2010;39:417–20.

37. Dahabreh IJ, Kent DM. Index event bias as an explanation for the paradoxes of recurrence risk research. JAMA 2011;305:822–3.

38. MacKay JW, Low SB, Smith TO, Toms AP, McCaskie AW, Gilbert FJ. Systematic review and meta-analysis of the reliability and discriminative validity of cartilage compositional MRI in knee osteoarthritis. Osteoarthritis Cartilage 2018;26:1140–52.

39. Gersing AS, Schwaiger BJ, Heilmeier U, Joseph GB, Facchetti L, Kretzschmar M, et al. Evaluation of chondrocalcinosis and associated knee joint degeneration using MR imaging: data from the Osteoarthritis Initiative. Eur Radiol 2017;27:2497–506.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Effect of Therapy on Radiographic Progression in Axial Spondyloarthritis: A Systematic Review and Meta-Analysis

Paras Karmacharya,[1] (iD) Ali Duarte-Garcia,[1] (iD) Maureen Dubreuil,[2] (iD) M. Hassan Murad,[1] Ravi Shahukhal,[3] Pragya Shrestha,[1] Elena Myasoedova,[1] (iD) Cynthia S. Crowson,[1] (iD) Kerry Wright,[1] and John M. Davis III[1] (iD)

Objective. To investigate the effect of therapies on radiographic progression in patients with axial spondyloarthritis (SpA).

Methods. A comprehensive database search for studies assessing radiographic progression in axial SpA (particular treatment versus no treatment of interest) was performed. Study-specific standardized mean differences in treatment outcomes at 2 and ≥4 years were estimated and combined using random-effects models.

Results. Twenty-four studies in patients with axial SpA were identified, of which 18 involved tumor necrosis factor inhibitors (TNFi), 8 involved nonsteroidal antiinflammatory drugs (NSAIDs), and 1 involved secukinumab. Spinal radiographic progression, as measured by the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS), was not significantly different between TNFi-treated and biologics-naive patients at 2 years (mSASSS difference −0.73 [95% confidence interval (95% CI) −1.52, 0.12], I^2 = 28%) and ≥4 years (mSASSS difference −2.03 [95% CI −4.63, 0.72], I^2 = 63%). Sensitivity analyses restricted to studies with a low risk of bias showed a significant difference in spinal radiographic progression between TNFi-treated and biologics-naive patients at ≥4 years (mSASSS difference −2.17 [95% CI −4.19, −0.15]). No significant difference in spinal radiographic progression was observed between NSAID-treated and control patients (mSASSS difference −0.30 [95% CI −2.62, 1.31], I^2 = 71%) or between secukinumab-treated and biologics-naive patients (mSASSS difference −0.34 [95% CI −0.85, 0.17]). With regard to treatment differences in patients with nonradiographic axial SpA or in patients with radiographic progression measured using the sacroiliac joint score, an insufficient number of studies were available for analysis.

Conclusion. Although no significant protective effect of TNFi treatment on spinal radiographic progression was seen over the course of 2 years or ≥4 years in patients with axial SpA, our analysis restricted to studies with a low risk of bias showed a protective effect of TNFi after ≥4 years. Therefore, long-term TNFi exposure might confer beneficial effects on spinal radiographic progression in axial SpA. No difference in radiographic progression at 2 years was seen in either the NSAID or secukinumab treatment groups compared to their controls. Future studies should explore the effects of biologic treatment on radiographic progression, as well as the effects of long-term biologics exposure, in patients with early axial SpA or those with nonradiographic axial SpA.

INTRODUCTION

Treatment goals for axial spondyloarthritis (SpA), including ankylosing spondylitis (AS) and nonradiographic axial SpA, are

to achieve control of spinal and peripheral symptoms as well as to prevent radiographic damage (1). Radiographic progression of spinal disease takes place in 20–45% of patients with AS after 2 years of follow-up, and progression is expected to continue in

successive years, leading to complete fusion of the spine ("bamboo spine") in up to 40% of patients (2,3). The progression rate is slower in patients with nonradiographic axial SpA, with a progression rate of ~7% observed in this subset of patients within 2 years (3).

Conventional radiography is the gold standard for assessment of radiographic progression (4), although magnetic resonance imaging can detect active inflammation before the appearance of changes associated with chronic spinal damage on radiographs (5). The Assessment of SpondyloArthritis international Society and Outcome Measures in Rheumatology Working Groups recommend use of the modified Stoke AS Spine Score (mSASSS) for quantitative assessment of axial damage (6,7). The scores on the mSASSS range from 0 to 72, signifying the extent of structural changes as detected on lateral radiographs of the cervical and lumbar spine (8). This scoring system has the highest reliability and strongest validation when compared to the Bath AS Radiology Index (BASRI) of the spine and sacroiliac (SI) joint radiographic progression scores (6,7,9). The minimum interval for assessing significant radiographic change is 2 years (10,11), and a change of 2 mSASSS units in 2 years or at least 1 new syndesmophyte formation in 2 years is considered to be radiographic progression (2,12).

Although radiographic progression is an important predictor of poor functional outcomes in axial SpA (13,14), treatment at present is largely targeted to symptoms. If any therapy is shown to slow the natural progression of the disease, it might support the early introduction of therapy with a treat-to-target strategy, similar to what has been implemented in patients with rheumatoid arthritis. The current guidelines recommend against this, however, because supporting evidence is lacking (1).

Considering the clinical implications of radiographic progression of spinal disease and the conflicting study results, we aimed to perform a systematic review and meta-analysis of the effect of different therapies on radiographic progression in patients with axial SpA.

METHODS

Search strategy and study selection. A comprehensive search of several databases from inception to January 16, 2019 was conducted. The databases included Ovid Medline In-Process & Other Non-Indexed Citations, Ovid MEDLINE, Ovid EMBASE, Ovid Cochrane Central Register of Controlled Trials, Ovid Cochrane Database of Systematic Reviews, and Scopus. American College of Rheumatology and European League Against Rheumatism abstracts indexed in MEDLINE were included as well. The search strategy was designed and conducted by a medical reference librarian with input from the principal investigator (see details on the search strategy in Supplementary File 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41206/abstract).

We included all original reports in which specific inclusion criteria were met, including the following requirements: 1) eligible participants were adult patients (age ≥18 years) with axial SpA, including AS and nonradiographic axial SpA; 2) participants were assessed for radiographic progression of axial disease (outcome) using a quantitative scoring method (mean mSASSS score, number of syndesmophytes, SI joint radiographic progression score, or any other scoring method); 3) radiographic scores from each patient were reported with respect to a particular therapy, and were compared to those from a placebo group or a group without the therapy of interest; and 4) the duration of follow-up was at least 1 year. While cohort and case–control studies were included, cross-sectional studies, case series, case reports, and studies in animals were excluded. The effect of combination therapy was beyond the scope of this study.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement for reporting systematic reviews (15) was followed (Figure 1). Two authors (PK and AD) screened abstracts for eligibility, retrieved full texts, and excluded irrelevant articles. The interrater agreement was very good (κ = 0.91, 95% confidence interval [95% CI] 0.81, 0.99). Disagreements between raters were resolved by discussion about eligibility. Bibliographies belonging to included studies, reviews, and relevant articles were screened for additional studies. No language restrictions were made. Language translation was done by translators (AD and EM) who were proficient both in the particular language and in English.

Data extraction/handling methods. Relevant data were extracted by one of the authors (PK) and then checked by a second author (RS). Data from available full texts were considered in preference to conference abstracts, unless the abstracts contained updated results. Where a matched data analysis was provided (for example, in studies with comparison to a historical control cohort), this was extracted with preference to data on all participants. Data on radiographic progression were collected at 2 years and ≥4 years of follow-up. Authors were contacted for additional data where necessary.

Outcome measures. The main outcome measure was the difference in mSASSS scores between the treatment groups of patients with AS at 2 years and at ≥4 years. Other radiographic scores were used as secondary outcome measures, including the mSASSS score in patients with nonradiographic axial SpA, change in the number of syndesmophytes, BASRI spine score, computed tomography (CT) score of the facet joints, and SI joint radiographic progression score in patients with AS and/or nonradiographic axial SpA. We did not pool the mSASSS scores with the SI joint radiographic progression scores or with other scores in our analysis, as pooling different scoring systems has not been validated. For example, the SI joint score is semiquantitative, and has a relatively poor correlation with functional status (16) and with interreader reliability (17).
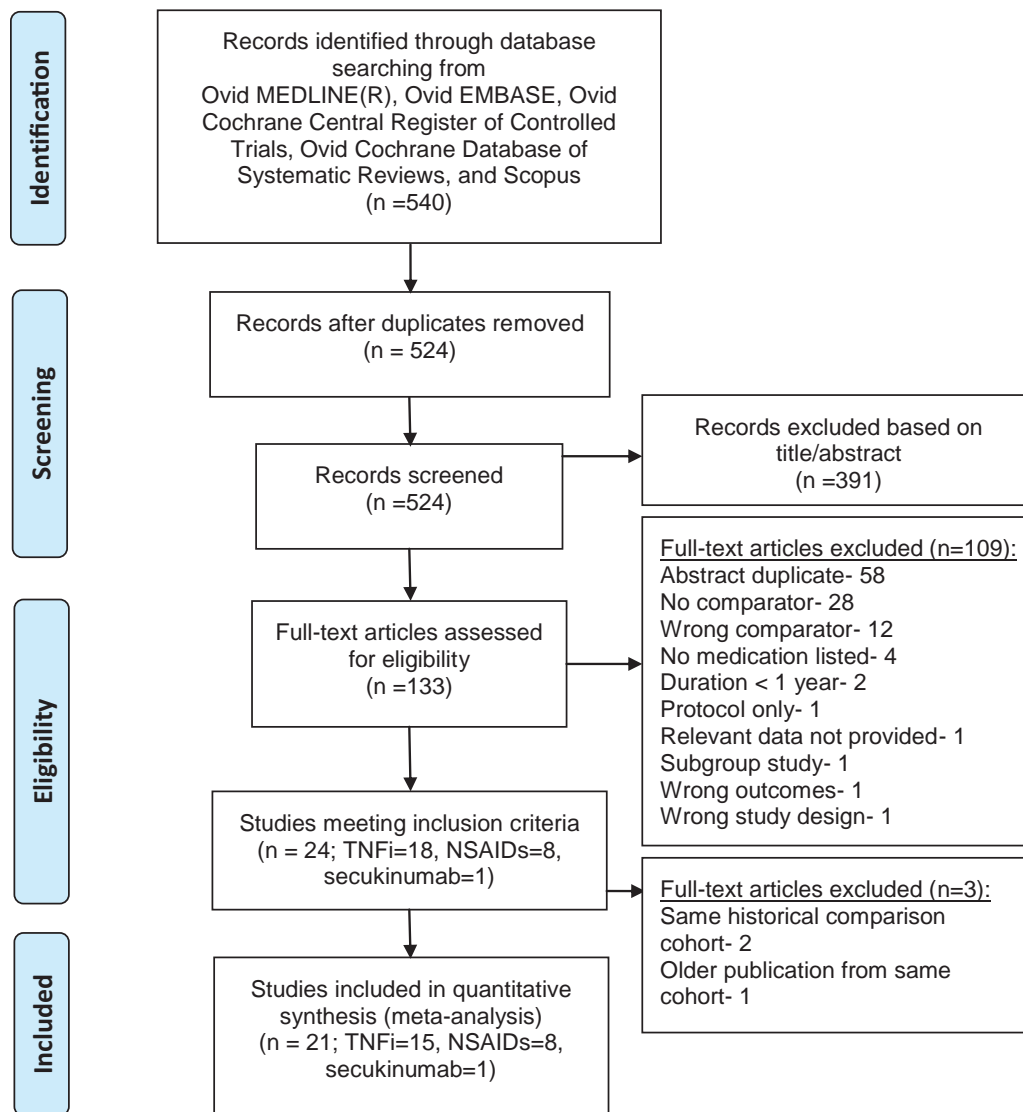
Figure 1. Flow chart describing the systematic search and study selection process. TNFi = tumor necrosis factor inhibitors; NSAIDs = nonsteroidal antiinflammatory drugs.

The effect size was calculated as the standard mean difference (SMD) of the change in radiographic score from baseline between the treatment and control groups.

The SMD was estimated as follows: SMD = (radiographic score$_{T,baseline}$ – radiographic score$_{T,follow-up}$) – (radiographic score$_{C,baseline}$ – radiographic score$_{C,follow-up}$) divided by SD$_{pooled}$, where SD$_{pooled}$ is the pooled standard difference (SD) values for radiographic change in the treatment (T) and control (C) groups both at baseline and at follow-up.

Statistical analysis. All outcome comparisons and treatment effects were calculated using RevMan, version 5.3. Binary measures were converted to the SMD ± SE (18). Study-specific SMDs were estimated and combined using the random-effects model described by DerSimonian and Laird (19). Rescaling to original units (e.g., mSASSS and number of syndesmophytes)

was done by multiplying the SMD by the SD to allow better clinical interpretation. The SD was obtained from the average of the pooled SD values for change in radiographic scores (in the treatment and control groups) from several trials reporting the original scale (20).

Radiographic outcomes in AS and nonradiographic axial SpA were reported separately and not pooled. To maintain independence, only 2 of the 4 studies with the Outcome Assessment in Ankylosing Spondylitis (OASIS) cohort as the control group (with the largest number of patients) were included in 2 separate forest plots (1 study with 2 years of follow-up [21] and 1 study with ≥4 years of follow-up [22]).

Sensitivity analyses were carried out with other studies (23,24), one at a time. Similarly, 2 studies from the Prospective Study Of AS (PSOAS) cohort were included only in separate analyses (25,26). Data from the study by Gensler et al (26) were used in pref-

erence to older data from the PSOAS cohort in which change in the mSASSS was defined as an mSASSS of ≥1 unit/year, and variable follow-up periods of 1.5–9 years were used (27). Sample mean ± SD values were calculated based on the sample size, and median values (with range and/or interquartile range) were also reported (28).

Assessment of risk of bias, certainty in the evidence, and heterogeneity. Studies were independently evaluated by 2 of the authors (PK and RS) for risk of bias (29). The Newcastle-Ottawa scale (30) was used for cohort studies, and the revised Cochrane risk-of-bias tool (RoB 2.0) was used for randomized trials (31) (see details in Supplementary File 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41206/abstract). The certainty in the evidence was evaluated using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) method (32). Publication bias was assessed visually using funnel plots (see Supplementary File 3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41206/abstract). Between-study heterogeneity was assessed using I^2 statistics (i.e., I^2 of <30% = low heterogeneity, 30–60% = moderate heterogeneity, and >60% = high heterogeneity) (33).

RESULTS

Characteristics of the included studies. Of the 524 studies screened, 24 of them (23 in English, 1 in Russian) fulfilled our inclusion criteria: 18 involving treatment with tumor necrosis factor inhibitors (TNFi) (n = 4,874 patients), 8 involving nonsteroidal antiinflammatory drugs (NSAIDs) (n = 2,321 patients), and 1 involving secukinumab (n = 237 patients). Among these studies, 3 contained data for both NSAIDs and TNFi (25,26,34). The included studies were mostly cohort studies and open-label extensions of randomized controlled trials (RCTs). Three of the TNFi studies were excluded from the meta-analysis. One RCT for TNFi (35) and 2 RCTs for NSAIDs (36,37) were noted (Table 1). Most studies were judged to have a low risk of bias, except for 2 observational studies (38,39) and 2 RCTs (36,37) (see Supplementary File 2 [http://onlinelibrary.wiley.com/doi/10.1002/art.41206/abstract]). One study on NSAIDs (40) was excluded, because no quantitative data were available and phenylbutazone (which is no longer approved for human use in the US) was studied.

Among the TNFi studies, there were 17 studies in patients with AS and 1 in patients with nonradiographic axial SpA (41). Six studies used a historical cohort as a comparator, including 4 studies using the OASIS cohort (comprising subjects not receiving TNFi) (21,32–34), 1 study using the German Spondyloarthritis Inception Cohort (GESPIC) (42), and 1 study using the Herne cohort (43). Two studies used contemporary cohorts (comprising subjects not receiving TNFi) as comparators (41,44). Control groups were subjects receiving continuous NSAIDs (38), subjects receiving any NSAIDs (44), or subjects receiving standard of care with no TNFi.

The only RCT for TNFi used a placebo group as the control. These control subjects were followed up until 24 weeks, after which they crossed over to receive golimumab 50 mg (same as the intervention arm) through 100 weeks (35). The type of TNFi used was variable: 3 studies used infliximab (22,24,42), 2 studies used etanercept (23,41), 1 study used adalimumab (21), 1 study used golimumab (35), and 2 studies used etanercept, infliximab, and adalimumab (45,46), while in the remainder of the 7 studies, type of TNFi was not specified. Only 1 of the studies compared 2 doses of TNFi (golimumab), and none of the other studies compared the TNFi dose response in relation to radiographic progression (35).

The maximum duration of follow-up was 2 years in 7 of the TNFi studies (21,33–35,39–41) and ≥4 years in 9 studies (23,29,32,36–38,43–45), while 1 study had a median follow-up of 3 years (>2 years up to >10 years) (25). The mean duration of disease was >5 years in most of the TNFi studies, except for 3 of them (41,44,47). The baseline mSASSS varied between studies and also within study groups, ranging from 4 to 18.87 (median 13.20) in the TNFi group and from 3.70 to 19 (median 14.20) in the control group.

Among the 8 NSAID studies, 6 studies included only patients with AS (25,34,36,37,49,50), 1 study evaluated patients with early axial SpA (39), and 1 study included both patients with AS and patients with nonradiographic axial SpA (48). Six studies used scores based on the mSASSS (26,34,36,37,39,48), 1 study used BASRI spine scores (49), and 1 study used the baseline SI joint score (25).

The maximum duration of follow-up was 2 years in 4 of the NSAID studies (36,37,39,48), and >2 years in the remaining 4 studies (25,26,34,49). Among patients with AS, the baseline mSASSS ranged from 6.60 to 14.20 (median 7.90) in the NSAID group, and from 5.70 to 14.20 (median 11.65) in the control group. In 1 of the studies in patients with nonradiographic axial SpA, the mean ± SD baseline mSASSS was 1.60 ± 4, compared to 2.60 ± 4.80 in the control group (48).

The only included study evaluating secukinumab (an abstract) was a retrospective analysis of an RCT, comparing 2-year data to the historical biologics-naive cohort from the Effects of NSAIDs on Radiographic Damage in AS (ENRADAS) study (50). The mSASSS was comparable between the 2 groups, with a mean ± SD mSASSS of 9.55 ± 14.14 in the secukinumab group and 9.95 ± 13.76 in the control group. The full report was published beyond the inclusion period for our review; nevertheless, data from the abstract were verified against the final published version (June 2019) (51). The original trial and 4-year data were not included in our study, because although 2 different doses of secukinumab were compared, the study did not have an NSAID or placebo arm (52,53).

Radiographic outcomes. *TNFi.* In the assessment of spinal radiographic progression following treatment with TNFi in patients with AS, most studies used the mSASSS as the measure of radiographic progression (15 studies), whereas 1 study reported CT

Table 1. Baseline characteristics of the included studies*

Treatment, study (ref.)	Study design	Duration of follow-up, years	Disease characteristics of cohort	Drug and dosage	Treatment group			Control group			
					No. of patients	Disease duration, mean ± SD years	Baseline mSASSS, mean ± SD	Drug	No. of patients	Disease duration, mean ± SD years	Baseline mSASSS. mean ± SD
TNFi											
Baraliakos 2005 (42)	Open-label extension of RCT	2	Active AS	Infliximab 5 mg/kg IV every 6 weeks	41	15.5 ± 7.5	12.1 ± 16.9	No TNFi, GESPIC historical cohort	41	5.5 ± 2.25	5.9 ± 13.4
Baraliakos 2007 (22)	Open-label extension of RCT	4	Active AS (same cohort as 2005 study)	Infliximab 5 mg/kg IV every 6 weeks	33	19.0 ± 23.4	11.6 ± 15.3	No TNFi, OASIS historical cohort	132	11.7 ± 9.3	12.7 ± 17.4
van der Heijde 2008 (24)	Open-label extension of 24-week RCT	2 (96 weeks)	Active AS (ASSERT cohort)	Infliximab 5 mg/kg IV every 6 weeks after loading dose	201	10.2 ± 8.7	17.7 ± 17.9	No TNFi, OASIS historical cohort–matched	70	9.9 ± 8.8	17.5 ± 19.1
van der Heijde 2008 (23)	Open-label extension of 24-week RCT	2 (96 weeks)	AS	Etanercept 25 mg SC twice weekly	257	10 ± 8.5	14 ± 17.6	No TNFi, OASIS historical cohort (meeting RCT entry criteria)	76	12 ± 9.8	19 ± 20.8
van der Heijde 2009 (21)	Open-label extension of 24-week RCT	2	AS (Canadian [M03-606] study and the ATLAS Study Group)	Adalimumab 40 mg SC every other week	307	19.8 ± 19.3	19.8 ± 19.3	No TNFi, OASIS historical cohort (eligible patients)	77	11.3 ± 8.7	15.8 ± 17.6
Pedersen 2011 (45)	Cohort study	2	AS	Infliximab 3 or 5 mg/kg (n = 11), etanercept 25 mg twice weekly (n = 10), adalimumab 40 mg every other week (n = 2)	23	18.2 ± 11.4	14.5 ± 16.1	No TNFi, standard therapy	27	15 ± 10	10.0 ± 12.1
Haroon 2013 (27)	Cohort, prospective	1.5–9	AS	TNFi (type and dose not specified)	201	–	–	No TNFi	133	–	–
Kang 2013 (46)	Cohort study	2	AS	Infliximab, etanercept, or adalimumab	26	9.5 ± 5.1	4.0 ± 6.6	No TNFi (NSAID and/or MTX or SSZ)	37	8.0 ± 4.5	3.7 ± 6.8
Min 2014 (38)	Cohort, retrospective, single-center	8	AS	TNFi (type and dose not specified)	14	–	–	Continuous NSAIDs	12	–	–

(Continued)

Table 1. *(Cont'd)*

Treatment, study (ref.)	Study design	Duration of follow-up, years	Disease characteristics of cohort	Drug and dosage	Treatment group No. of patients	Disease duration, mean ± SD years	Baseline mSASSS, mean ± SD	Control group Drug	No. of patients	Disease duration, mean ± SD years	Baseline mSASSS. mean ± SD
Braun 2014 (35)	Phase III, multicenter, randomized, placebo-controlled, double-blind, placebo crossover	2 (104 weeks), 4	Active AS (GO-RAISE trial)	Golimumab 50 or 100 mg every 4 weeks	233	7.25 ± 35.59	12.64 ± 17.71	Placebo, crossover to golimumab 50 mg at week 16 or 24	66	5.20 ± 45.94	16.1 ± 18.7
Baraliakos 2014 (43)	Open-label extension of RCT	8	AS (DIKAS)	Infliximab 5 mg/kg IV every 6 weeks	22	15.8 ± 8.5	13.2 ± 17.6 (adjusted to 13.8 for comparison)	No TNFi, Herne historical cohort	34	20.7 ± 5.7	14.2 ± 13.8 (adjusted to 13.8 for comparison)
Chung 2015 (47)	Clinical trial	4 (42–66 months)	AS	TNFi (type and dose not specified)	25	4.15 ± 4.02	–	No TNFi	25	2.13 ± 1.73	–
Minhas 2016 (25)	Cohort, prospective	≥2, up to >10 (median 3)	AS (PSOAS cohort)	TNFi (>50% of follow-up) (type and dose not specified)	630 (total)	–	–	No TNFi (<50% of follow-up period)	630 (total)	–	–
Kim 2016 (55)	Cohort, prospective	5	AS (OSKAR cohort)	TNFi (type and dose not specified)	269	11.33 ± 7.51	18.87 ± 17.96	TNFi naive	341	8.04 ± 6.57	15.68 ± 15.49
Dougados 2018 (41)	Open-label extension of RCT	2	Nonradiographic axial SpA (EMBARK trial)	Etanercept 25 mg SC twice weekly	162	2.4 ± 1.8	Baseline total SI joint score 1.5 ± 1.2	No biologics, contemporary DESIR cohort	193	1.70 ± 1.0	Baseline total SI joint score 1.9 ± 1.6
Molnar 2018 (34)	Cohort study	10 (2-year radiographic interval progression)	AS (patients fulfilling modified New York criteria for AS from SCQM axial SpA cohort)	Any TNFi before radiographic interval, NSAIDs	163	13.8 ± 9.7	6.6 ± 12.5	No TNFi before radiographic interval	269	–	–
Gensler 2018 (26)	Cohort, prospective	2 and 4	AS (PSOAS cohort)	TNFi (type and dose not specified)	239	16.8 ± 12.5	14.2 ± 19.6	No TNFi	280	16.8 ± 12.5	14.2 ± 19.6
Park 2019 (44)	Cohort study, single center	4	Early AS (symptom duration <10 years)	TNFi (type and dose not specified)	135	2.7 ± 2.6	6.2 ± 9.9	NSAIDs, control group from different institution but same location (Seoul)	80	0.7 ± 1.8	7.3 ± 10.8

(Continued)

Table 1. (Cont'd)

Treatment, study (ref.)	Study design	Duration of follow-up, years	Disease characteristics of cohort	Treatment group				Control group			
				Drug and dosage	No. of patients	Disease duration, mean ± SD years	Baseline mSASSS, mean ± SD	Drug	No. of patients	Disease duration, mean ± SD years	Baseline mSASSS. mean ± SD
NSAIDs											
Wanders 2005 (37)	RCT, open-label, radiographs blinded	2	AS patients	Continuous NSAIDs (started on celecoxib 200 mg BID but allowed to increase or change NSAID)	76	13 ± 10.2	7.9 ± 14.7	On-demand NSAIDs (celecoxib or another NSAID)	74	10.2 ± 9.3	9.3 ± 15.2
Poddubnyy 2012 (48)	Cohort study	2 (each group)	Axial SpA (GESPIC); AS; nonradiographic axial SpA	High NSAID intake (NSAID index ≥50)	43; 24; 19	-; 5.5 ± 2.7; 3.7 ± 2.1	-; 6.7 ± 7.7; 1.6 ± 4.0	Low NSAID intake (NSAID index <50)	121; 64; 57	-; 5.0 ± 2.9; 3.0 ± 2.2	-; 5.7 ± 11.6; 2.6 ± 4.8
Gensler 2018 (26)	Cohort, prospective	2, 4	AS (PSOAS cohort)	NSAIDs	343	16.8 ± 12.5	14.2 ± 19.6	No NSAIDs	176	16.8 ± 12.5	14.2 ± 19.6
Schiotis 2013 (49)	Cohort study	3	AS (REGISPONSER), BASRI spine ≥12 excluded	Continuous NSAIDs	81	25.32 ± 9.38	BASRI spine score 7.23 ± 3.34	On-demand NSAIDs	37	25.54 ± 12.3	BASRI spine score 6.69 ± 3.4
Minhas 2016 (25)	Cohort, prospective	>2 up to >10 (median 3)	AS (PSOAS cohort)	NSAID index >50	630 (total in both groups)	–	–	–	630 (total in both groups)	–	–
Sieper 2016 (36)	RCT, open-label, radiographs blinded	2	AS (ENRADAS)	Continuous diclofenac 150 mg/day (or equivalent dose if not tolerated)	85	12.2 ± 10.3	11.3 ± 14.9	On-demand NSAIDs	82	15.2 ± 12.4	14.0 ± 16.8
Molnar 2018 (34)	Cohort study	10 (2-year radiographic interval progression)	AS (patients fulfilling modified New York criteria for AS from SCQM axial SpA cohort)	NSAIDs	286	13.8 ± 9.7	6.6 ± 12.5	No NSAIDs	55	–	–
Rumyantseva 2018 (39)	Cohort study	2	Early axial SpA	Continuous NSAIDs	35	24.1 ± 15.4	SIJoint score 3.67 ± 0.74	On-demand NSAIDs	33	24.1 ± 15.4	SIJoint score 3 ± 1.48
Secukinumab											
Braun 2018 (50)	Retrospective analysis of RCT	2	AS (MEASURE 1)	Secukinumab 150 mg or 75 mg IV every 4 weeks	168	–	9.55 ± 14.1	Biologics-naive (taking only NSAIDs), ENRADAS cohort	69	–	9.95 ± 13.76

* Ankylosing spondylitis (AS) was defined as meeting the modified New York criteria. mSASSS = modified Stoke AS Spine Score; TNFi = tumor necrosis factor inhibitor; RCT = randomized controlled trial; IV = intravenous; GESPIC = German Spondylitis Inception Cohort; OASIS = Outcome Assessment in AS; ASSERT = AS Study for the Evaluation of Recombinant Infliximab Therapy; SC = subcutaneous; ATLAS = Adalimumab Trial Evaluating Long-term Efficacy and Safety for AS; NSAID = nonsteroidal antiinflammatory drug; MTX = methotrexate; SSZ = sulfasalazine; GO-RAISE = Golimumab in AS; DIKAS = Seutsche Infliximab Kohorte fur AS; PSOAS = Prospective Study of Outcomes in AS; OSKAR = Observation Study of Korean Spondyloarthropathy Registry; SpA = spondyloarthritis; EMBARK = Evaluation of Change in Structural Radiographic Sacroiliac Joint Damage after 2 years of Etanercept Therapy; SI = sacroiliac; DESIR = French Cohort of Undifferentiated Spondyloarthritis; SCQM = Swiss Clinical Quality Management; BID = twice a day; REGISPONSER = Spanish National Registry of Spondyloarthropathies; BASRI = Bath AS Radiology Index; ENRADAS = Effects of NSAIDs on Radiographic Damage in AS; MEASURE 1 = 16 Week Efficacy and 2 Year Long Term Safety and Efficacy of Secukinumab in Patients with Active AS.

Table 2. Radiographic progression at the spine and sacroiliac (SI) joint in the included studies*

Treatment, study (ref.)	Duration of follow-up, years	Treatment group			Control group			Measure of difference between groups
		No. of patients	Change in mSASSS, mean ± SD	No. of new syndesmophytes or other radiographic parameters, mean ± SD	No. of patients	Change in mSASSS, mean ± SD	No. of new syndesmophytes or other radiographic parameters, mean ± SD	
TNFi								
Baraliakos 2005 (42)	2	41	0.4 ± 2.7	NA	41	0.7 ± 2.8	NA	NA
Baraliakos 2007 (22)	4	33	1.6 ± 2.6	NA	132	4.4 ± 26.21	NA	NA
van der Heijde 2008 (24)	2	201	0.9 ± 2.6	NA	70	1.2 ± 3.9	NA	NA
van der Heijde 2008 (23)	2	257	0.91 ± 2.45	Change in cervical spine radiography score 0.49 ± 1.40; change in lumbar spine radiography score 0.42 ± 1.84	76	1.27 ± 3.64	Change in cervical spine radiography score 0.53 ± 2.29; change in lumbar spine radiography score 0.73 ± 2.00	NA
van der Heijde 2009 (21)	2	307	0.8 ± 2.6	NA	77	0.9 ± 4.1	NA	NA
Pedersen 2011 (45)	2	23	1.4 ± 1.9	0.52 ± 0.8	27	1.5 ± 3.1	0.70 ± 1.4	NA
Haroon 2013 (27)	>1.5–9	201	NA	NA	133	NA	NA	mSASSS OR 0.52 (95% CI 0.30, 0.88)
Kang 2013 (46)	2	26	3.3 ± 4.2	0.9 ± 1.4	37	2.3 ± 5.1	0.4 ± 0.9	NA
Min 2014 (38)	8	14	9.29 ± 6.22	2.79 ± 2.94	12	4.58 ± 2.15	0.50 ± 0.674	NA
Braun 2014 (35)	2; 4	233	0.9 ± 3.33; 1.67 ± 4.89	NA; NA	66	1.6 ± 4.6; 2.1 ± 5.2	NA; NA	NA
Baraliakos 2014 (43)	2; 4; 6; 8	22	3.2 ± 7.04; 4.4 ± 6.57; 6.2 ± 6.57; 7.2 ± 6.57	NA; NA; NA; 1 ± 8.64	34	2.7 ± 9.33; 17.5 ± 10.49; 19.3 ± 7.58; 11.70 ± 6.41	NA; NA; NA; 2.70 ± 6.25	NA
Chung 2015 (47)	4	25	NA	Change in CT score of facet joints in total spine 11.72 ± 53.94	25	NA	Change in CT score of facet joints in total spine 11.56 ± 43.02	NA
Minhas 2016 (25)	>2, up to >10	630	NA	NA	–	–	–	Change in BASRI SI joint score of radiographic progression OR 0.06 (95% CI 0.004, 0.99)
Kim 2016 (55)	5	269	4.73 ± 18.56	NA	341	6.14 ± 32.80	NA	NA
Dougados 2018 (41)	2	162	NA	Change in total SI joint score −0.14 ± 0.81	193	NA	Change in total SI joint score 0.08 ± 0.83	Difference in total SI joint score −0.22 (95% CI −0.38,−0.06)

(Continued)

Table 2. *(Cont'd)*

		Treatment group			Control group			
Treatment, study (ref.)	Duration of follow-up, years	No. of patients	Change in mSASSS, mean ± SD	No. of new syndesmophytes or other radiographic parameters, mean ± SD	No. of patients	Change in mSASSS, mean ± SD	No. of new syndesmophytes or other radiographic parameters, mean ± SD	Measure of difference between groups
Molnar 2018 (34)	10 (2-year radiographic interval progression)	163	NA	NA	269	NA	NA	2-year radiographic progression OR 0.50 (95% CI 0.28, 0.88); 2-year progression of ≥1 new syndesmophyte OR 0.55 (95% CI 0.33, 0.94)
Gensler 2018 (26)	2; 4	239	NA; NA	NA; NA	280	NA; NA	NA; NA	Mean mSASSS difference −0.71 (95% CI −2.10, 0.68; $P = 0.32$); mean mSASSS difference −1.37 (95% CI −2.07, −0.63; $P < 0.001$)
Park 2019 (44)	4	135	NA	NA	80	NA	NA	Radiographic progression OR 0.41 (95% CI 0.22, 0.75)
NSAIDs								
Wanders 2005 (37)	2	76	0.4 ± 1.7	NA	74	1.5 ± 2.5	NA	NA
Poddubnyy 2012 (48)	2	24 (AS); 19 (non-radiographic axial SpA)	0.02 ± 1.39; 0.51 ± 1.72	NA; NA	64 (AS); 57 (non-radiographic axial SpA)	0.96 ± 2.78; 0.74 ± 1.95	NA; NA	NA
Gensler 2018 (26)	2; 4	343	NA; NA	NA; NA	176	NA; NA	NA; NA	Mean mSASSS difference 0.64 (95% CI −1.52, 2.81; $P = 0.56$); mean mSASSS difference 1.95 (95% CI 0.90, 2.99; $P < 0.001$)
Schiotis 2013 (49)	3	81	NA	Change in BASRI spine score 0.64 ± 1.22	37	NA	Change in BASRI spine score 0.66 ± 1.04	NA
Minhas 2016 (25)	>2, up to >10	630	NA	NA	–	–	–	Change in BASRI SI joint score OR 0.03 (95% CI 0.002, 3.50)

(Continued)

Table 2.   *(Cont'd)*

Treatment, study (ref.)	Duration of follow-up, years	Treatment group			Control group			Measure of difference between groups
		No. of patients	Change in mSASSS, mean ± SD	No. of new syndesmophytes or other radiographic parameters, mean ± SD	No. of patients	Change in mSASSS, mean ± SD	No. of new syndesmophytes or other radiographic parameters, mean ± SD	
Sieper 2016 (36)	2	85	1.29 ± 2.82	NA	82	0.71 ± 2.77	NA	NA
Molnar 2018 (34)	10 (2-year radiographic interval progression)	286	NA	NA	55	NA	NA	2-year radiographic progression OR 0.85 (95% CI 0.42, 1.72); 2-year progression of ≥1 new syndesmophyte OR 1.04 (95% CI 0.50, 2.17)
Rumyantseva 2018 (39)	2	35	NA	Change in SI joint score 1 ± 1.66	33	NA	Change in SI joint score 1.33 ± 2.67	NA
Secukinumab								
Braun 2018 (50)	2	168	0.55 ± 1.82	Patients with no radiographic progression 61%	69	0.89 ± 1.83	Patients with no radiographic progression 52%	NA

* mSASSS = modified Stoke Ankylosing Spondylitis Spine Score; TNFi = tumor necrosis factor inhibitor; NA = not available; OR = odds ratio; 95% CI = 95% confidence interval; CT = computed tomography; BASRI = Bath AS Radiology Index (of the spine); NSAIDs = nonsteroidal antiinflammatory drugs; SpA = spondyloarthritis.

scores of the facet joints in the total spine (47) (Table 2). No studies on spinal radiographic progression with TNFi treatment were found in patients with nonradiographic axial SpA. Five studies also measured change in the number of syndesmophytes (34,38,43,45,46).

Spinal radiographic progression was not significantly different between the TNFi-treated and biologics-naive populations at 2 years (mSASSS difference –0.73 [95% CI –1.52, 0.12], I^2 = 28%) and at ≥4 years (mSASSS difference –2.03 [95% CI –4.63, 0.72],

I^2 = 63%) (Figures 2A and B). However, a sensitivity analysis restricted to 6 studies with a low risk of bias (excluding 1 study [38]) showed a significant difference in spinal radiographic progression at ≥4 years in the TNFi-treated patients compared to biologics-naive patients (mSASSS difference –2.17 [95% CI –4.19, –0.15], I^2 = 49%) (Figure 2C). Certainty with regard to this finding, as determined using the GRADE approach, was low (see Supplementary File 2 [http://onlinelibrary.wiley.com/doi/10.1002/art.41206/abstract]).
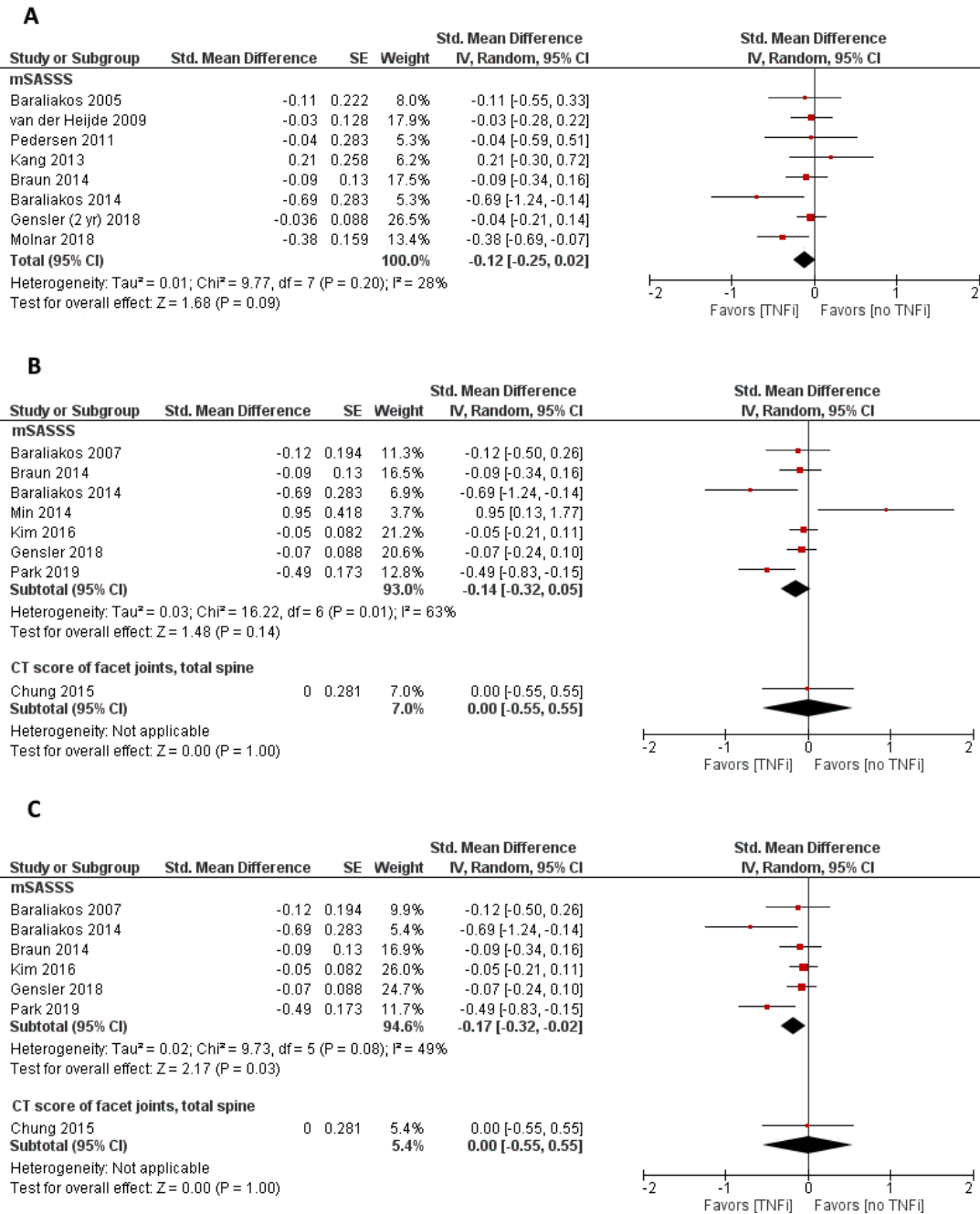
Figure 2. Forest plots of radiographic progression at the spine in patients with ankylosing spondylitis treated with tumor necrosis factor inhibitors (TNFi), measured at 2 years (**A**) and at ≥4 years of follow-up (**B**) with the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) and the computed tomography (CT) score of the facet joints. In sensitivity analyses, data on radiographic progression were obtained only from studies with a low risk of bias (**C**). Std. = standard; IV = inverse variance method (in RevMan); 95% CI = 95% confidence interval. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41206/abstract.

As explained above, conversion of all measures to the mSASSS was done by multiplying the SMD by the average pooled SD, yielding mSASSS radiographic progression scores of 6.06 at 2 years (21,26,35,42,45,46) and 14.47 at 4 years (22,26,35,38,43,55). Only 1 of the studies used change in CT score of the facet joints as the measure of radiographic progression, which waboveificantly different between the TNFi-treated and biologics-naive groups (CT score difference −0.16 [95% CI −27.91, 27.59]) (47).

The largest study in which the OASIS cohort was used as the control group (21) was included in the primary analysis (1 study with 2 years and 1 study with ≥4 years of follow-up). Sensitivity analyses performed with other studies that used the OASIS cohort as the control, carried out one at a time, did not show any difference between the treatment and control groups.

Subgroup analyses performed to assess heterogeneity did not reveal significant differences between the historical and contemporary control groups ($P = 0.54$). Similarly, there was no difference in the number of syndesmophytes between the TNFi-treated and biologics-naive groups at 2 years of follow-up (SMD −0.04 [95% CI −0.51, 0.43]; change in number of syndesmophytes −0.05 [95% CI −0.59, 0.49], $I^2 = 69\%$) or 8 years of follow-up (SMD 0.34 [95% CI −0.86, 1.55]; change in number of syndesmophytes 0.78 [95% CI −3.01, 4.57], $I^2 = 83\%$). The average pooled SD (1.15 at 2 years [45,46] and 2.29 at 8 years [38,43]) from studies reporting a change in the number of syndesmophytes was used for the conversion (46).

Only 2 studies (25,41) showed radiographic changes at the SI joint. Minhas et al reported an odds ratio of 0.06 (95% CI 0.004, 0.99) for the change in the BASRI SI joint score (measuring radiographic change in the SI joint) in the TNFi-treated group compared to the biologics-naive group in the PSOAS cohort (25). Dougados et al reported a change in the total SI joint score of −0.22 (95% CI −0.38, −0.06) in the TNFi-treated group compared to the biologics-naive group of patients with nonradiographic axial SpA ($P = 0.008$) (41). Pooled estimates were not calculated for the 2 studies, as the study populations were different (AS versus nonradiographic axial SpA) (Figures 3A and B).

NSAIDs. Among the 6 studies reporting mSASSS scores at 2 years following treatment with NSAIDs in patients with AS, no significant difference was observed between the NSAID-treated and control groups (SMD −0.08 [95% CI −0.32, 0.16]; mSASSS difference −0.30 [95% CI −2.62, 1.31], $I^2 = 71\%$) (26,34,36,37,48) (Figure 4). The average pooled SD value of 8.18 was used for the conversion to mSASSS (26,36,37).

Dosing strategies for NSAIDs in both the treatment and control arms were different among the studies: comparisons were made between continuous and on-demand NSAIDs (36,37), between an NSAID index of >50 (high) and an NSAID index of <50 (low) (48), and between NSAID use and no NSAID use (26,34). Subgroup analyses using these comparisons (continuous versus on-demand NSAIDs, NSAID index high versus low, and NSAID use versus no NSAID use) showed no differences (each $P = 0.79$). There was one study in which the BASRI spine score was compared between the NSAID and control groups, and no difference was found (BASRI spine score difference 0.020 [95% CI −0.44, 0.48]) (49). One study
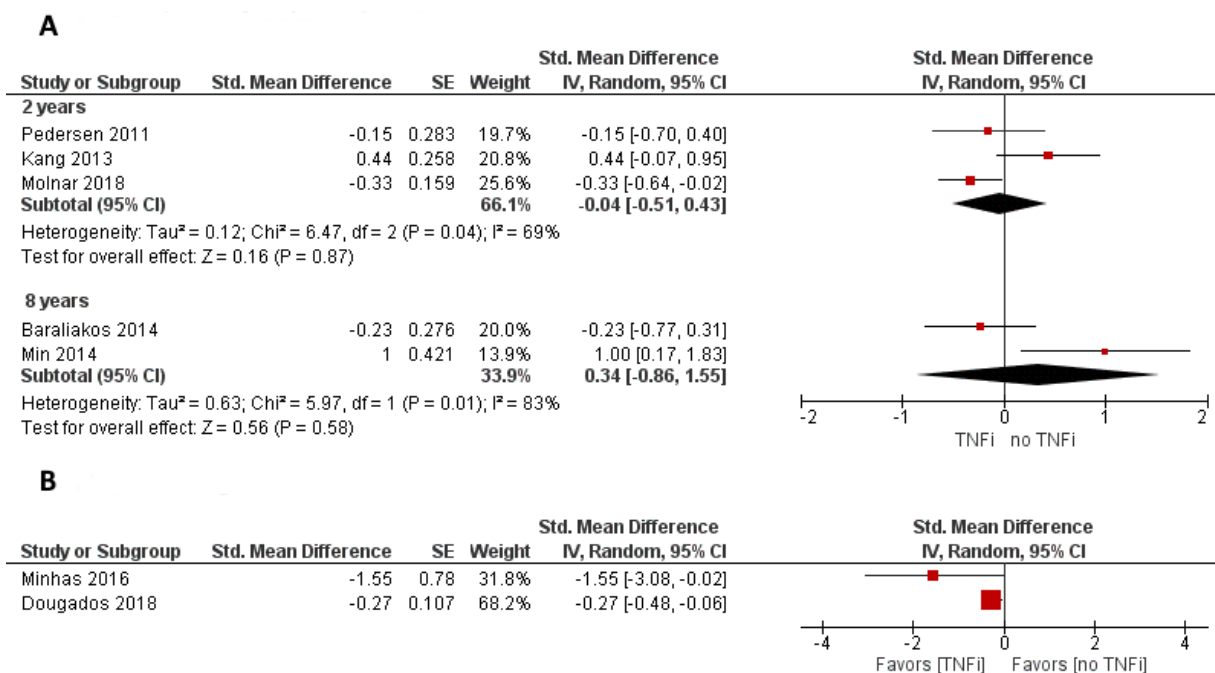
Figure 3. Forest plots of radiographic progression in patients treated with TNFi, measured as change in the number of syndesmophytes in patients with ankylosing spondylitis (**A**) and as radiographic progression at the sacroiliac joint in patients with nonradiographic axial spondyloarthritis (**B**). See Figure 2 for definitions. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41206/abstract.
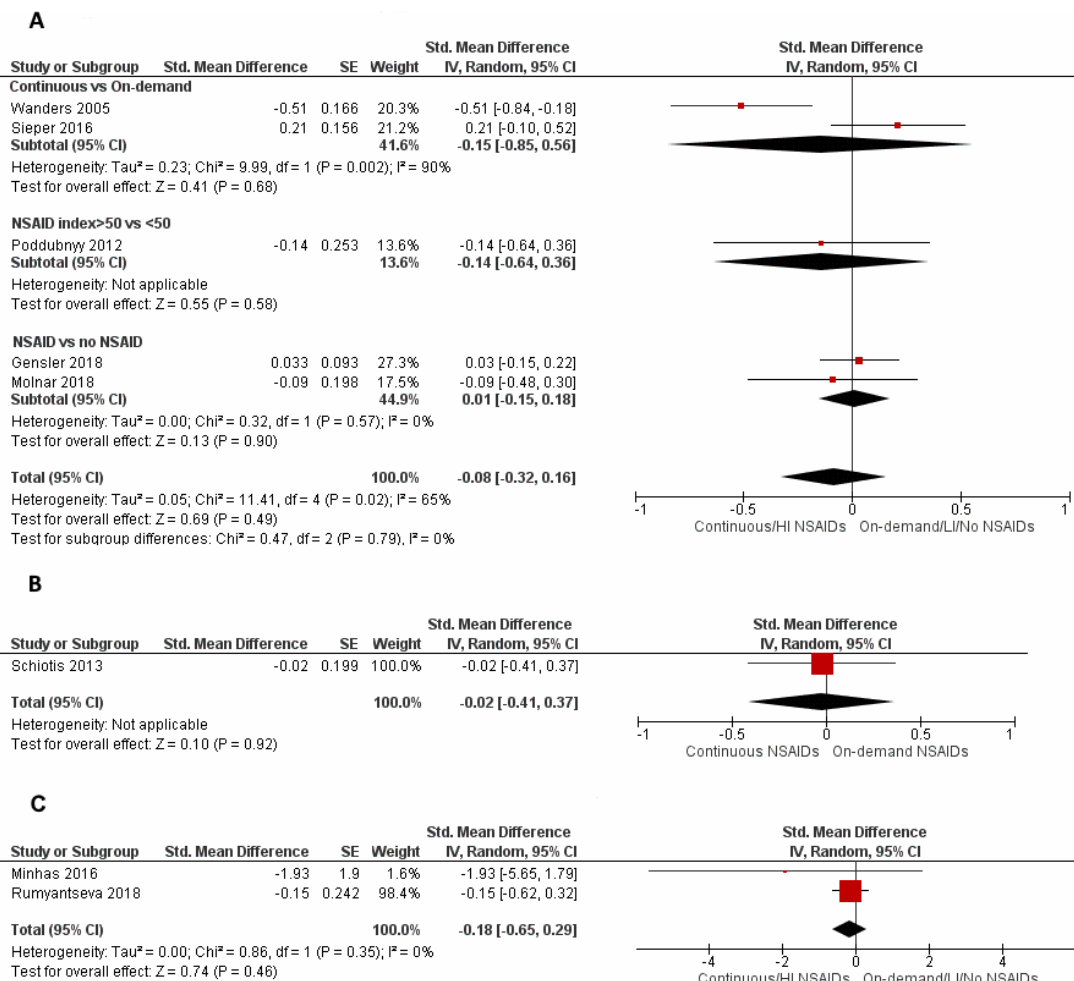
Figure 4. Forest plots of radiographic progression in patients treated with nonsteroidal antiinflammatory drugs (NSAIDs). Comparisons were made between continuous and on-demand NSAIDs, between an NSAID index of >50 (high index [HI]) and an NSAID index of <50 (low index [LI]), and between NSAID use and no NSAID use. In patients with ankylosing spondylitis (AS), radiographic progression at the spine was determined at 2 years (measured with the mSASSS) (**A**) and at 3 years (measured with the Bath AS Radiology Index for the spine) (**B**). In patients with early axial spondyloarthritis, radiographic progression was measured at the sacroiliac joint (**C**). See Figure 2 for other definitions. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41206/abstract.

showed NSAID effects in a subgroup of patients with nonradiographic axial SpA, in whom no difference in mSASSS scores was observed between the NSAID-treated patients and controls (mSASSS difference 0.13 [95% CI –0.39, 0.65]) (48). A sensitivity analysis in which an observational study that was judged to have a high risk of bias (39) was removed did not change the results.

Two studies assessing NSAIDs (25,39) demonstrated changes in radiographic progression based on the BASRI SI joint score, and no difference in the SI joint radiographic progression score was seen between NSAID-treated and control groups (SMD –0.18 [95% CI –0.65, 0.29]; SI joint score difference –0.40 [95% CI –1.44, 0.64], I^2 = 0%). The pooled SD of 2.21 was used for the conversion to mSASSS (39).

Secukinumab. The only included study in which secukinumab was evaluated did not show a significant difference in radiographic progression compared to controls over 2 years (mean mSASSS difference –0.34 [95% CI –0.85, 0.17]) (50).

DISCUSSION

This systematic review and meta-analysis showed that TNFi may slow radiographic progression at the spine in patients with AS at ≥4 years of follow-up (when only studies judged to have a low risk of bias were included) but not at 2 years. Results of a recent study suggested that progression might be less beyond 2 years (54), and a nonlinear, continued benefit with TNFi has been observed beyond 4 years (26,56). These studies adjusted for important confounders (e.g., smoking status, NSAID use, baseline mSASSS, and disease/symptom duration), which might account for the difference in results. Our study showed a larger difference in the mSASSS between the TNFi-treated and biologics-naive groups at ≥4 years, suggesting that there is a continued benefit with long-term use of TNFi. The difference in radiographic progression at the SI joint in patients with AS and patients with nonradiographic axial SpA was not significant in each of the 2 included

studies (Figure 3). The semiquantitative nature of the SI joint score, with interreader variability, makes the study of radiographic progression at the SI joint challenging.

In contrast, treatment with NSAIDs did not show any significant inhibition of radiographic progression at the spine in patients with AS at 2 years or at the SI joint at >2 years. A potentially clinically significant effect might be observed over long-term follow-up, and therefore long-term studies are needed. Furthermore, differential effects of NSAIDs have been noted, and are believed to be related to the degree of cyclooxygenase 2 (COX-2) selectivity. Wanders et al showed that patients with AS receiving continuous celecoxib had lower radiographic progression in comparison to those receiving on-demand treatment over 2 years, although the cumulative dose was not much different between the groups (37). Moreover, the study was not blinded, which may have resulted in differential use of cointerventions, such as exercise or smoking cessation. In a post hoc analysis of this trial, those with elevated levels of acute-phase reactants seemed to achieve the greatest benefit from treatment with continuous celecoxib compared to those who received on-demand NSAIDs (57). It is postulated that elevated levels of prostaglandin E_2 will lead to increased osteoblastic activity, and therefore inhibition of prostaglandin (especially prostaglandin E_2) synthesis by COX-2 inhibitors might inhibit new bone formation (58,59). However, a subsequent, nonblinded RCT (ENRADAS trial) that evaluated diclofenac failed to confirm these findings (36). It is unclear whether the differential effect on bone formation with selective COX-2 inhibitors (as used in Wanders et al), as opposed to that with nonselective NSAIDs (as used in the ENRADAS trial), led to the difference in results. At present, there is insufficient evidence to confirm the effect of NSAIDs (selective or nonselective) on radiographic progression alone or in combination with TNFi. The long-term risk of different NSAIDs in this population should be studied further, in terms of cardiovascular and gastrointestinal safety, to justify the risk–benefit profile.

The only study in which secukinumab was used did not show a significant difference in radiographic progression at 2 years (50). More studies are required to better understand the long-term effects of secukinumab and its effects in early disease. Phase III data from the secukinumab trial have shown no increase in spinal radiographic damage in 80% of patients with AS at 2 and 4 years (52,53), but these data need to be interpreted with caution in the absence of a control group.

A paucity of data on the effect of treatment on radiographic progression in patients with nonradiographic axial SpA was noted. A subgroup of patients with nonradiographic axial SpA from the GESPIC cohort showed no difference in spinal radiographic progression, as measured by the mSASSS, between the high and low NSAIDs use groups at 2 years (48). Only 1 study exploring the effects of TNFi on SI joint radiographic progression (from the Evaluation of Change in Structural Radiographic SI Joint Damage after 2 years of Etanercept Therapy [EMBARK] trial) was included in our study, which did not show any significant difference (41). A

study by Almirall et al and the Certolizumab Pegol in Subjects with Active Axial SpA (RAPID-axSpA) study were not included in our meta-analysis, as all patients in these studies were treated with TNFi with no comparator arm (54,60). These studies showed no radiographic progression at the spine/SI joint at 2 years or at the SI joint at 4 years. However, the data are difficult to interpret in the absence of a controlled comparison. No data on secukinumab or on biologic agents with other modes of action were found.

Although our study showed a significant effect of TNFi on long-term radiographic progression (in sensitivity analyses), none of the included studies provided a prospective, long-term controlled comparison. Most included studies were judged to have a low risk of bias; however, a predominance of observational and open-label extensions of RCTs limited the overall level of evidence. Most analyses in our study showed low-to-moderate heterogeneity, and a few (studies with NSAIDs) showed high heterogeneity. There were methodologic differences between the studies, and various subgroup and sensitivity analyses, e.g., historical versus contemporary controls were performed, and various dose levels of NSAIDs were studied. These may explain some of the heterogeneity in the data, but not all.

Definitively answering the question as to which treatments have the optimal beneficial effects on radiographic progression in axial SpA will require a concerted effort. Long-term, controlled trials of axial SpA therapies are costly, and therefore alternative strategies will be necessary to learn which therapies are best in preventing radiographic progression. Well-designed RCTs with head-to-head comparisons will be required to establish the comparative efficacy of biologic therapies, either alone or in combination with NSAIDs, in slowing radiographic progression. A head-to-head study of secukinumab with TNFi is planned (61), which will hopefully give us more information regarding their comparative efficacy. There is an ongoing RCT in which radiographic progression at 2 years is being compared between treatment with TNFi alone and the combination of TNFi and NSAIDs, based on promising data from observational studies supporting the combination approach (26). There is also increasing evidence to suggest that TNFi in patients with early AS (<10 years' disease duration) is associated with a higher benefit.

Observational studies by Haroon et al (27) and Park et al (44) have both suggested the importance of early initiation of TNFi therapy for controlling radiographic progression. Data from observational studies will be limited because of the expense of serial imaging tests. Using observational data to assess radiographic progression will likely require funding specifically to obtain serial imaging at standardized intervals among patients who are treated with TNFi or interleukin-17 inhibitors alone or in combination with NSAIDs.

Moreover, risk stratification to identify those at high risk of progression is important, given that not all patients with axial SpA experience spinal disease progression. Risk factors for radiographic progression of axial SpA, as noted in multiple studies,

are male sex, HLA–B27 positivity, baseline radiographic changes, long duration of disease and symptoms, high C-reactive protein levels, high disease activity, and smoking status (62), and, more recently, alcohol has been implicated (63). Close monitoring and early therapy with a treat-to-target strategy might have a greater impact on slowing structural damage in this high-risk group.

Furthermore, we need more sensitive and reliable measures to document radiographic progression. While most studies used the mSASSS, which is the most validated measure for radiographic progression in AS, it is based on plain radiographs with limited sensitivity to change. The mSASSS does not include assessment of changes at the thoracic spine or posterior elements (facet joints), and cannot assess early damage (64,65). Interreader reliability in the assessment of change in the mSASSS is also of concern, with most studies showing poor-to-moderate reliability (21,24,51). Although plain radiography is cheap and fast and has been backed up by years of experience with reading, it might not be the best measure for assessment due to the slow nature of radiographic progression in AS. Newer measures, such as those based on quantitative low-dose CT scanning, may provide higher sensitivity in detecting small changes in syndesmophyte size/volume (66,67). Low-dose CT has also been shown to have a good correlation with various measures of patient function, such as the Schober's test and lateral thoracolumbar flexion (68). Comparison of radiographic measurement (the mSASSS) with these newer modalities will provide guidance as to whether CT is an acceptable outcome measure in clinical trials.

In conclusion, although no significant protective effect of TNFi treatment on radiographic progression of AS at the spine at 2 years and 4 years was found in our study, analyses restricted to studies with a low risk of bias showed a protective effect of TNFi at ≥4 years. Therefore, long-term TNFi exposure might have beneficial effects on radiographic progression. No difference was seen with NSAIDs or with secukinumab (only 1 study) at 2 years, but long-term data were not available. Further studies should explore the effect of NSAIDs and biologics alone and in combination in patients with early axial SpA; their use in the group with high risk of progression should be evaluated with a follow-up of >4 years to see whether the effects are more pronounced over time. Newer measures with higher sensitivity to detect structural changes, such as those based on quantitative low-dose CT, should be compared to mSASSS for use in clinical trials.

ACKNOWLEDGMENTS

REFERENCES

1. 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.

2. Baraliakos X, Listing J, von der Recke A, Braun J. The natural course of radiographic progression in ankylosing spondylitis—evidence for major individual variations in a large proportion of patients. J Rheumatol 2009;36:997–1002.

3. Poddubnyy D, Haibel H, Listing J, Märker-Hermann E, Zeidler H, Braun J, et al. Baseline radiographic damage, elevated acute-phase reactant levels, and cigarette smoking status predict spinal radiographic progression in early axial spondylarthritis. Arthritis Rheum 2012;64:1388–98.

4. Braun J, Golder W, Bollow M, Sieper J, van der Heijde D. Imaging and scoring in ankylosing spondylitis. Clin Exp Rheumatol 2002;20 Suppl 28:20:S178–84.

5. Oostveen J, Prevo R, den Boer J, van de Laar M. Early detection of sacroiliitis on magnetic resonance imaging and subsequent development of sacroiliitis on plain radiography: a prospective, longitudinal study. J Rheumatol 1999;26:1953–8.

6. Wanders AJ, Landewé RB, Spoorenberg A, Dougados M, van der Linden S, Mielants H, et al. What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the Outcome Measures in Rheumatology Clinical Trials filter. Arthritis Rheum 2004;50:2622–32.

7. Van der Heijde D, Landewé R. Selection of a method for scoring radiographs for ankylosing spondylitis clinical trials, by the Assessment in Ankylosing Spondylitis Working Group and OMERACT. J Rheumatol 2005;32:2048–9.

8. Creemers MC, Franssen MJ, van 't Hof MA, Gribnau FW, van de Putte LB, van Riel PL. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis 2005;64:127–9.

9. Ramiro S, Claudepierre P, Sepriano A, van Lunteren M, Molto A, Feydy A, et al. Which scoring method depicts spinal radiographic damage in early axial spondyloarthritis best? Five-year results from the DESIR cohort. Rheumatology (Oxford) 2018;57:1991–2000.

10. MacKay K, Mack C, Brophy S, Calin A. The Bath Ankylosing Spondylitis Radiology Index (BASRI): a new, validated approach to disease assessment. Arthritis Rheum 1998;41:2263–70.

11. Spoorenberg A, de Vlam K, van der Linden S, Dougados M, Mielants H, van de Tempel H, et al. Radiological scoring methods in ankylosing spondylitis: reliability and change over 1 and 2 years. J Rheumatol 2004;31:125–32.

12. Ramiro S, Stolwijk C, van Tubergen A, van der Heijde D, Dougados M, van den Bosch F, et al. Evolution of radiographic damage in ankylosing spondylitis: a 12 year prospective follow-up of the OASIS study. Ann Rheum Dis 2015;74:52–9.

13. Landewé 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.

14. Poddubnyy D, Listing J, Haibel H, Knüppel 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.

15. Beller EM, Glasziou PP, Altman DG, Hopewell S, Bastian H, Chalmers I, et al, for the PRISMA for Abstracts Group. PRISMA for Abstracts: reporting systematic reviews in journal and conference abstracts. PLoS Med 2013;10:e1001419.

16. Exarchou S, Redlund-Johnell I, Karlsson M, Mellström D, Ohlsson C, Turesson C, et al. The prevalence of moderate to severe radiographic sacroiliitis and the correlation with health status in elderly Swedish men—the MrOS study. BMC Musculoskelet Disord 2013;14:352.

17. Van den Berg R, Lenczner G, Feydy A, van der Heijde D, Reijnierse M, Saraux A, et al. Agreement between clinical practice and trained central reading in reading of sacroiliac joints on plain pelvic radiographs: results from the DESIR cohort. Arthritis Rheumatol 2014;66:2403–11.

18. Higgins JP, Green S, editors. 9.4.6 combining dichotomous and continuous outcomes. In: Cochrane handbook for systematic reviews of interventions. URL: https://handbook-5-1.cochrane.org/chapter_9/9_4_6_combining_dichotomous_and_continuous_outcomes.htm.

19. DerSimonian R, Laird NM. Meta-analysis in clinical trials. Controlled Clinical Trials 1986;7:177–88.

20. Murad MH, Wang Z, Chu H, Lin L. When continuous outcomes are measured using different scales: guide for meta-analysis and interpretation. BMJ 2019;364:k4817.

21. Van der Heijde D, Salonen D, Weissman BN, Landewé R, Maksymowych WP, Kupper H, et al. Assessment of radiographic progression in the spines of patients with ankylosing spondylitis treated with adalimumab for up to 2 years. Arthritis Res Ther 2009;11:R127.

22. Baraliakos X, Listing J, Brandt J, Haibel H, Rudwaleit M, Sieper J, et al. Radiographic progression in patients with ankylosing spondylitis after 4 yrs of treatment with the anti-TNF-α antibody infliximab. Rheumatology (Oxford) 2007;46:1450–3.

23. Van der Heijde D, Landewé R, Einstein S, Ory P, Vosse D, Ni L, et al. Radiographic progression of ankylosing spondylitis after up to two years of treatment with etanercept. Arthritis Rheum 2008;58:1324–31.

24. Van der Heijde D, Landewé R, Baraliakos X, Houben H, van Tubergen A, Williamson P, et al, and the Ankylosing Spondylitis Study for the Evaluation of Recombinant Infliximab Therapy Study Group. Radiographic findings following two years of infliximab therapy in patients with ankylosing spondylitis. Arthritis Rheum 2008;58:3063–70.

25. Minhas D, Lee M, Rahbar MH, Gensler LS, Reveille JD, Weisman M, et al. Do TNF inhibitors change the progression of sacroiliitis? [abstract]. Arthritis Rheumatol 2016;68 Suppl 10. URL: https://acrabstracts.org/abstract/do-tnf-inhibitors-change-the-progression-of-sacroiliitis/.

26. Gensler LS, Gianfrancesco M, Weisman MH, Brown MA, Lee M, Learch T, et al. Combined effects of tumour necrosis factor inhibitors and NSAIDs on radiographic progression in ankylosing spondylitis [abstract]. Ann Rheum Dis 2018;77 Suppl 2:148.

27. Haroon N, Inman RD, Learch TJ, Weisman MH, Lee M, Rahbar MH, et al. The impact of tumor necrosis factor α inhibitors on radiographic progression in ankylosing spondylitis. Arthritis Rheum 2013;65:2645–54.

28. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol 2014;14:135.

29. Higgins JP, Green S, editors. Cochrane handbook for systematic reviews of interventions. URL: https://handbook-5-1.cochrane.org/front_page.htm.

30. Wells GA, Shea B, O'Connell D, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

31. Risk of bias tools. URL: https://www.riskofbias.info/welcome/rob-2-0-tool.

32. Murad MH. Clinical practice guidelines: a primer on development and dissemination. Mayo Clin Proc 2017;92:423–33.

33. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58.

34. Molnar C, Scherer A, Baraliakos X, de Hooge M, Micheroli R, Exer P, et al. TNF blockers inhibit spinal radiographic progression in ankylosing spondylitis by reducing disease activity: results from the Swiss Clinical Quality Management cohort. Ann Rheum Dis 2018;77:63–9.

35. Braun J, Baraliakos X, Hermann KG, Deodhar A, van der Heijde D, Inman R, et al. The effect of two golimumab doses on radiographic progression in ankylosing spondylitis: results through 4 years of the GO-RAISE trial. Ann Rheum Dis 2014;73:1107–13.

36. Sieper J, Listing J, Poddubnyy D, Song IH, Hermann KG, Callhoff J, et al. Effect of continuous versus on-demand treatment of ankylosing spondylitis with diclofenac over 2 years on radiographic progression of the spine: results from a randomised multicentre trial (ENRADAS). Ann Rheum Dis 2016;75:1438–43.

37. Wanders A, van der Heijde D, Landewé R, Béhier JM, Calin A, Olivieri I, et al. Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis Rheum 2005;52:1756–65.

38. Min HK, Kang JY, Koh JH, Jung SM, Lee J, Kwok SK, et al. Comparisons of radiographic progression of ankylosing spondylitis between treatment with TNF antagonist, continuous treatment with NSAID, and on demand treatment of NSAID [abstract]. Clin Exp Rheumatol 2014;32:829.

39. Rumyantseva DG, Dubinina TV, Erdes SF. Impact of the frequency of using nonsteroidal anti-inflammatory drugs on the radiographic progression of sacroiliitis in patients with early axial spondyloarthritis. Nauchno-Prakticheskaya Revmatologiya 2018;56:346–50. In Russian.

40. Boersma JW. Retardation of ossification of the lumbar vertebral column in ankylosing spondylitis by means of phenylbutazone. Scand J Rheumatol 1976;5:60–4.

41. Dougados M, Maksymowych WP, Landewé RB, Moltó A, Claudepierre P, de Hooge M, et al. Evaluation of the change in structural radiographic sacroiliac joint damage after 2 years of etanercept therapy (EMBARK trial) in comparison to a contemporary control cohort (DESIR cohort) in recent onset axial spondyloarthritis. Ann Rheum Dis 2018;77:221–7.

42. Baraliakos X, Listing J, Rudwaleit M, Brandt J, Sieper J, Braun J. Radiographic progression in patients with ankylosing spondylitis after 2 years of treatment with the tumour necrosis factor α antibody infliximab. Ann Rheum Dis 2005;64:1462–6.

43. Baraliakos X, Haibel H, Listing J, Sieper J, Braun J. Continuous long-term anti-TNF therapy does not lead to an increase in the rate of new bone formation over 8 years in patients with ankylosing spondylitis. Ann Rheum Dis 2014;73:710–5.

44. Park JW, Kim MJ, Lee JS, Ha YJ, Park JK, Kang EH, et al. Impact of tumor necrosis factor inhibitor versus nonsteroidal antiinflammatory drug treatment on radiographic progression in early ankylosing spondylitis: its relationship to inflammation control during treatment. Arthritis Rheumatol 2019;71:82–90.

45. Pedersen SJ, Chiowchanwisawakit P, Lambert RG, Østergaard M, Maksymowych WP. Resolution of inflammation following treatment of ankylosing spondylitis is associated with new bone formation. J Rheumatol 2011;38:1349–54.

46. Kang KY, Ju JH, Park SH, Kim HY. The paradoxical effects of TNF inhibitors on bone mineral density and radiographic progression in patients with ankylosing spondylitis. Rheumatology (Oxford) 2013;52:718–26.

47. Chung SW, Song R, Lee YA, Hong SJ, Yang HI, Lee SH. Evaluation of longitudinal association between changes in facet joints and administration of tumor necrosis factor inhibitors in ankylosing spondylitis using CT [abstract]. Ann Rheum Dis 2015;74 Suppl 2:1162.

48. Poddubnyy D, Rudwaleit M, Haibel H, Listing J, Märker-Hermann E, Zeidler H, et al. Effect of non-steroidal anti-inflammatory drugs on radiographic spinal progression in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. Ann Rheum Dis 2012;71:1616–22.

49. Schiotis R, Font P, Escudero A, Buzoianu A, Zarco P, Almodovar R, et al. Long term influence of NSAIDs on radiographic progression in patients with ankylosing spondylitis [abstract]. Ann Rheum Dis 2013;72 Suppl 3:A514–5.

50. Braun J, Haibel H, De Hooge M, Landewé RB, Rudwaleit M, Readie A, et al. Low rate of spinal radiographic progression over 2 years in ankylosing spondylitis patients treated with secukinumab: a historical cohort comparison [abstract]. Arthritis Rheumatol 2018;70 Suppl 10. URL: https://acrabstracts.org/abstract/low-rate-of-spinal-radiographic-progression-over-2-years-in-ankyl osing-spondylitis-patients-treated-with-secukinumab-a-historical-cohort-comparison/.

51. Braun J, Haibel H, de Hooge M, Landewé R, Rudwaleit M, Fox T, et al. Spinal radiographic progression over 2 years in ankylosing spondylitis patients treated with secukinumab: a historical cohort comparison. Arthritis Res Ther 2019;21:142.

52. 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 2017;76:1070–7.

53. Braun J, Baraliakos X, Deodhar A, Poddubnyy D, Emery P, Delicha EM, et al. Secukinumab shows sustained efficacy and low structural progression in ankylosing spondylitis: 4-year results from the MEASURE 1 study. Rheumatology (Oxford) 2019;58:68.

54. Van der Heijde D, Baraliakos X, Hermann KA, Landewé RB, Machado PM, Maksymowych WP, et al. Limited radiographic progression and sustained reductions in MRI inflammation in patients with axial spondyloarthritis: 4-year imaging outcomes from the RAPID-axSpA phase III randomised trial. Ann Rheum Dis 2018;77:699–705.

55. Kim TJ, Shin JH, Kim S, Sung IH, Lee S, Song Y, et al. Radiographic progression in patients with ankylosing spondylitis according to tumor necrosis factor blocker exposure: Observation Study of Korean Spondyloarthropathy Registry (OSKAR) data. Joint Bone Spine 2016;83:569–72.

56. Maas F, Arends S, Brouwer E, Essers I, van der Veer E, Efde M, et al. Reduction in spinal radiographic progression in ankylosing spondylitis patients receiving prolonged treatment with tumor necrosis factor inhibitors. Arthritis Care Res (Hoboken) 2017;69:1011–9.

57. Kroon F, Landewé R, Dougados M, van der Heijde D. Continuous NSAID use reverts the effects of inflammation on radiographic progression in patients with ankylosing spondylitis. Ann Rheum Dis 2012;71:1623–9.

58. Blackwell KA, Raisz LG, Pilbeam CC. Prostaglandins in bone: bad cop, good cop? [review]. Trends Endocrinol Metab TEM 2010; 21:294–301.

59. Vuolteenaho K, Moilanen T, Moilanen E. Non-steroidal anti-inflammatory drugs, cyclooxygenase-2 and the bone healing process. Basic Clin Pharmacol Toxicol 2008;102:10–4.

60. Almirall M, López-Velandia JG, Maymó J. Absence of radiographic progression at two years in a cohort of patients with non-radiographic axial spondyloarthritis treated with TNF-α blockers. Reumatol Clin 2014;10:134–5.

61. Baraliakos X, Ostergaard M, Gensler LS, Poddubnyy D, Lee EY, Kiltz U, et al. Comparison of secukinumab and adalimumab biosimilar on radiographic progression in ankylosing spondylitis patients: design of a randomized, phase-IIIb study (SURPASS) [abstract]. Int J Rheum Dis 2018;21 Suppl 1:121–2.

62. Sari I, Haroon N. Radiographic progression in ankylosing spondylitis: from prognostication to disease modification. Curr Rheumatol Rep 2018;20:82.

63. Min HK, Lee J, Ju JH, Park SH, Kwok SK, et al. Alcohol consumption as a predictor of the progression of spinal structural damage in axial spondyloarthritis: data from the Catholic Axial Spondyloarthritis COhort (CASCO). Arthritis Res Ther 2019;21:187.

64. Van der Heijde D, Braun J, Deodhar A, Baraliakos X, Landewé R, Richards HB, et al. Modified stoke ankylosing spondylitis spinal score as an outcome measure to assess the impact of treatment on structural progression in ankylosing spondylitis. Rheumatology (Oxford) 2019;58:388–400.

65. Maas F, Arends S, Brouwer E, Bootsma H, Bos R, Wink FR, et al. Incorporating assessment of the cervical facet joints in the modified Stoke ankylosing spondylitis spine score is of additional value in the evaluation of spinal radiographic outcome in ankylosing spondylitis. Arthritis Res Ther 2017;19:77.

66. De Bruin F, de Koning A, van den Berg R, Baraliakos X, Braun J, Ramiro S, et al. Development of the CT Syndesmophyte Score (CTSS) in patients with ankylosing spondylitis: data from the SIAS cohort. Ann Rheum Dis 2018;77:371–7.

67. De Koning A, de Bruin F, van den Berg R, Ramiro S, Baraliakos X, Braun J, et al. Low-dose CT detects more progression of bone formation in comparison to conventional radiography in patients with ankylosing spondylitis: results from the SIAS cohort. Ann Rheum Dis 2018;77:293–9.

68. Tan S, Yao J, Flynn JA, Yao L, Ward MM. Quantitative measurement of syndesmophyte volume and height in ankylosing spondylitis using CT. Ann Rheum Dis 2014;73:544–50.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Negative Regulation of Osteoclast Commitment by Intracellular Protein Phosphatase Magnesium-Dependent 1A

Oh Chan Kwon,[1] Bongkun Choi,[2] Eun-Jin Lee,[2] Ji-Eun Park,[2] Eun-Ju Lee,[2] Eun-Young Kim,[2] Sang-Min Kim,[2] Min-Kyung Shin,[2] Tae-Hwan Kim,[3] (iD) Seokchan Hong,[2] Chang-Keun Lee,[2] Bin Yoo,[2] William H. Robinson,[4] Yong-Gil Kim,[2] (iD) and Eun-Ju Chang[2]

Objective. Increased protein phosphatase magnesium-dependent 1A (PPM1A) levels in patients with ankylosing spondylitis regulate osteoblast differentiation in bony ankylosis; however, the potential mechanisms that regulate osteoclast differentiation in relation to abnormal bone formation remain unclear. This study was undertaken to investigate the relationship of PPM1A to osteoclast differentiation by generating conditional gene-knockout (PPM1A^fl/fl; LysM-Cre) mice and evaluating their bone phenotype.

Methods. The bone phenotypes of LysM-Cre mice (n = 6) and PPM1A^fl/fl;LysM-Cre mice (n = 6) were assessed by micro–computed tomography. Osteoclast differentiation was induced by culturing bone marrow–derived macrophages in the presence of RANKL and macrophage colony-stimulating factor (M-CSF), and was evaluated by counting tartrate-resistant acid phosphatase–positive multinucleated cells. Levels of messenger RNA for *PPM*1A, *RANK*, and osteoclast-specific genes were examined by real-time quantitative polymerase chain reaction, and protein levels were determined by Western blotting. Surface RANK expression was analyzed by fluorescence flow cytometry.

Results. The PPM1A^fl/fl;LysM-Cre mice displayed reduced bone mass ($P < 0.001$) and increased osteoclast differentiation ($P < 0.001$) and osteoclast-specific gene expression ($P < 0.05$) compared with their LysM-Cre littermates. Mechanistically, reduced *PPM*1A function in osteoclast precursors in PPM1A^fl/fl;LysM-Cre mice induced osteoclast lineage commitment by up-regulating RANK expression ($P < 0.01$) via p38 MAPK activation in response to M-CSF. PPM1A expression in macrophages was decreased by Toll-like receptor 4 activation ($P < 0.05$). The Ankylosing Spondylitis Disease Activity Score was negatively correlated with the expression of PPM1A in peripheral blood mononuclear cells from patients with axial spondyloarthritis (SpA) (γ = −0.7072, $P < 0.0001$).

Conclusion. The loss of *PPM*1A function in osteoclast precursors driven by inflammatory signals contributes to osteoclast lineage commitment and differentiation by elevating RANK expression, reflecting a potential role of PPM1A in dynamic bone metabolism in axial SpA.

INTRODUCTION

Axial spondyloarthritis (SpA) is a chronic inflammatory disorder that primarily affects the axial skeleton, including the spine and sacroiliac joints (1). The characteristic features of axial SpA include new bone formation by the osteoproliferation of osteoblasts, which

can eventually lead to bony ankylosis (1). Blocking the differentiation and activation of osteoblasts prevented radiographic progression in a mouse model of ankylosing spondylitis (AS) (2), which illustrates that osteoblasts represent an important component of bony ankylosis in AS. In addition, recent studies have focused on the divergent ability of osteoclast precursors to differentiate into

osteoclasts during osteoclast commitment and accordingly influence bone-resorbing capacity in AS (3,4). Given that the bone remodeling process is coordinated by bone-resorbing osteoclasts and bone-forming osteoblasts (5), the pathophysiologic mechanisms underlying osteoproliferation in AS must be considered in terms of the coupling activities of osteoblasts and osteoclasts (2,3,6–8).

Bone-resorbing osteoclasts are derived from monocyte/macrophage precursors of the hematopoietic progenitor lineage. Proliferating macrophages serve as osteoclast precursors, which undergo further differentiation into osteoclasts (9). Two factors, namely macrophage colony-stimulating factor (M-CSF) and RANKL, are critical for osteoclast lineage commitment and differentiation (10–12). In the earlier stages of osteoclast differentiation, M-CSF binds to its receptor c-Fms on proliferating monocyte/macrophage precursors, thereby playing a key role in osteoclast lineage commitment prior to osteoclast differentiation (13) by activating transcription factors such as microphthalmia-associated transcription factor and PU.1 (14). Importantly, M-CSF can induce the expression of RANK, a receptor for RANKL, through MAPKs, including p38, JNK, and ERK (15,16). This induction of RANK typifies osteoclast lineage commitment. Following RANKL stimulation, the binding of TRAF6 to RANK results in the activation of MAPKs (17,18) and osteoclast-specific transcription factors, such as NFATc1 and NF-κB, to induce osteoclast-specific genes (10,19–21), leading to osteoclast differentiation (9,22). Therefore, in inflammatory conditions, the lineage commitment of macrophages is closely related to osteoclast activity, resulting in pathologic processes in bone remodeling in many diseases, such as osteoporosis, rheumatoid arthritis (RA), and AS (3,4,23).

Protein phosphatase magnesium-dependent 1A (PPM1A) is a member of the protein phosphatase 2C family of serine/threonine phosphatases (24). The expression of this protein is increased in cells or tissues under certain circumstances (25,26). We previously demonstrated that PPM1A levels are increased in the synovial tissue of patients with AS and that its up-regulation enhances osteoblast differentiation (25). Further, PPM1A expression is increased in macrophages during *Mycobacterium tuberculosis* infection, and this up-regulation plays a key role in the innate immune responses of macrophages (26). Considering that macrophages are the precursors of osteoclasts (9) and that PPM1A is known to inactivate MAPKs by dephosphorylating p38 and JNK (27), which are critical for RANK expression (17,18), PPM1A in macrophages might be a possible regulator of osteoclast differentiation.

In this study, we found that the down-regulation of the osteoclast precursor macrophage-specific *PPM1A* in mice resulted in apparently increased osteoclastogenesis, which occurs through the regulation of RANK expression. An inverse correlation between AS Disease Activity Score (ASDAS) and PPM1A expression level was also observed in peripheral blood mononuclear cells (PBMCs) obtained from patients with axial SpA. These discoveries establish PPM1A as a potential determinant of osteoclast lineage commitment from macrophages.

MATERIALS AND METHODS

Reagents and antibodies. RANKL and M-CSF were obtained from PeproTech. A tartrate-resistant acid phosphatase (TRAP) assay kit, lipopolysaccharide (LPS; from *Escherichia coli* O11:B4), PD98059, SB203580, SP60025, and β-actin antibodies were purchased from Sigma-Aldrich. Lipofectamine 2000 was purchased from Invitrogen. Phycoerythrin (PE)–conjugated anti-mouse CD265 (RANK), protease, and phosphatase inhibitor cocktails were purchased from ThermoFisher Scientific. Antibodies against PPM1A were purchased from Novus Biologicals. Antibodies against phospho-ERK, ERK, phospho-p38, p38, phospho-JNK, and JNK were purchased from Cell Signaling Technology. All antibodies used in this study were polyclonal antibodies raised from rabbits.

Mice and bone mineral density (BMD) measurements and histologic analysis. PPM1A$^{flox/flox}$ (PPM1A$^{fl/fl}$, MGI:4458753, Ppm1a$^{tm1a(EUCOMM)Hmgu}$) mice were purchased from MRC Harwell. To obtain mice with PPM1A-deficient macrophages, PPM1A$^{fl/fl}$ mice were bred with LysM-Cre mice (004781, B6.129P2-Lyz2$^{tm1(cre)Ifo}$/J; Jackson Laboratory). PPM1A$^{fl/fl}$; LysM-Cre mice were born at the expected Mendelian ratios, and they survived and reproduced as well as their wild-type (WT) littermates (PPM1A$^{fl/fl}$ or LysM-Cre mice). All mice in this study were from a pure C57BL/6 background. All animal procedures were approved by the Institutional Animal Care and Use Committee of the Asan Institute for Life Sciences (Seoul, Korea). Distal femurs dissected from the LysM-Cre mice and PPM1A$^{fl/fl}$;LysM-Cre mice (n = 6 per group) were fixed in 4% paraformaldehyde. Bone volume was measured by a micro–computed tomography analysis as previously described (28). Mouse hind limbs were collected for histologic analysis as previously described (28), and tissue sections were stained with hematoxylin and eosin or TRAP using an Acid Phosphatase Assay Kit (Sigma-Aldrich) according to the manufacturer's instructions. TRAP staining indicates the presence of mature osteoclasts. The osteoclast surface was assessed on TRAP-stained sections using ImageJ densitometry software, version 1.6 (National Institutes of Health).

Osteoclast differentiation. Bone marrow (BM) cells were isolated by flushing the marrow space in the femurs and tibiae collected from 6-week-old C57BL/6, LysM-Cre, PPM1A$^{fl/fl}$, and PPM1A$^{fl/fl}$;LysM-Cre mice. Mature osteoclasts were generated from BM-derived macrophages (BMMs) and were evaluated by TRAP staining as previously described (28,29). Briefly, isolated BMMs were cultured for 4 days with M-CSF (30 ng/ml) and RANKL (100 ng/ml) to induce differentiation into mature osteoclasts.

Reverse transcription–polymerase chain reaction (RT-PCR) and real-time quantitative PCR (qPCR) analysis.

Total RNA was isolated from cells using TRIzol reagent (Life Technology), and 0.5–1 μg of RNA was reverse-transcribed using SuperScript II reverse transcriptase (Life Technologies). The resulting complementary DNA (cDNA) was amplified by PCR using the following primers: for mouse *PPM1A*, forward 5′-ATG-GTG-CAG-ATA-GAA-GCG-GG-3′ and reverse 5′-AGC-CAG-AGA-GCC-ATT-GAC-AC-3′; for mouse *DC-STAMP*, forward 5′-CCA-AGG-AGT-CGT-CCA-TGA-TT-3′ and reverse 5′-GGC-TGC-TTT-GAT-CGT-TTC-TC-3′; for mouse *OC-STAMP*, forward 5′-TTC-TCT-GGC-CTG-GAG-TTC-CT-3′ and reverse 5′-TGA-CAA-CTT-AGG-CTG-GGC-TG-3′; for mouse *CTSK*, forward 5′-AAT-ACC-TCC-CTC-TCG-ATC-CTA-CA-3′ and reverse 5′-GGT-TCT-TGA-CTG-GAG-TAA-CGT-A-3′; for mouse *TRAP*, forward 5′-TCC-TGG-CTC-AAA-AAG-CAG-TT-3′ and reverse 5′-ACA-TAG-CCC-ACA-CCG-TTC-TC-3′; for mouse *PU.1*, forward 5′-GAT-GGA-GAA-GCT-GAT-GGC-TTG-G-3′ and reverse 5′-TTC-TTC-ACC-TCG-CCT-GTC-TG-C-3′; for mouse *NFATc1*, forward 5′-GGG-TCA-GTG-TGA-CCG-AAG-AT-3′ and reverse 5′-GGA-AGT-CAG-AAG-TGG-GTG-GA-3′; for mouse *c-fms*, forward 5′-CCC-ACC-CTG-AAG-TCC-TGA-GT-3′ and reverse 5′-CTT-TGT-CCT-AGG-GAG-ACG-GC-3′; for mouse *RANK*, forward 5′-CAG-ATG-TCT-TTT-CGT-CCA-CAG-A-3′ and reverse 5′-AGA-CTG-GGC-AGG-TAA-GCC-T-3′; for mouse *GAPDH*, forward 5′-AGC-CAC-ATC-GCT-CAG-ACA-3′ and reverse 5′-GCC-CAA-TAC-GAC-CAA-ATC-C-3′; for human *PPM1A*, forward 5′-TGG-CGT-GTT-GAA-ATG-GAG-3′ and reverse 5′-AGC-GGA-TTA-CTT-GGT-TTG-TG-3′; for human *RANK*, forward 5′-AGA-TCG-CTC-CTC-CAT-GTA-CCA-3′ and reverse 5′-GCC-TTG-CCT-GTA-TCA-CAA-ACT-TT-3′; and for human *GAPDH*,
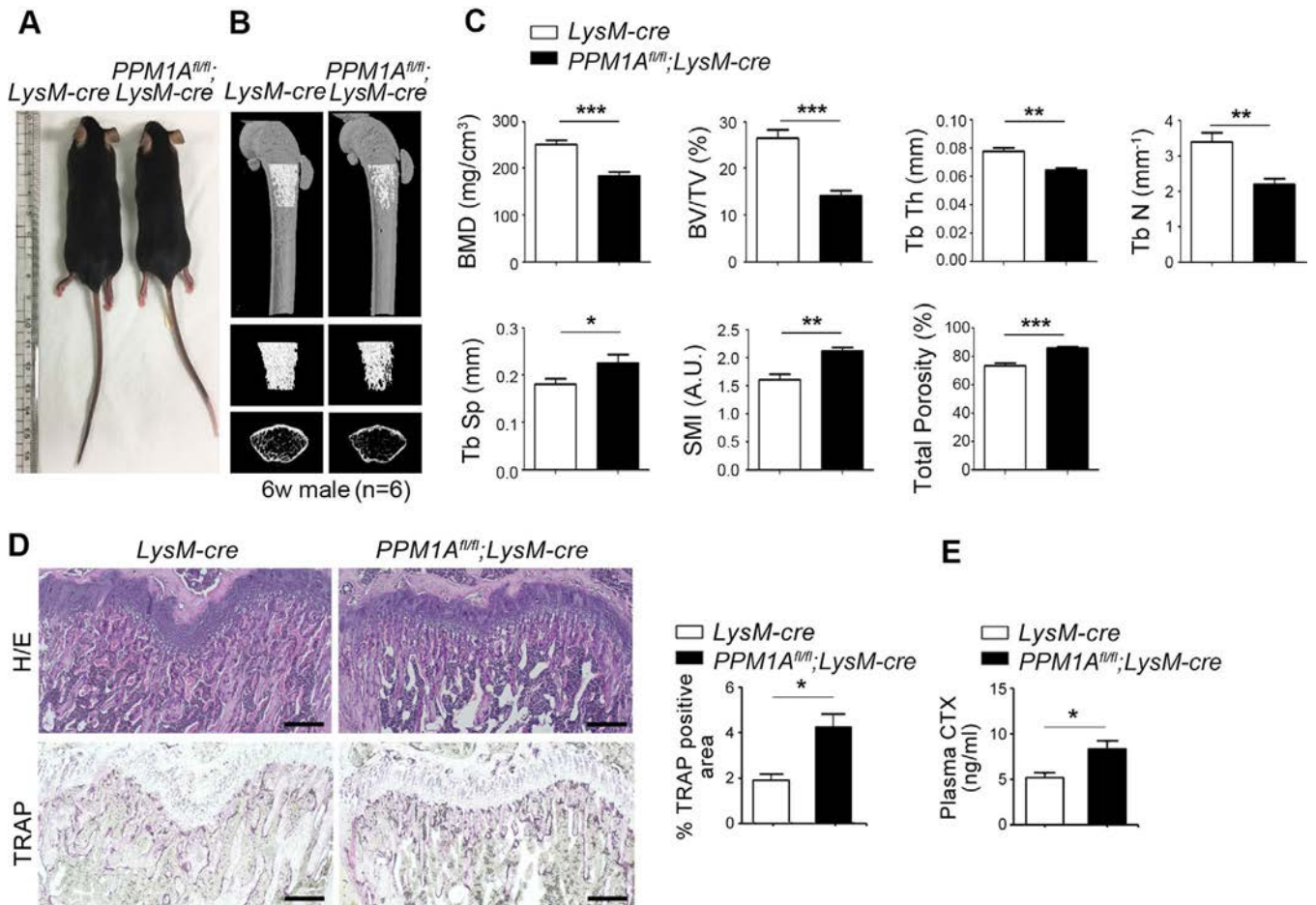
Figure 1. Bone phenotype in LysM-Cre mice and PPM1A$^{fl/fl}$;LysM-Cre mice. **A,** Comparison of the size of LysM-Cre mice and PPM1A$^{fl/fl}$;LysM-Cre mice. **B,** Micro–computed tomography images showing trabecular bone density in the femurs of 6-week-old LysM-Cre and PPM1A$^{fl/fl}$;LysM-Cre mice. **C,** Quantification of the features shown in **B**. BMD = bone mineral density; BV/TV = bone volume/total volume; TbTh = trabecular thickness; TbN = trabecular number; TbSp = trabecular separation; SMI = structure model index. **D,** Left, Hematoxylin and eosin (H&E) and tartrate-resistant acid phosphatase (TRAP) staining of hind limb sections from LysM-Cre and PPM1A$^{fl/fl}$;LysM-Cre mice. Mouse hind limbs were dissected, fixed, and decalcified, and the sections with the trabecular region were stained with H&E or TRAP (purple). Representative images from 3 independent experiments are shown. Bars = 200 μm. Right, Quantitation of the TRAP-positive surface area in sections from each mouse strain. **E,** Levels of C-terminal crosslinking telopeptide of type I collagen (CTX), a marker of bone turnover, in plasma from 6-week-old male LysM-Cre and PPM1Afl/fl;LysM-Cre mice, measured by enzyme-linked immunosorbent assay. Values in **C**, the right panel of **D**, and **E** are the mean ± SD of triplicate determinations (n = 6 mice per group). * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$, by Mann-Whitney U test.

forward 5′-TGT-TGC-CAT-CAA-TGA-CCC-CTT-3′ and reverse 5′-CTC-CAC-GAC-GTA-CTC-AGC-G-3′. The PCR conditions and detailed procedure have been described previously (28). Real-time qPCR was performed using a Power SYBR Green 1-Step Kit and an ABI 7000 Real-Time PCR System (Applied Biosystems) according to the manufacturer's instructions.

Western blotting and fluorescence flow cytometry. Preparation of the cell lysates and sodium dodecyl sulfate–polyacrylamide gel electrophoresis gels and Western blot analyses were conducted according to a standard protocol (28). RANK expression on BMMs isolated from LysM-Cre and PPM1A$^{fl/fl}$; LysM-Cre mice was evaluated by incubating 5×10^5 cells with PE-conjugated anti-mouse RANK or an isotype control antibody in phosphate buffered saline (PBS) containing 2% fetal bovine serum at 4°C for 1 hour. The cells were washed twice with PBS. Flow cytometry was performed on a FACScan instrument according to the manufacturer's instructions (Becton Dickinson).

Reporter assay. The RANK promoter–luciferase reporter plasmid was transiently transfected into BMMs from WT and TLR-4–knockout mice using Lipofectamine 2000 according to the manufacturer's instructions. Two days after transfection, the cells were lysed using passive lysis buffer (Promega), and the luciferase activity in the extracts was measured using a Dual Luciferase assay system (Promega). Co-transfection with the *Renilla* vector allowed normalization of the assays for differences in transfection efficiency.

Human samples. PBMCs were collected from patients with axial SpA (n = 30) and age- and sex-matched healthy controls (n = 13) at the Asan Medical Center and Hanyang University Hospital. Clinical information was extracted from an electronic clinical database. All patients met the Assessment of SpondyloArthritis International Society classification criteria for axial SpA (30). Disease activity was determined using the ASDAS using the C-reactive protein level (ASDAS-CRP) (31). This study was approved by the institutional review boards of Asan Medical Center (IRB No. 2015-0274) and Hanyang University Hospital (IRB No. 2017-12-001).

Enzyme-linked immunosorbent assay (ELISA). The concentrations of C-terminal crosslinking telopeptide of type I collagen (CTX) (Mouse CTX-1 ELISA kit; Novus) in the plasma of LysM-Cre and PPM1Afl/fl;LysM-Cre mice were measured according to the manufacturer's protocols. All samples were examined in triplicate for each experiment.

Statistical analysis. Differences between the 2 groups were analyzed using the Mann-Whitney U test or Student's unpaired *t*-test, and the differences among 3 groups were analyzed by one-way analysis of variance. In the figures, bars are triplicate averages from single experiments, and a representative of 3 independent experiments is shown. The relationships between parameters were tested using Spearman's rank correlation coefficient. *P* values less than 0.05 were considered significant.

RESULTS

Macrophage-specific reduction in *PPM1A* expression in mice results in increased bone resorption due to enhanced osteoclast formation. To evaluate the effect of reduced *PPM1A* expression in macrophages (osteoclast precursors) on bone phenotype, we compared LysM-Cre and PPM1A$^{fl/fl}$; LysM-Cre mice at 6 weeks of age. Compared with LysM-Cre mice, the PPM1A$^{fl/fl}$;LysM-Cre mice were smaller (Figure 1A). CT revealed sparse trabecular bone density in PPM1A$^{fl/fl}$;LysM-Cre mice compared with LysM-Cre mice (Figure 1B). Accordingly, BMD, bone volume/total volume, trabecular thickness, and trabecular number were significantly lower and trabecular spacing and the structure model index were significantly greater in PPM1A$^{fl/fl}$; LysM-Cre mice, suggesting that bone resorption appears as a bone phenotype in PPM1A$^{fl/fl}$;LysM-Cre mice (Figure 1C). TRAP staining revealed enhanced osteoclast activity in PPM1A$^{fl/fl}$;LysM-Cre mice, as evidenced by the increased TRAP-positive staining (Figure 1D). Indeed, ELISA revealed that the level of CTX in the plasma of PPM1A$^{fl/fl}$;LysM-Cre mice was significantly higher than that in LysM-Cre mice, indicating that the reduced bone mass in PPM1A$^{fl/fl}$; LysM-Cre mice is caused by increased osteoclast activity (Figure 1E).

We next evaluated the expression status of osteoclast-specific genes (*TRAP*, *DC-STAMP*, *OC-STAMP*, and *CTSK*) (10,19–21,32) in macrophages from PPM1A$^{fl/fl}$;LysM-Cre and LysM-Cre mice. RT-PCR revealed lower *PPM1A* expression in macrophages from PPM1A$^{fl/fl}$;LysM-Cre mice than C57BL/6, LysM-Cre, and PPM1A$^{fl/fl}$ mice, as expected (Figure 2A). Consistent with this finding, PPM1A protein expression in macrophages was similarly decreased in PPM1A$^{fl/fl}$;LysM-Cre mice (Figure 2B). TRAP-positive mononuclear cells, which are osteoclast-specific lineage cells (33), were numerically increased in PPM1A$^{fl/fl}$;LysM-Cre mice (Figures 2C and D). Real-time qPCR illustrated that BMMs from PPM1A$^{fl/fl}$;LysM-Cre mice displayed higher *DC-STAMP*, *OC-STAMP*, *CTSK*, and *TRAP* expression than those from LysM-Cre mice (Figure 2E). These data suggest that reduced PPM1A expression in macrophages results in increased osteoclast differentiation due to the increased capacity for osteoclast formation in the early stages of osteoclast differentiation.

***PPM1A* down-regulation in macrophages leads to increased expression of RANK via p38 MAPK signaling.** To determine the mechanism by which PPM1A influences osteoclast commitment, we investigated the status of gene expression related to osteoclast lineage commitment (34) in M-CSF–cultured BMMs obtained from LysM-Cre and PPM1A$^{fl/fl}$; LysM-Cre mice. Real-time PCR revealed that *PU.1* messenger
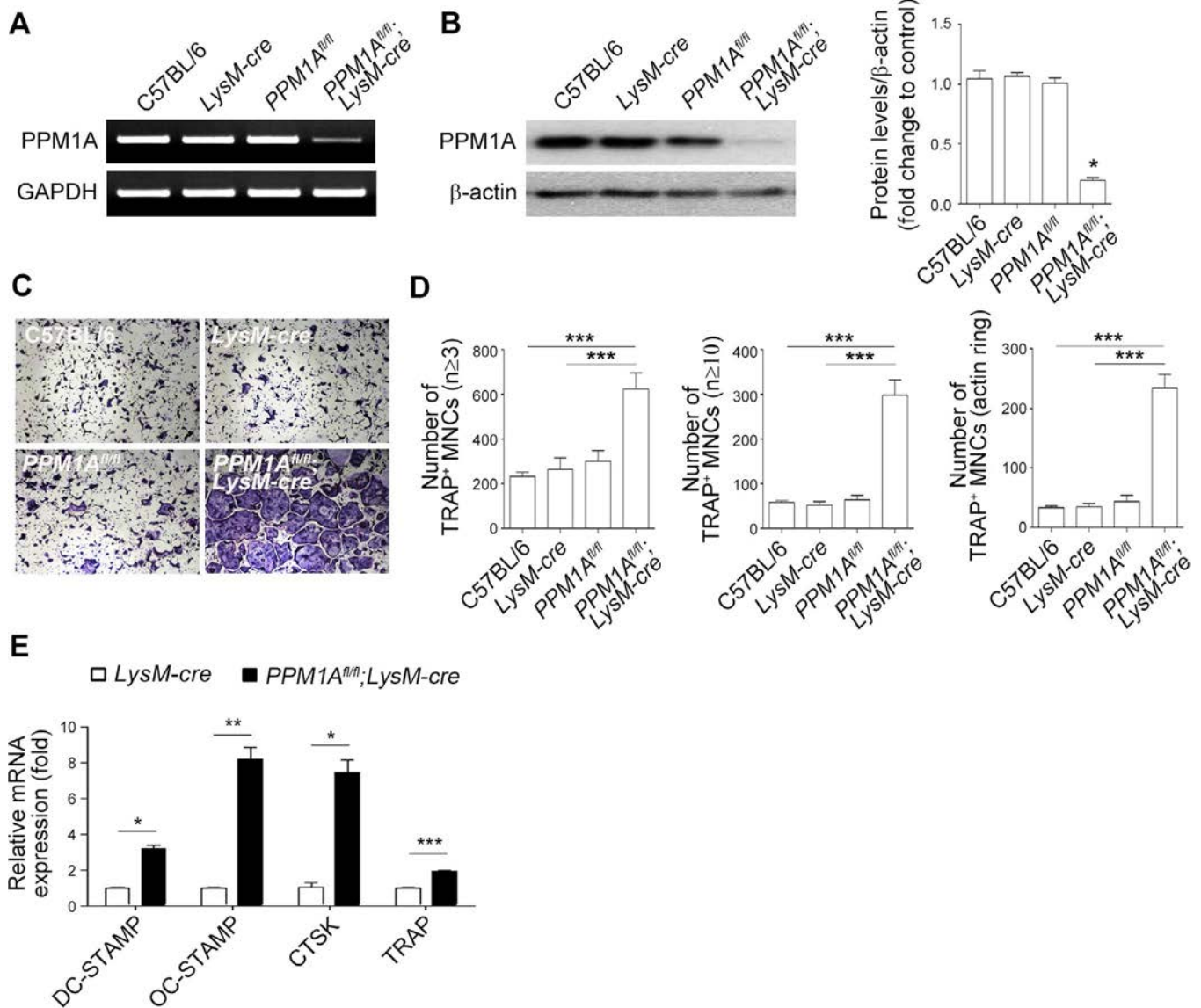
Figure 2. Association of decreased protein phosphatase magnesium-dependent 1A (PPM1A) expression in macrophages with increased expression of osteoclast-specific genes and osteoclast differentiation. **A** and **B**, *PPM1A* mRNA expression, determined by reverse transcription–polymerase chain reaction (PCR) (**A**), and PPM1A protein levels, determined by Western blotting (**B**), in macrophages from C57BL/6, LysM-Cre, PPM1A$^{fl/fl}$, and PPM1A$^{fl/fl}$;LysM-Cre mice. The right panel of **B** shows the densitometric quantification of PPM1A compared to β-actin. GAPDH was used as a loading control in **A**; β-actin was used as a loading control in **B**. **C**, Tartrate-resistant acid phosphatase (TRAP) staining for osteoclast formation. Macrophages from C57BL/6, LysM-Cre, PPM1A$^{fl/fl}$, and PPM1A$^{fl/fl}$;LysM-Cre mice were treated with macrophage colony-stimulating factor (30 ng/ml) and RANKL (100 ng/ml) for 4 days, fixed, and stained with TRAP. Original magnification × 200. **D**, Numbers of TRAP-positive multinucleated cells (MNCs) containing ≥3 nuclei, ≥10 nuclei, or an actin ring in **C**, counted under a light microscope. **E**, Relative expression of mRNA for the indicated osteoclast-specific genes in bone marrow–derived macrophages from LysM-Cre and PPM1A$^{fl/fl}$;LysM-Cre mice, determined by real-time quantitative PCR. The transcript levels were normalized to GAPDH. Values in the right panel of **B**, **D**, and **E** are the mean ± SD of triplicate determinations. * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$, by Mann-Whitney U test. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41180/abstract.

RNA (mRNA) expression was increased in BMMs from PPM1A$^{fl/fl}$;LysM-Cre mice with no alteration in *c-fms* expression, suggesting that the osteoclast lineage commitment was enhanced by *PPM1A* down-regulation. Interestingly, *RANK* mRNA expression was increased in BMMs from PPM1A$^{fl/fl}$;LysM-Cre mice (Figure 3A). RANK protein expression was

increased in macrophages from PPM1A$^{fl/fl}$;LysM-Cre mice compared with those from LysM-Cre mice (Figure 3B). Time-course Western blot analysis revealed no major changes in the activation of ERK or JNK upon M-CSF stimulation for 0–30 minutes compared with the LysM-Cre control (Figure 3C). Only p38 activation was increased in a time-dependent manner
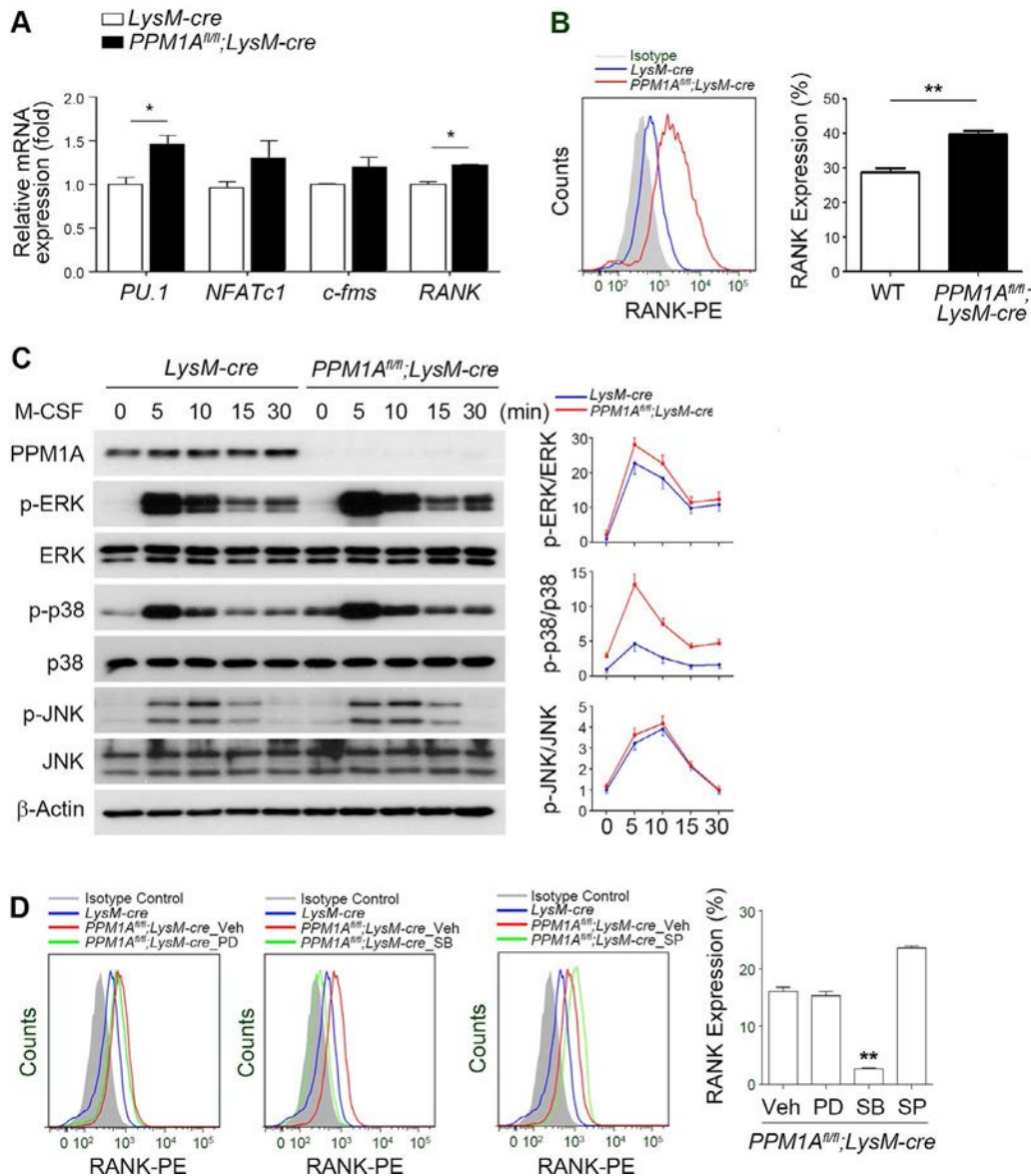
Figure 3. Increased expression of RANK in macrophages from PPM1A^fl/fl^;LysM-Cre mice. **A**, Levels of expression of mRNA for genes involved in macrophage colony-stimulating factor (M-CSF) signaling and *RANK* in LysM-Cre (wild-type [WT]) and PPM1A^fl/fl^;LysM-Cre mouse macrophages, determined by real-time quantitative polymerase chain reaction. **B**, Left, RANK expression in bone marrow–derived macrophages (BMMs) from LysM-Cre and PPM1A^fl/fl^;LysM-Cre mice, analyzed by fluorescence flow cytometry. The shaded histogram indicates the isotype control. Right, Quantification of the percentage of phycoerythrin (PE)–conjugated RANK–positive cells. **C**, Left, Western blot analysis of M-CSF–induced activation of MAPKs (phospho-ERK, ERK, phospho-p38, p38, phospho-JNK, and JNK) in BMMs from LysM-Cre and PPM1A^fl/fl^;LysM-Cre mice after stimulation with M-CSF for 0–30 minutes. Right, Quantification of the Western blot analysis results. **D**, Left, RANK expression in BMMs from LysM-Cre mice and PPM1A^fl/fl^;LysM-Cre mice pretreated with DMSO (vehicle [veh]) and treated with an ERK inhibitor (PD98059 [PD]; 10 mM), a p38 inhibitor (SB203580 [SB]; 10 mM), or a JNK inhibitor (SP60025 [SP];10 mM) in the presence of M-CSF. Cells were analyzed for RANK expression by fluorescence flow cytometry. Right, Percentage of PE-conjugated RANK–positive cells. Values in **A** and the right panels of **B** and **D** are the mean ± SD of triplicate determinations. * = $P < 0.05$; ** = $P < 0.01$ by Mann-Whitney U test. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41180/abstract.

in PPM1A^fl/fl^;LysM-Cre mouse macrophages compared with LysM-Cre mouse macrophages (Figure 3C).

To gain more direct evidence that PPM1A regulates *RANK* through p38 MAPK signaling, we treated BMMs with various MAPK inhibitors and analyzed RANK expression by flow cytometry. Consistent with the up-regulation of *RANK*, there was a sig-

nificant difference in RANK levels in PPM1A^fl/fl^;LysM-Cre mouse macrophages due to the activation of p38 MAPK (Figure 3D). These data indicate that PPM1A influences *RANK* expression through the p38 signaling pathway and that it may directly dephosphorylate p38 MAPK, as previously reported (35). Taken together, these findings show that PPM1A primarily regulates

RANK expression in mouse macrophages through the p38 signaling pathway.

LPS reduces PPM1A expression, resulting in increased RANK expression in macrophages. We next explored whether the inflammatory environment affects PPM1A expression. LPS, tumor necrosis factor, interleukin-1β (IL-1β), and IL-6 were used as the inflammatory stimuli. *PPM1A* mRNA and protein expression were diminished in macrophages from wild-type mice following LPS stimulation (Figures 4A and 4B). Because Toll-like receptor 4 (TLR-4) is the receptor for

LPS (36–38), we then investigated whether PPM1A is down-regulated in response to LPS exposure using TLR-4–knockout mice. Whereas LPS stimulation resulted in decreased *PPM1A* mRNA and protein expression in WT mouse macrophages, these effects were not observed in macrophages from TLR-4–knockout mice (Figures 4C and 4D). This regulatory axis affects RANK expression in macrophages, as evidenced by the finding that LPS stimulation increased *RANK* promoter activation (Figure 4E) and RANK expression (Figure 4F) in WT mouse macrophages but not in TLR-4–knockout mouse macrophages. Similarly, in human PBMCs, TLR-4 activation by LPS
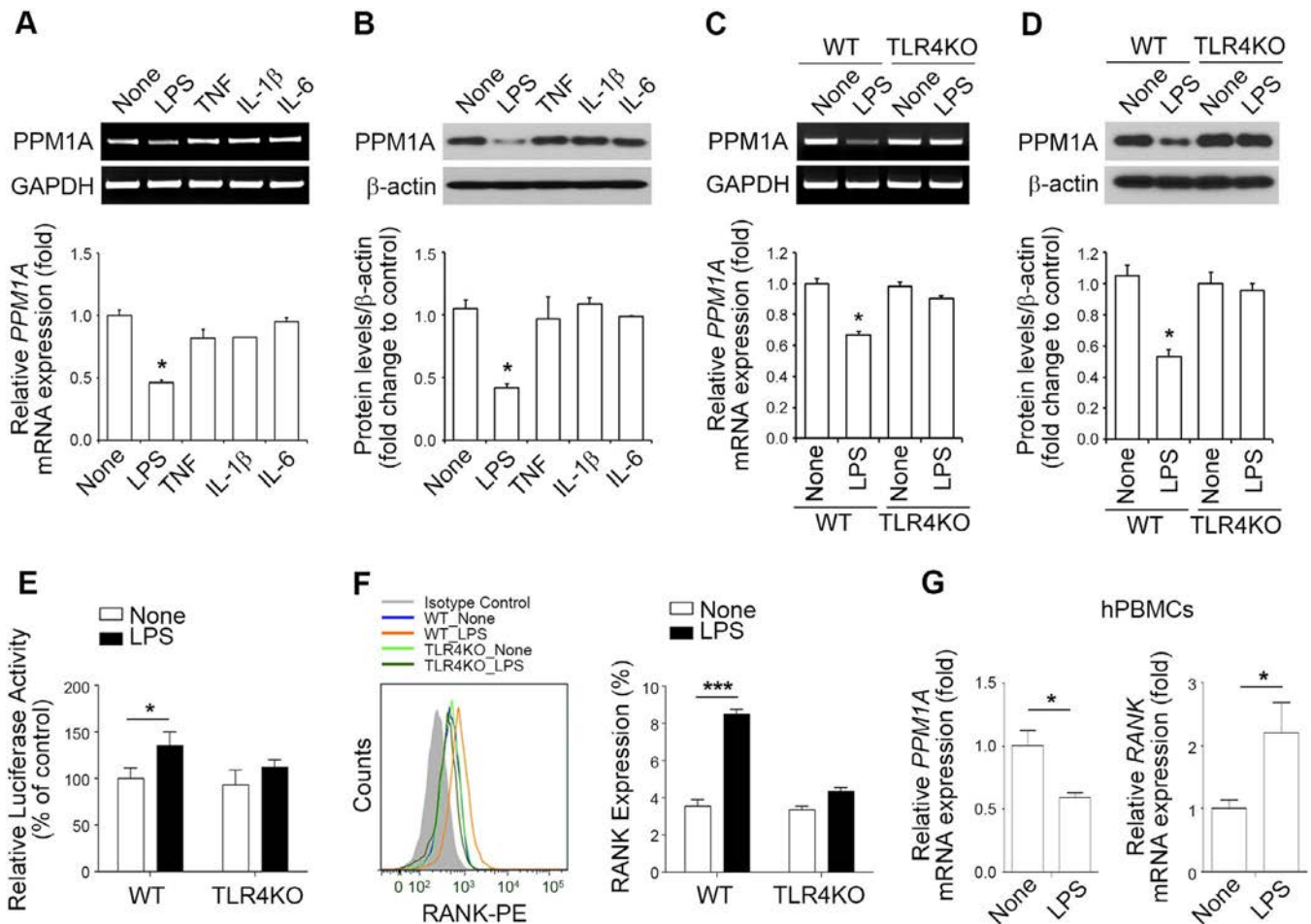
Figure 4. Reduced protein phosphatase magnesium-dependent 1A (PPM1A) expression and increased RANK expression in macrophages from wild-type mice stimulated with lipopolysaccharide (LPS). **A**, *PPM1A* mRNA expression in macrophages exposed to inflammatory stimuli including LPS, tumor necrosis factor (TNF), interleukin-1β (IL-1β), and IL-6, determined by reverse transcription–polymerase chain reaction (RT-PCR) (top) and quantitative RT-PCR (qRT-PCR) (bottom). **B**, PPM1A protein expression in macrophages treated with LPS, TNF, IL-1β, and IL-6, determined by Western blot analysis (top) and quantification of protein expression (bottom). **C**, *PPM1A* mRNA expression in macrophages from wild-type (WT) and TL-R4–knockout (TLR4KO) mice exposed to LPS, determined by RT-PCR (top) and qRT-PCR (bottom). **D**, PPM1A protein expression level in macrophages from WT and TLR-4–knockout mice exposed to LPS, determined by Western blotting (top), and quantification of protein expression (bottom). **E**, Relative luciferase activity in WT and TLR-4–knockout mice. Bone marrow–derived macrophages (BMMs) from WT and TLR-4–knockout mice were treated with LPS, and RANK promoter–luciferase reporter plasmids were transiently transfected into these cells. After 24 hours, the cells were harvested and subjected to the luciferase assay. Relative luciferase activity was normalized to control activity. **F**, Surface RANK protein level in WT and TLR-4–knockout mouse BMMs exposed to LPS, determined by flow cytometry. **G**, *PPM1A* and *RANK* mRNA expression levels in human peripheral blood mononuclear cells (hPBMCs) isolated from healthy controls and stimulated with LPS, determined by qRT-PCR. Values are the mean ± SD of triplicate determinations. * = P < 0.05; *** = P < 0.001, by Mann-Whitney U test. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41180/abstract.

stimulation was linked to decreased *PPM1A* mRNA expression, and this decrease was accompanied by an increase in *RANK* mRNA expression (Figure 4G).

PPM1A expression in PBMCs from patients with axial SpA. Because our results revealed a regulatory axis between PPM1A expression and the inflammatory condition,
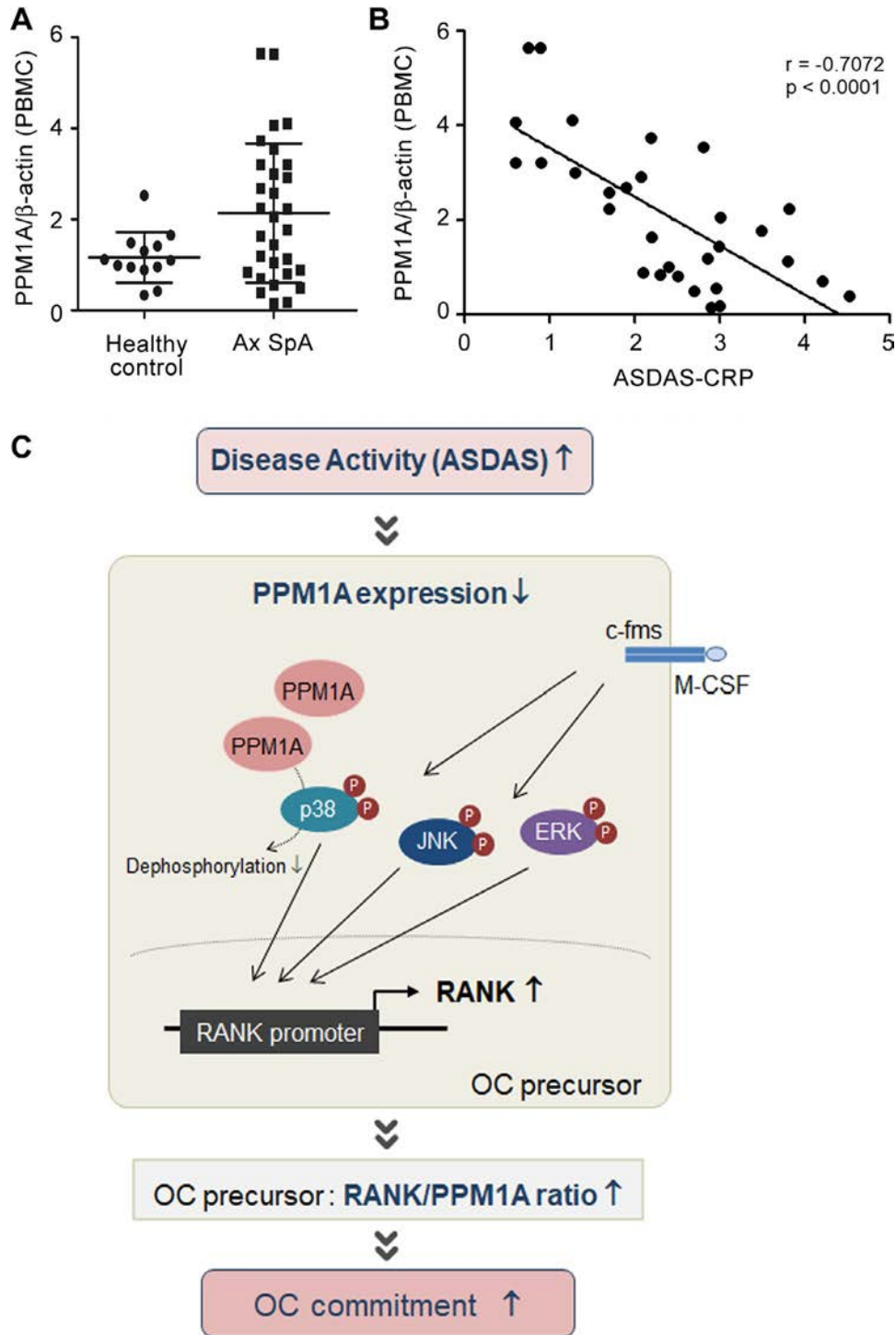
Figure 5. Lower expression levels of protein phosphatase magnesium-dependent 1A (PPM1A) in peripheral blood mononuclear cells (PBMCs) from patients with axial spondyloarthritis (Ax SpA) with higher disease activity. **A**, Protein levels of PPM1A in PBMCs from patients with axial SpA (n = 30) and age- and sex-matched healthy controls (n = 13), determined by immunoblot assay. Symbols represent individual subjects; horizontal lines and error bars show the mean ± SD. **B**, Correlation between PPM1A levels and Ankylosing Spondylitis Disease Activity Score using the C-reactive protein level (ASDAS-CRP) in patients with axial SpA (n = 30), determined by Spearman's correlation analysis. **C**, Suggested working model in axial SpA. M-CSF = macrophage colony-stimulating factor; OC = osteoclast. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41180/abstract.

Table 1. Characteristics of the 30 patients with axial spondyloarthritis*

Age, median (IQR) years	33.0 (28.5–42.0)
Male, no. (%)	28 (93.3)
Disease duration, median (IQR) months	19.2 (1.5–111.7)
HLA–B27 positive, no. (%)	29 (96.7)
Ankylosing spondylitis, no. (%)†	27 (90.0)
ASDAS-CRP, mean ± SD	2.35 ± 1.06

* IQR = interquartile range; ASDAS-CRP = Ankylosing Spondylitis Disease Activity Score using the C-reactive protein level.
† Patients who fulfilled the radiologic criterion of the 1984 modified New York criteria (sacroiliitis grade ≥2 bilaterally or grade 3–4 unilaterally).

it was important to determine whether there is a correlation between PPM1A expression and disease activity in axial SpA (7). We compared the PPM1A:β-actin ratio in PBMCs from patients with axial SpA (n = 30) and age- and sex-matched healthy controls (n = 13) and found that there was no significant difference (Figure 5A). Next, we evaluated whether the inflammatory burden, as measured by ASDAS-CRP, is correlated with PPM1A expression in PBMCs from patients with axial SpA (Figure 5B). Clinical variables for the patients with axial SpA are shown in Table 1. Of the 30 patients with axial SpA, 27 (90.0%) fulfilled the 1984 modified New York criteria for AS (39). In the correlation analysis, PPM1A expression in PBMCs and the ASDAS-CRP were negatively correlated (r = −0.7072, P < 0.0001), emphasizing that the inflammatory burden of axial SpA is associated with decreased PPM1A expression.

DISCUSSION

In this study, we demonstrated that the macrophage-specific down-regulation of *PPM1A* results in osteoclast commitment via increased RANK expression and enhanced RANK signaling. To our knowledge, this is the first study to identify the role of PPM1A in macrophages in osteoclast differentiation. We previously reported that PPM1A levels are increased in the synovial tissue of patients with AS and that PPM1A overexpression promotes osteoblast differentiation (25). The serum levels of PPM1A in patients with AS were also increased compared with those in patients with RA and in healthy controls (25). Although the serum PPM1A levels varied among patients with AS, the clinical significance of this variability has not been determined. In the present study, a greater range of PPM1A expression was also observed in PBMCs from patients with axial SpA. Those with higher expression of PPM1A in PBMCs may attenuate RANK expression more potently, resulting in the inhibition of osteoclast commitment and potential changes in the joint microstructure.

Osteoclasts are derived from hematopoietic stem cells (HSCs), and they are responsible for the resorption of endosteal bone surfaces and periosteal surfaces beneath the periosteum (10,40). RANK–RANKL interaction is the primary factor involved in osteoclast differentiation (9,22). RANKL binds to RANK expressed on the surface of osteoclast precursors and initiates downstream signaling (RANK signaling), which leads to

the expression of osteoclast-specific genes and consequently the differentiation and activation of mature osteoclasts (9,22). Thus, the RANK level in osteoclast precursors can determine the capacity for osteoclast formation through osteoclast lineage commitment. We demonstrated that expression of mRNA for *RANK* and *PU.1* was increased in PPM1A^fl/fl;LysM-Cre mouse macrophages compared with LysM-Cre mouse macrophages. Considering that HSCs differentiate into osteoclast precursors in the presence of PU.1 under M-CSF signaling (33), we can conclude that macrophages from PPM1A^fl/fl;LysM-Cre mice display enforced osteoclast commitment due to M-CSF signaling. Notably, M-CSF, by binding to c-fms, autophosphorylates cytoplasmic tail tyrosine residues (41) and activates downstream events including p38 phosphorylation (42,43), resulting in increased RANK expression in early osteoclast precursors (44). Based on our results, we speculate that in circumstances in which PPM1A expression is decreased in macrophages, p38 activity increases, resulting in increased RANK expression and thereby enhanced osteoclast commitment.

The osteoclast-specific genes that are induced by RANK-mediated intracellular signaling include *CTSK*, *TRAP*, *CALCR*, *DC-STAMP*, *OC-STAMP*, and β3 integrin (10,19–21,32). In our study, the expression of mRNA for *CTSK*, *TRAP*, *DC-STAMP*, and *OC-STAMP* was increased in PPM1A^fl/fl;LysM-Cre mouse macrophages compared with LysM-Cre mouse macrophages, indicating that PPM1A down-regulates osteoclast-specific genes in these cells. Given that PPM1A inactivates MAPKs (27) and that MAPKs are important mediators of RANK-mediated intracellular signaling (17,18), PPM1A may attenuate further osteoclast differentiation by down-regulating RANK-mediated intracellular signaling stimulated by RANKL in osteoclast precursors. In particular, we found that LPS stimulation reduced PPM1A expression in macrophages, thus identifying inflammation as an important variable affecting *PPM1A* mRNA expression. Thus, in inflammatory conditions, osteoclast differentiation may be enhanced by increased RANK expression signaling attributable to PPM1A down-regulation in macrophages. However, the individual LPS-regulated inflammatory cytokines failed to suppress PPM1A. The reasons are not yet apparent, but we speculate that there could be additional mediators that stimulate TLR-4 signaling, such as pathogen-associated molecular patterns, rather than cytokine signaling that affects PP1MA.

In axial SpA, the severity of joint inflammation tends to fluctuate over time (45). Data in Figure 5A show that there was no significant difference in PPM1A expression in PBMCs from patients with axial SpA versus healthy controls. Although the explanation for this observation is unclear, the variability of axial SpA disease activity might have contributed to the greater range of PPM1A expression in PBMCs from patients with axial SpA. A strong negative correlation was observed between ASDAS-CRP and PPM1A expression in PMBCs, supporting the notion that the inflammatory burden is a possible regulator

of PPM1A expression. Interestingly, elevated serum levels of soluble RANKL and increased bone resorption as assessed by decreased BMD have also been reported in patients with AS (46). In that study, prominent elevation of soluble RANKL in patients with AS might have led to osteoclast commitment more actively in the presence of strong RANK expression in macrophages, thereby resulting in a resorptive bone phenotype. The notion that osteoclasts play a role in the pathogenesis of AS is supported by the clinical benefits resulting from treatment with pamidronate in active AS (47). Taken together, those findings and our results indicate that the PPM1A level may determine the resorptive bone phenotype in active axial SpA under an inflammatory burden by altering the capacity for osteoclast commitment in macrophages.

In conclusion, we demonstrated that PPM1A down-regulation in macrophages results in RANK up-regulation and RANK signaling enhancement, causing osteoclast commitment and further bone resorption. This finding suggests that PPM1A is both a potential enhancer of osteoblastogenesis (25) and a potential regulator of osteoclast commitment. Thus, in axial SpA with active inflammation, decreased PPM1A expression in PBMCs may enhance osteoclastogenesis via the up-regulation of RANK, thereby shifting the homeostasis of bone metabolism toward bone resorption (Figure 5B). These findings identify PPM1A as an important marker of bone metabolism in axial SpA and a potential therapeutic target for treating bony ankylosis.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Y.-G. Kim 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. Robinson, Y.-G. Kim, Chang.
Acquisition of data. Kwon, Choi, Eun-Jin Lee, T.-H. Kim, Hong, C.-K. Lee, Yoo, Y.-G. Kim, Chang.
Analysis and interpretation of data. Kwon, Choi, Eun-Jin Lee, Park, Eun-Ju Lee, E.-Y. Kim, S.-M. Kim, Shin, Y.-G. Kim, Chang.

REFERENCES

1. Taurog JD, Chhabra A, Colbert RA. Ankylosing spondylitis and axial spondyloarthritis [review]. N Engl J Med 2016;374:2563–74.

2. Lories RJ, Derese I, Luyten FP. Modulation of bone morphogenetic protein signaling inhibits the onset and progression of ankylosing enthesitis. J Clin Invest 2005;115:1571–9.

3. Perpétuo IP, Caetano-Lopes J, Vieira-Sousa E, Campanilho-Marques R, Ponte C, Canhão H, et al. Ankylosing spondylitis patients have impaired osteoclast gene expression in circulating osteoclast precursors. Front Med (Lausanne) 2017;4:5.

4. Perpétuo IP, Raposeiro R, Caetano-Lopes J, Vieira-Sousa E, Campanilho-Marques R, Ponte C, et al. Effect of tumor necrosis factor inhibitor therapy on osteoclasts precursors in ankylosing spondylitis. PLoS One 2015;10:e0144655.

5. Khosla S. Pathogenesis of age-related bone loss in humans. J Gerontol A Biol Sci Med Sci 2013;68:1226–35.

6. Lories R. The balance of tissue repair and remodeling in chronic arthritis [review]. Nat Rev Rheumatol 2011;7:700–7.

7. Zhang X, Aubin JE, Inman RD. Molecular and cellular biology of new bone formation: insights into the ankylosis of ankylosing spondylitis. Curr Opin Rheumatol 2003;15:387–93.

8. Tam LS, Gu J, Yu D. Pathogenesis of ankylosing spondylitis [review]. Nat Rev Rheumatol 2010;6:399–405.

9. Boyle WJ, Simonet WS, Lacey DL. Osteoclast differentiation and activation [review]. Nature 2003;423:337–42.

10. Arai F, Miyamoto T, Ohneda O, Inada T, Sudo T, Brasel K, et al. Commitment and differentiation of osteoclast precursor cells by the sequential expression of c-Fms and receptor activator of nuclear factor κB (RANK) receptors. J Exp Med 1999;190:1741–54.

11. Cappellen D, Luong-Nguyen NH, Bongiovanni S, Grenet O, Wanke C, Šuša M. Transcriptional program of mouse osteoclast differentiation governed by the macrophage colony-stimulating factor and the ligand for the receptor activator of NFκB. J Biol Chem 2002;277:21971–82.

12. Dougall WC, Glaccum M, Charrier K, Rohrbach K, Brasel K, De Smedt T, et al. RANK is essential for osteoclast and lymph node development. Genes Dev 1999;13:2412–24.

13. Tanaka S, Takahashi N, Udagawa N, Tamura T, Akatsu T, Stanley ER, et al. Macrophage colony-stimulating factor is indispensable for both proliferation and differentiation of osteoclast progenitors. J Clin Invest 1993;91:257–63.

14. Nakanishi A, Hie M, Iitsuka N, Tsukamoto I. A crucial role for reactive oxygen species in macrophage colony-stimulating factor-induced RANK expression in osteoclastic differentiation. Int J Mol Med 2013;31:874–80.

15. Lee K, Chung YH, Ahn H, Kim H, Rho J, Jeong D. Selective regulation of MAPK signaling mediates RANKL-dependent osteoclast differentiation. Int J Biol Sci 2016;12:235–45.

16. Sobacchi C, Frattini A, Guerrini MM, Abinun M, Pangrazio A, Susani L, et al. Osteoclast-poor human osteopetrosis due to mutations in the gene encoding RANKL. Nat Genet 2007;39:960–2.

17. Li X, Udagawa N, Itoh K, Suda K, Murase Y, Nishihara T, et al. P38 MAPK-mediated signals are required for inducing osteoclast differentiation but not for osteoclast function. Endocrinology 2002;143:3105–13.

18. Matsumoto M, Sudo T, Saito T, Osada H, Tsujimoto M. Involvement of p38 mitogen-activated protein kinase signaling pathway in osteoclastogenesis mediated by receptor activator of NF-κ B ligand (RANKL). J Biol Chem 2000;275:31155–61.

19. Wada T, Nakashima T, Hiroshi N, Penninger JM. RANKL-RANK signaling in osteoclastogenesis and bone disease. Trends Mol Med 2006;12:17–25.

20. Boyce BF. Advances in osteoclast biology reveal potential new drug targets and new roles for osteoclasts [review]. J Bone Miner Res 2013;28:711–22.

21. Edwards JR, Mundy GR. Advances in osteoclast biology: old findings and new insights from mouse models [review]. Nat Rev Rheumatol 2011;7:235–43.

22. Suda T, Takahashi N, Udagawa N, Jimi E, Gillespie MT, Martin TJ. Modulation of osteoclast differentiation and function by the new members of the tumor necrosis factor receptor and ligand families [review]. Endocr Rev 1999;20:345–57.

23. Takayanagi H. Osteoimmunology: shared mechanisms and cross-talk between the immune and bone systems [review]. Nat Rev Immunol 2007;7:292–304.

24. Zolnierowicz S. Type 2A protein phosphatase, the complex regulator of numerous signaling pathways. Biochem Pharmacol 2000;60:1225–35.

25. Kim YG, Sohn DH, Zhao X, Sokolove J, Lindstrom TM, Yoo B, et al. Role of protein phosphatase magnesium-dependent 1A and

anti–protein phosphatase magnesium-dependent 1A autoantibodies in ankylosing spondylitis. Arthritis Rheumatol 2014;66:2793–803.

26. Sun J, Schaaf K, Duverger A, Wolschendorf F, Speer A, Wagner F, et al. Protein phosphatase, Mg2+/Mn2+-dependent 1A controls the innate antiviral and antibacterial response of macrophages during HIV-1 and Mycobacterium tuberculosis infection. Oncotarget 2016;7:15394–409.

27. Takekawa M, Maeda T, Saito H. Protein phosphatase 2Cα inhibits the human stress-responsive p38 and JNK MAPK pathways. EMBO J 1998;17:4744–52.

28. Lee EJ, Kim SM, Choi B, Kim EY, Chung YH, Lee EJ, et al. Interleukin-32 γ stimulates bone formation by increasing miR-29a in osteoblastic cells and prevents the development of osteoporosis. Sci Rep 2017;7:40240.

29. Lee EJ, Song DH, Kim YJ, Choi B, Chung YH, Kim SM, et al. PTX3 stimulates osteoclastogenesis by increasing osteoblast RANKL production. J Cell Physiol 2014;229:1744–52.

30. Rudwaleit M, van der Heijde D, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis. Part II. Validation and final selection. Ann Rheum Dis 2009;68:777–83.

31. Lukas C, Landewé 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.

32. Lacey DL, Timms E, Tan HL, Kelley MJ, Dunstan CR, Burgess T, et al. Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation. Cell 1998;93:165–76.

33. Amarasekara DS, Yun H, Kim S, Lee N, Kim H, Rho J. Regulation of osteoclast differentiation by cytokine networks. Immune Netw 2018;18:e8.

34. Boyce BF. Advances in the regulation of osteoclasts and osteoclast functions. J Dent Res 2013;92:860–7.

35. Dvashi Z, Sar Shalom H, Shohat M, Ben-Meir D, Ferber S, Satchi-Fainaro R, et al. Protein phosphatase magnesium dependent 1A governs the wound healing-inflammation-angiogenesis cross talk on injury. Am J Pathol 2014;184:2936–50.

36. Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X, et al. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science 1998;282:2085–8.

37. Qureshi ST, Larivière L, Leveque G, Clermont S, Moore KJ, Gros P, et al. Endotoxin-tolerant mice have mutations in Toll-like receptor 4 (Tlr4). J Exp Med 1999;189:615–25.

38. Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, et al. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. J Immunol 1999;162:3749–52.

39. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.

40. Zaidi M. Skeletal remodeling in health and disease. Nat Med 2007; 13:791–801.

41. Feng X, Takeshita S, Namba N, Wei S, Teitelbaum SL, Ross FP. Tyrosines 559 and 807 in the cytoplasmic tail of the macrophage colony-stimulating factor receptor play distinct roles in osteoclast differentiation and function. Endocrinology 2002;143:4868–74.

42. Wang Y, Piper MG, Marsh CB. The role of Src family kinases in mediating M-CSF receptor signaling and monocytic cell survival. Adv Biosci Biotechnol 2012;3:592–602.

43. Wang Y, Zeigler MM, Lam GK, Hunter MG, Eubank TD, Khramtsov VV, et al. The role of the NADPH oxidase complex, p38 MAPK, and Akt in regulating human monocyte/macrophage survival. Am J Respir Cell Mol Biol 2007;36:68–77.

44. Soysa NS, Alles N, Aoki K, Ohya K. Osteoclast formation and differentiation: an overview. J Med Dent Sci 2012;59:65–74.

45. Stone MA, Pomeroy E, Keat A, Sengupta R, Hickey S, Dieppe P, et al. Assessment of the impact of flares in ankylosing spondylitis disease activity using the Flare Illustration. Rheumatology (Oxford) 2008;47:1213–8.

46. Kim HR, Lee SH, Kim HY. Elevated serum levels of soluble receptor activator of nuclear factors-κB ligand (sRANKL) and reduced bone mineral density in patients with ankylosing spondylitis (AS). Rheumatology (Oxford) 2006;45:1197–200.

47. Maksymowych WP, Jhangri GS, Fitzgerald AA, LeClercq S, Chiu P, Yan A, et al. A six-month randomized, controlled, double-blind, dose-response comparison of intravenous pamidronate (60 mg versus 10 mg) in the treatment of nonsteroidal antiinflammatory drug–refractory ankylosing spondylitis. Arthritis Rheum 2002;46:766–73.

Maintenance of Efficacy and Safety of Ustekinumab Through One Year in a Phase II Multicenter, Prospective, Randomized, Double-Blind, Placebo-Controlled Crossover Trial of Patients With Active Systemic Lupus Erythematosus

Ronald F. van Vollenhoven,[1] [ID] Bevra H. Hahn,[2] George C. Tsokos,[3] [ID] Peter Lipsky,[4] Kaiyin Fei,[5] Robert M. Gordon,[5] Irene Gregan,[5] Kim Hung Lo,[5] Marc Chevrier,[5] and Shawn Rose[5]

Objective. To evaluate the efficacy and safety of ustekinumab through 1 year in a phase II trial in patients with systemic lupus erythematosus (SLE).

Methods. Eligible patients were diagnosed as having clinically active SLE (based on Systemic Lupus International Collaborating Clinics criteria), despite standard background therapy. Active disease was defined by an SLE Disease Activity Index 2000 (SLEDAI-2K) score of ≥6 as well as having ≥1 British Isles Lupus Assessment Group (BILAG) A organ domain score and/or ≥2 BILAG B organ domain scores present at screening. Patients (n = 102) were randomized (3:2) to receive either ustekinumab (~6 mg/kg of single intravenous infusion at week 0, then 90-mg subcutaneous injections every 8 weeks beginning at week 8) or a matching placebo added to standard therapy. At week 24, the placebo group crossed over to receive a subcutaneous 90-mg dose of ustekinumab every 8 weeks, and the original ustekinumab group continued to receive therapy through week 40. Maintenance of efficacy was assessed using the SLEDAI-2K, the SLE Responder Index 4 (SRI-4), physician global assessment, and mucocutaneous and joint disease measures in a modified intent-to-treat population.

Results. SRI-4 response rates were significantly greater in the ustekinumab group (62%) versus the placebo group (33%) in the week 24 primary end point analysis (*P* = 0.006) and were maintained at week 48 (63.3%) in the ustekinumab group. In the ustekinumab group, response rates across other disease measures were also maintained through week 48. Among patients in the placebo group who crossed over to ustekinumab treatment (n = 33), increased response rates across efficacy measures were noted. Among all ustekinumab-treated patients, 81.7% had ≥1 adverse event (AE), and 15.1% had ≥1 serious AE through week 56. No deaths, malignancies, opportunistic infections, or tuberculosis cases were observed.

Conclusion. Ustekinumab provided sustained clinical benefit in patients with SLE through 1 year, with a safety profile consistent with other indications.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic inflammatory disease that affects multiple tissues, often resulting in organ damage. The pathogenesis and clinical manifestations of SLE are complex and heterogeneous, which can complicate diagnosis and treatment (1). Conventional therapy for SLE often includes glucocorticoids and immunosuppressants, which carry a risk of

serious toxicities, particularly with long-term use. Belimumab, a monoclonal antibody targeting B lymphocyte stimulator, provides an additional therapeutic option for autoantibody-positive patients with SLE (2). Treat-to-target paradigms for many chronic inflammatory diseases now aim for long-term low disease activity and remission; however, the proportion of patients with SLE who achieve and sustain these important treatment goals is low with currently available standard therapy. Consequently, there remains a significant need to develop therapies that allow patients to maintain a low disease activity or remissive state, prevent flares, and limit organ damage.

Interleukin-12 (IL-12) and IL-23 have been implicated in the pathogenesis of SLE (3–6). Ustekinumab is a monoclonal antibody targeting the p40 subunit of IL-12 and IL-23 and has demonstrated efficacy in phase III trials of patients with plaque psoriasis (7,8), psoriatic arthritis (9,10), Crohn's disease (11), and ulcerative colitis (12). In a randomized, placebo-controlled phase II trial, the safety and efficacy of ustekinumab were evaluated in patients with active moderate-to-severe SLE (13). Through week 24, patients treated with ustekinumab in addition to standard therapy had significantly greater improvements in global SLE disease activity measures compared to patients receiving standard therapy plus placebo. Patients treated with ustekinumab also had a lower risk of flares and greater improvements in skin and joint disease compared to standard therapy alone. Here, we report additional results of this phase II trial, which demonstrated maintenance of efficacy through 1 year of treatment with ustekinumab added to background medication in patients with active SLE.

PATIENTS AND METHODS

Patients and study design. This was a phase II multicenter, prospective, randomized, double-blind, placebo-controlled crossover study of patients with active SLE despite conventional therapy. The details of the study design and patient population have been previously published (13). Briefly, eligible patients were adults 18–75 years of age with a diagnosis of SLE, based on the Systemic Lupus International Collaborating Clinics classification criteria (14), for ≥3 months prior to the first study drug administration. At screening, patients had to have a Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) (15) score of ≥6 (assessed ~2–6 weeks before randomization) and a SLEDAI-2K score of ≥4 for clinical features (excluding laboratory results) at baseline. Eligible patients were also required to have ≥1 British Isles Lupus Assessment Group 2004 index (BILAG 2004) (16) A organ domain score (severe manifestation) and/or ≥2 BILAG B organ domain scores (moderate manifestation) present at screening. An additional screening requirement was a positive test result for autoantibodies (antinuclear antibody titer ≥1:80 by immunofluorescence, anti–double-stranded DNA antibodies >75 IU/ml, and/or anti-Sm antibodies >120 AU/ml) as well as a positive titer for ≥1 of these autoantibodies documented in their medical history.

Prior to screening, all patients were required to be receiving ≥1 of the following conventional therapies for SLE: oral glucocorticoids (≤20 mg/day of prednisone or equivalent), immunosuppressant drugs (e.g., ≤2 gm/day of mycophenolate mofetil/mycophenolic acid, ≤2 mg/kg/day of azathioprine/6-mercaptopurine, or ≤25 mg/week of methotrexate), antimalarials (e.g., hydroxychloroquine, quinacrine, or chloroquine), angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, nonsteroidal antiinflammatory drugs or other analgesics, or topical low- or medium-potency glucocorticoids. With the exception of glucocorticoids, concomitant medications were to be kept stable through week 28; temporary reductions or discontinuation of concomitant medications were permitted for safety reasons.

Patients were randomly assigned (in a 3:2 ratio) to receive a single intravenous infusion of ustekinumab (260 mg for patients weighing ≥35 kg to ≤55 kg; 390 mg for patients weighing >55 kg to ≤85 kg; 520 mg for patients weighing >85 kg) at week 0, followed by subcutaneous injections of a 90-mg dose of ustekinumab at week 8 and every 8 weeks thereafter through week 40, or placebo infusion at week 0 followed by placebo subcutaneous injections at weeks 8 and 16. Patients in the placebo group crossed over to ustekinumab treatment at week 24 and began subcutaneous injections of a 90-mg dose of ustekinumab every 8 weeks from week 24 through week 40 without receiving the initial intravenous dose.

Assessments. Global clinical efficacy assessments included the SLE Responder Index 4 (SRI-4) response, which is a composite measure requiring a reduction of ≥4 points from baseline in the SLEDAI-2K score, no worsening (<10% increase) in the physician global assessment (PhGA) score from baseline (17), no new BILAG A organ domain scores, and ≤1 new BILAG B organ domain score compared to baseline (18). SRI-5 and SRI-6 responses were assessed similarly to the SRI-4 response, with the exception that SLEDAI-2K improvement from baseline was ≥5 or ≥6 points, respectively. SLEDAI-2K response (≥4-point improvement from baseline), PhGA response (≥30% improvement from baseline), and BILAG-based Combined Lupus Assessment (BICLA) response (19) were also used to assess overall efficacy. BICLA response is a composite measure requiring BILAG improvement (all BILAG A and B scores present at baseline improve to lower grade scores, no new BILAG A scores, and ≤1 new BILAG B score), no worsening in the PhGA score, no worsening in the SLEDAI-2K score from baseline (change ≤0), and no treatment failure (i.e., prohibited changes to concomitant medications or discontinuation of study agent due to lack of efficacy or the adverse event [AE] of worsening of SLE prior to week 24) (13).

Mucocutaneous disease was evaluated using the Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) (20), with CLASI response defined as ≥50% improvement from baseline CLASI activity score in patients with a baseline score of ≥4. Joint disease was evaluated by recording the numbers of

tender, swollen, and active (defined as demonstrating both pain/tenderness and signs of inflammation [e.g., swelling or effusion]) joints. Joint response was defined as ≥50% improvement from baseline active joint count in patients with ≥4 active joints at baseline. Flare was defined as ≥1 new BILAG A domain score (severe flare) or ≥2 new BILAG B domain scores (moderate flare) compared to baseline (21).

Patients were monitored throughout the study for AEs, and AEs are reported here according to the actual treatment received. Serum samples were collected for the evaluation of ustekinumab concentrations and the presence of antibodies to ustekinumab. Immunogenicity was assessed using a highly sensitive, drug-tolerant enzyme immunoassay.

This trial was conducted in accordance with the principles of the Declaration of Helsinki and Good Clinical Practices (clinicaltrials.gov: NCT02349061). The protocol was approved by the institutional review board or ethics committee at each site. Patients provided written informed consent before any study-specific procedures were performed.

Statistical analysis. Efficacy results are reported according to the randomized treatment group through week 48. For patients randomized to receive ustekinumab, efficacy analyses included all patients who were randomized to receive ustekinumab at baseline and who received ≥1 administration of ustekinumab. For patients randomized to receive placebo, efficacy analyses after week 24 included only patients who crossed over to ustekinumab treatment. No formal hypothesis testing was performed between the treatment groups after week 24 because all patients were receiving ustekinumab. Proportions of patients achieving SRI, SLEDAI-2K, and BICLA responses were determined using the intent-to-treat population, and patients who met the treatment failure criteria (13) before week 24 were classified as nonresponders for subsequent time points through week 48. Patients who withdrew from the study or had missing values were considered to be nonresponders. Mean changes from baseline in PhGA and SLEDAI-2K scores were reported using observed data, with no imputation for missing data. For the proportion of patients achieving PhGA, CLASI, and joint response and the proportion of patients with no worsening in BILAG or PhGA scores, values for patients who met the treatment failure criteria were classified as missing data from the point of treatment failure forward. Safety data are reported through week 56.

RESULTS

Patient disposition. At baseline, a total of 102 patients were randomized to receive either placebo (n = 42) or ustekinumab (n = 60) in addition to standard therapy for SLE. Patient demographic data and disease characteristics at baseline were well balanced between the treatment groups, as previously

reported (13); some baseline characteristics are also shown in Supplementary Table 1 (available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41179/abstract). Most patients were female (placebo group, n = 35 [83.3%]; ustekinumab group, n = 58 [96.7%]), and the mean age was 42.9 years in the placebo group and 40.0 years in the ustekinumab group. The mean disease duration was also similar between the 2 groups (placebo, mean 9.5 years [range 0.3–24.6]; ustekinumab, mean 9.7 years [range 0.4–32.7]), and 4 patients in each group had a disease duration of <1 year. The most common standard therapy at baseline was glucocorticoids (placebo, n = 34 [80.9%]; ustekinumab, n = 51 [85.0%]).

Through week 24, 13 patients (placebo, n = 9 [21.4%]; ustekinumab, n = 4 [6.7%]) discontinued the study agent (13). Between weeks 24 and 40, an additional 6 patients discontinued the study agent (placebo crossover group, n = 3 [7.1%]; ustekinumab group, n = 3 [5.0%]) (Figure 1). AEs were the most common reason for discontinuation (placebo crossover, n = 5 [11.9%]; ustekinumab, n = 5 [8.3%]).

Clinical efficacy. As previously reported, the SRI-4 response rate at week 24 (primary end point) was significantly greater in patients receiving ustekinumab (62%) than in those receiving placebo (33%) in addition to background therapy (*P* = 0.006) (13). In the ustekinumab group, the proportion of patients achieving an SRI-4 response was maintained from week 24 through week 48
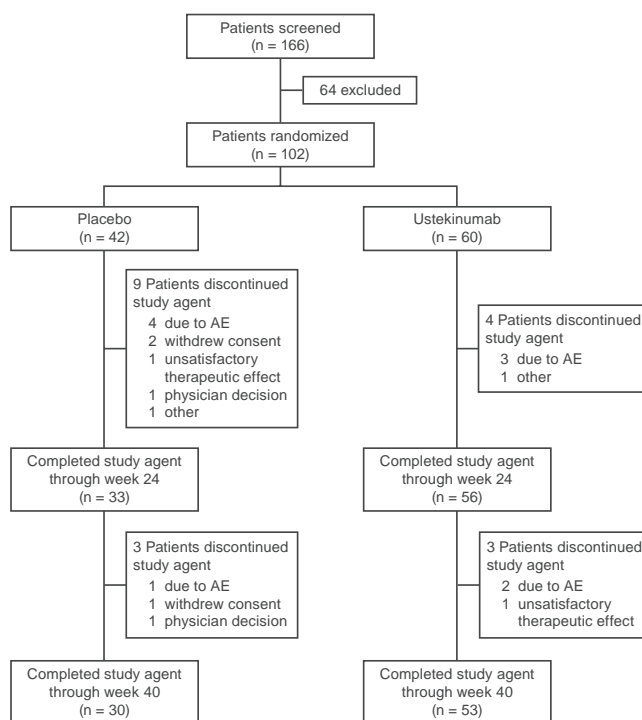
Figure 1. Patient disposition through week 40 (the last administration of study agent). AE = adverse event.

(63.3%) (Figure 2). Similarly, the proportions of patients achieving an SRI-5 and SRI-6 response were maintained from week 24 through week 48 (SRI-5: 43.4% at week 24, 48.3% at week 48; SRI-6: 43.3% at week 24, 46.7% at week 48). In addition, the
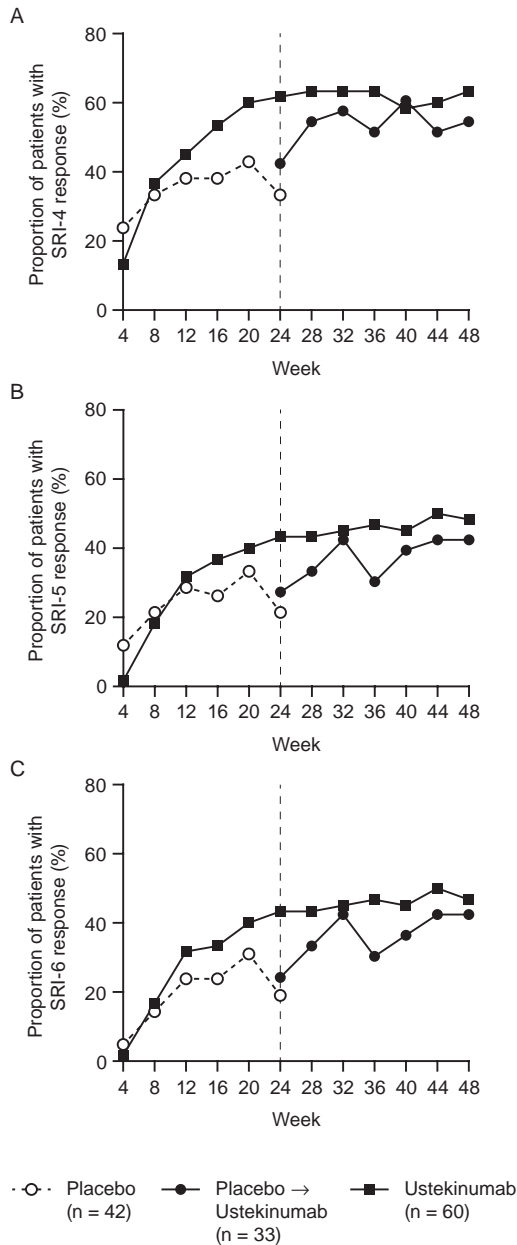
Figure 2. Proportions of patients achieving Systemic Lupus Erythematosus Responder Index (SRI) responses. Responses for SRI-4 (**A**), SRI-5 (**B**), and SRI-6 (**C**) through week 48 are shown. In the placebo group, only patients who crossed over to receive ustekinumab at week 24 (n = 33) were included in efficacy analyses from week 24 through week 48. The SRI-4/5/6 is a composite measure, defined as a reduction (by ≥4, ≥5, or ≥6 points, respectively) in the Systemic Lupus Erythematosus Disease Activity Index 2000 total score, with no new scores from the British Isles Lupus Assessment Group 2004 index (BILAG) A domain and ≤1 new BILAG B domain score, as well as no worsening (<10% increase) from baseline in the physician global assessment.

proportion of patients in the ustekinumab group achieving a SLEDAI-2K response (65.0% at week 24, 66.7% at week 48) and the mean ± SD change from baseline SLEDAI-2K score were sustained from week 24 (mean ± SD –4.7 ± 3.2) through week 48 (mean ± SD –5.4 ± 3.1) (Figure 3 and Supplementary Figure 1, http://onlinelibrary.wiley.com/doi/10.1002/art.41179/abstract). PhGA response rates (67.9% at week 24, 75.0% at week 48) and the mean ± SD change from baseline PhGA score (mean ± SD –2.2 ± 2.0 at week 24, –2.5 ± 2.2 at week 48) were also similar between week 24 and week 48 in the ustekinumab group (Figure 3 and Supplementary Figure 1). In addition, the proportion of patients with no worsening from baseline in PhGA score (92.9% at week 24, 94.2% at week 48) or BILAG score (100% at week 24, 96.2% at week 48), and the proportion of patients with a BICLA response (35.0% at week 24, 45.0% at week 48) were also sustained through week 48 in the ustekinumab group (Supplementary Figure 1).

Improvements in organ-specific disease manifestations including mucocutaneous and joint disease were also observed. The proportion of patients with ≥50% improvement from baseline in CLASI activity score stabilized at week 28 (67.7%) and then was maintained through week 48 (68.6%) in the ustekinumab group (Figure 3). In addition, the proportion of patients with ≥50% improvement from baseline in the number of active joints was maintained from week 24 to week 48 (86.5% at both time points) (Figure 3).

Of the 42 patients randomized to the placebo group, 33 crossed over to receive subcutaneous ustekinumab at week 24. Among these 33 patients, 14 (42.4%) had an SRI-4 response at week 24; this response rate increased to 54.5% (18 of 33 patients) at week 48 (Figure 2). Similar increases were observed in patients achieving an SRI-5 response (27.3% at week 24, 42.4% at week 48) and an SRI-6 response (24.2% at week 24, 42.4% at week 48) (Figure 2). Trends consistent with this were observed in the proportion of patients achieving a SLEDAI-2K response (45.5% at week 24, 54.5% at week 48), PhGA response (56.3% at week 24, 76.7% at week 48), and CLASI response (35.3% at week 24, 47.1% at week 48), and in patients with ≥50% improvement from baseline in the number of active joints (60.9% at week 24, 81.8% at week 48) (Figure 3). The proportion of placebo crossover patients who achieved a BICLA response was slightly increased from week 24 (42.4%) to week 48 (48.5%).

All flares that occurred during the trial were classified as severe (i.e., ≥1 new BILAG A domain score). In the ustekinumab group, flare rates for patients with severe flares were 2.1:10,000 patient-days from week 0 to week 24 and 1.1:10,000 patient-days from week 24 to week 48. Among patients in the placebo group, the rate for severe flares was 8.4:10,000 patient-days from week 0 to week 24. Among the placebo group patients who crossed over to receive ustekinumab at week 24, the severe flare rate was 4.6:10,000 patient-days from week 24 to week 48. The occurrence of severe flares seemed to diminish after ~8 weeks
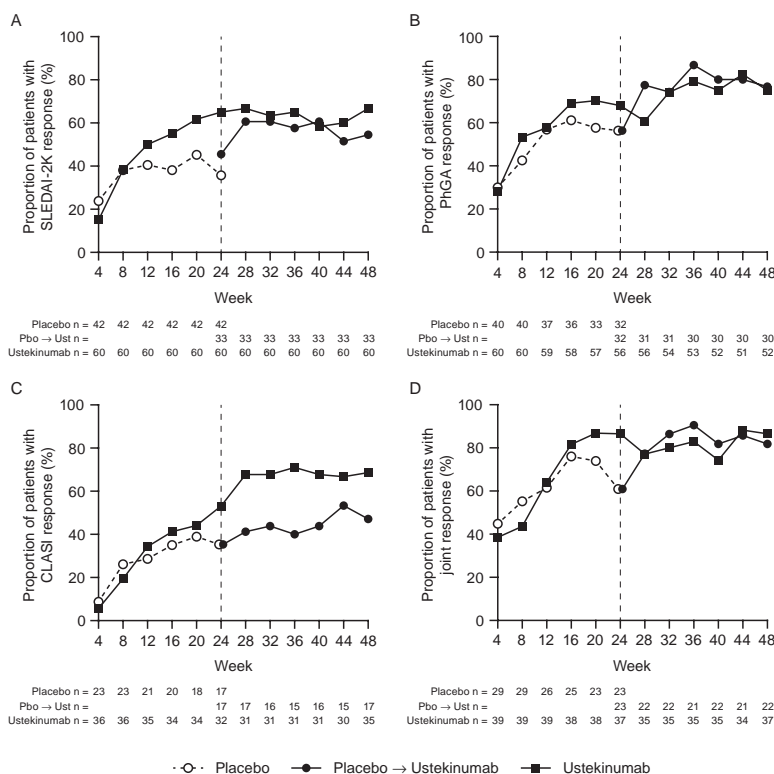
Figure 3. Proportion of patient responses to disease activity indexes. Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) response (**A**), physician global assessment (PhGA) response (**B**), Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) response (**C**), and joint response (**D**) through week 48 are shown. In the placebo group, only patients who crossed over to receive ustekinumab at week 24 (n = 33) were included in efficacy analyses from week 24 through week 48. SLEDAI-2K response was defined as a ≥4-point improvement from baseline. PhGA response was defined as ≥30% improvement from baseline. CLASI response was defined as ≥50% improvement from baseline in patients with a baseline CLASI activity score of ≥4. Joint response was defined as ≥50% improvement from baseline in the number of joints with pain and signs of inflammation (active joints) in patients with ≥4 active joints at baseline. Pbo = placebo; Ust = ustekinumab.

of treatment with ustekinumab (week 8 in the ustekinumab arm, week 32 in the placebo arm) (Figure 4).

Safety. Through week 56, 93 patients (placebo crossover, n = 33; ustekinumab, n = 60) received ≥1 administration of ustekinumab and were included in the safety analysis. Among all ustekinumab-treated patients, 76 (81.7%) had ≥1 AE through week 56 (Table 1). Infections were the most common type of AE. Common infections among all ustekinumab-treated patients were urinary tract infection (18.3%), upper respiratory tract infection (17.2%), and nasopharyngitis (8.6%).

Fourteen (15.1%) of all ustekinumab-treated patients had ≥1 serious AE (SAE) through week 56 (Table 1). Serious infections occurred in 8 patients (8.6%) in the combined ustekinumab group: bronchitis, pneumonia, bacteremia, cellulitis, neutropenic sepsis, *Salmonella* sepsis, and pyrexia (each in 1 patient), and *Stenotrophomonas* infection and urinary tract infection in the same patient. The SAE noted as pyrexia occurred in a patient who presented with an unknown source of fever, sweats, chills, myalgia, and abdominal pain, which resolved after antibiotic therapy. The SAE noted as *Stenotrophomonas* infection occurred in a patient who presented with a lower respiratory tract infection 4 months after

the last dose of ustekinumab. While *Stenotrophomonas* was isolated from a bronchoalveolar aspirate, blood, sputum, and bronchoalveolar lavage fluid were all negative for this organism. Thus, it was believed to be a colonizing organism rather than the source of infection. The lower respiratory tract infection resolved after treatment with levofloxacin. Other SAEs that occurred after week 24 included glomerulonephritis, hypotension, hypochromic anemia, and fracture (humerus). There were no opportunistic infections, cases of tuberculosis, malignancies, or deaths through week 56.

One patient randomized to the ustekinumab group experienced an infusion reaction (potential anaphylactic reaction) prior to week 24 as previously described (13). One patient in the ustekinumab group had multiple mild injection site reactions (urticaria) before week 24 and between week 24 and week 48.

Immunogenicity. Through week 56, a total of 92 patients (placebo crossover, n = 32; ustekinumab, n = 60) had received ≥1 administration of ustekinumab and had samples available for assessing the presence of antibodies to ustekinumab. Among these patients, 10 (31.3%) in the placebo crossover group and 6 (10.0%) in the ustekinumab group tested positive for antibodies to ustekinumab; of these, 3 and 2 patients, respectively, were
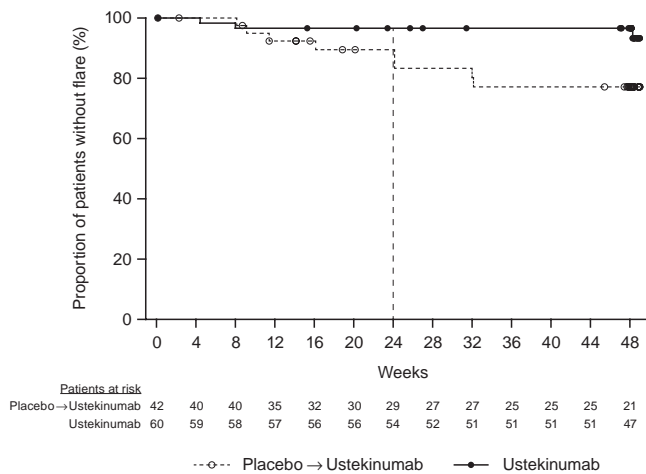
Figure 4. Time to flare (baseline through week 48). Flare was defined as ≥1 new British Isles Lupus Assessment Group 2004 index (BILAG) A domain score (severe flare) or ≥2 new BILAG B domain scores (moderate flare). Patients in the placebo group crossed over to ustekinumab treatment at week 24 (vertical dashed line). All flares experienced during the trial were severe (≥1 new BILAG A domain score). Counts included patients evaluable at each study visit; patients who met the treatment failure criteria were classified as having missing data for subsequent visits.

positive for neutralizing antibodies. All patients who tested positive for antibodies to ustekinumab through week 40 achieved an SRI-4 response by week 48.

DISCUSSION

In this phase II prospective trial, greater proportions of ustekinumab-treated patients had improvements in clinical efficacy assessments at week 24, compared to patients in the placebo group, for many of the indices evaluated. Results reported here demonstrate sustained response rates through 1 year in patients who received ustekinumab from baseline through week 40. Patients in the ustekinumab group also had a significantly lower risk of flares compared to the placebo group through week 24, and the patients in this group maintained a low risk of flares through week 48. Patients in the placebo group who began

subcutaneous ustekinumab therapy at week 24 had an increase in response rates of ~6–20% from week 24 to week 48, as well as a reduction in severe flares (severe flare defined as ≥1 new BILAG A domain score).

Among ustekinumab-treated patients, AEs reported through week 56 were consistent with those reported through week 24. Infections were the most common type of AE and SAE occurring in patients receiving ustekinumab. There were no deaths in either group, and SAEs occurred in 15.1% of all patients who received ≥1 dose of ustekinumab throughout the 56 weeks. There were no new safety findings in patients with SLE, and the safety results through week 56 were consistent with the known safety profile of ustekinumab in patients with psoriasis (7,8), psoriatic arthritis (9,10), Crohn's disease (11), and ulcerative colitis (12).

Despite recent advancements in treatment and increased survival rates, mortality rates among patients with SLE remain high, with progressive organ damage over time, including pulmonary, cardiovascular, and renal disease (22,23). Patients with SLE also have a 4.6-fold greater risk of mortality when compared to the general population (24). Given the heterogeneous nature of SLE, it is often difficult to predict the disease course in these patients. In addition to persistent disease activity, long-term use of glucocorticoids can also contribute to organ damage (25,26). Thus, there continues to be a significant unmet need for therapies to improve long-term outcomes for patients with SLE.

These results are limited by the sample size of this phase II trial and the lack of a placebo comparison after week 24. Because patients in this trial were receiving ≥1 standard medication, it is presently unclear whether ustekinumab may be effective as a monotherapy in SLE. Given the variations in background medications used at baseline, it is also unknown which of these medications is most suitable for combination with ustekinumab, potentially affecting the generalizability of these results. A steroid-sparing effect of ustekinumab could not be ascertained in this study because although tapering of oral glucocorticoids was permitted after week 24, only ~9–12% of evaluable patients had their glucocorticoid dose reduced by study investigators at any visit after week 24 through week 48. Further, this trial did not include patients with certain manifestations of lupus, such as

Table 1. AEs through week 56*

	Placebo, weeks 0–24	Placebo crossover, weeks 24–56†	Ustekinumab, weeks 0–56‡	Combined, weeks 0–56§
No. of patients	42	33	60	93
Mean follow-up, weeks	23.8	26.4	49.5	41.3
Patients with ≥1 AE	29 (69.0)	22 (66.7)	54 (90.0)	76 (81.7)
Infections	20 (47.6)	15 (45.5)	40 (66.7)	55 (59.1)
Patients with ≥1 SAE	4 (9.5)	4 (12.1)	10 (16.7)	14 (15.1)
Serious infections	0	1 (3.0)	7 (11.7)	8 (8.6)

* Except where indicated otherwise, values are the number (%) of patients. SAE = serious adverse event.
† Includes only patients who crossed over to receive ustekinumab.
‡ Includes all patients who were randomized to receive ustekinumab at baseline and received ≥1 administration of ustekinumab.
§ Includes all patients in the placebo crossover and ustekinumab groups who received ≥1 administration of ustekinumab.

severe unstable lupus nephritis or central nervous system disease, which could be evaluated in future studies. In addition, this study was not powered to detect rare AEs, and longer-term safety data on ustekinumab in SLE patients are needed. However, taken together with the week 24 results, these findings indicate that response rates were sustained during the first year of ustekinumab treatment, with an acceptable safety profile in patients with SLE. The efficacy and safety of ustekinumab, including the potential for reduction in glucocorticoid dose, are being evaluated further in a broader population of SLE patients in an ongoing phase III program.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. van Vollenhoven 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. Van Vollenhoven, Hahn, Tsokos, Lipsky, Gordon, Lo, Chevrier, Rose.
Acquisition of data. Gordon, Lo, Chevrier, Rose.
Analysis and interpretation of data. Van Vollenhoven, Hahn, Tsokos, Lipsky, Fei, Gordon, Gregan, Lo, Chevrier, Rose.

ROLE OF THE STUDY SPONSOR

Employees of Janssen Research & Development, LLC participated in study design, data collection, analysis and interpretation of the data, and writing the manuscript. All authors approved the manuscript for submission. Writing assistance was provided by Janssen Scientific Affairs, LLC. Publication of this article was not contingent upon approval by Janssen Research & Development, LLC.

ADDITIONAL DISCLOSURES

Author Lipsky is an employee of AMPEL BioSolutions.

REFERENCES

1. Fava A, Petri M. Systemic lupus erythematosus: diagnosis and clinical management. J Autoimmun 2019;96:1–13.

2. Hahn BH. Belimumab for systemic lupus erythematosus. N Engl J Med 2013;368:1528–35.

3. Dai H, He F, Tsokos GC, Kyttaris VC. IL-23 limits the production of IL-2 and promotes autoimmunity in lupus. J Immunol 2017;199:903–10.

4. Dai J, Liu B, Cua DJ, Li Z. Essential roles of IL-12 and dendritic cells but not IL-23 and macrophages in lupus-like diseases initiated by cell surface HSP gp96. Eur J Immunol 2007;37:706–15.

5. Deng Y, Tsao BP. Updates in lupus genetics [review]. Curr Rheumatol Rep 2017;19:68.

6. Kyttaris VC, Zhang Z, Kuchroo VK, Oukka M, Tsokos GC. Cutting edge: IL-23 receptor deficiency prevents the development of lupus nephritis in C57BL/6-lpr/lpr mice. J Immunol 2010;184:4605–9.

7. Leonardi CL, Kimball AB, Papp KA, Yeilding N, Guzzo C, Wang Y, et al. Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 76-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 1). Lancet 2008;371:1665–74.

8. Papp KA, Langley RG, Lebwohl M, Krueger GG, Szapary P, Yeilding N, et al. Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 52-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 2). Lancet 2008;371:1675–84.

9. McInnes IB, Kavanaugh A, Gottlieb AB, Puig L, Rahman P, Ritchlin C, et al. Efficacy and safety of ustekinumab in patients with active psoriatic arthritis: 1 year results of the phase 3, multicentre, double-blind, placebo-controlled PSUMMIT 1 trial. Lancet 2013;382:780–9.

10. Ritchlin C, Rahman P, Kavanaugh A, McInnes IB, Puig L, Li S, et al. Efficacy and safety of the anti-IL-12/23 p40 monoclonal antibody, ustekinumab, in patients with active psoriatic arthritis despite conventional non-biological and biological anti-tumour necrosis factor therapy: 6-month and 1-year results of the phase 3, multicentre, double-blind, placebo-controlled, randomised PSUMMIT 2 trial. Ann Rheum Dis 2014;73:990–9.

11. Feagan BG, Sandborn WJ, Gasink C, Jacobstein D, Lang Y, Friedman JR, et al. Ustekinumab as induction and maintenance therapy for Crohn's disease. N Engl J Med 2016;375:1946–60.

12. Sands BE, Sandborn WJ, Panaccione R, O'Brien CD, Zhang H, Johanns J, et al. Ustekinumab as induction and maintenance therapy for ulcerative colitis. N Engl J Med 2019;381:1201–14.

13. Van Vollenhoven RF, Hahn BH, Tsokos GC, Wagner CL, Lipsky P, Touma Z, et al. Efficacy and safety of ustekinumab, an IL-12 and IL-23 inhibitor, in patients with active systemic lupus erythematosus: results of a multicentre, double-blind, phase 2, randomised, controlled study. Lancet 2018;392:1330–9.

14. Petri M, Orbai AM, Alarcón GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum 2012;64:2677–86.

15. Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol 2002;29:288–91.

16. Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004: development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford) 2005;44:902–6.

17. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727–35.

18. Furie RA, Petri MA, Wallace DJ, Ginzler EM, Merrill JT, Stohl W, et al. Novel evidence-based systemic lupus erythematosus responder index. Arthritis Rheum 2009;61:1143–51.

19. Wallace D, Strand V, Furie R, Petri M, Kalunian K, Pike M, et al. Evaluation of treatment success in systemic lupus erythematosus clinical trials: development of the British Isles Lupus Assessment Group-based composite lupus assessment endpoint [abstract]. Arthritis Rheum 2011;63 Suppl 10:S885–6.

20. Albrecht J, Taylor L, Berlin JA, Dulay S, Ang G, Fakharzadeh S, et al. The CLASI (Cutaneous Lupus Erythematosus Disease Area and Severity Index): an outcome instrument for cutaneous lupus erythematosus. J Invest Dermatol 2005;125:889–94.

21. Navarra SV, Guzman RM, Gallacher AE, Hall S, Levy RA, Jimenez RE, et al. Efficacy and safety of belimumab in patients with active

systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. Lancet 2011;377:721–31.

22. Lopez R, Davidson JE, Beeby MD, Egger PJ, Isenberg DA. Lupus disease activity and the risk of subsequent organ damage and mortality in a large lupus cohort. Rheumatology (Oxford) 2012;51:491–8.

23. Peschken CA, Wang Y, Abrahamowicz M, Pope J, Silverman E, Sayani A, et al. Persistent disease activity remains a burden for patients with systemic lupus erythematosus. J Rheumatol 2019; 46:166–75.

24. Jacobsen S, Petersen J, Ullman S, Junker P, Voss A, Rasmussen JM, et al. Mortality and causes of death of 513 Danish patients with systemic lupus erythematosus. Scand J Rheumatol 1999;28: 75–80.

25. Al Sawah S, Zhang X, Zhu B, Magder LS, Foster SA, Iikuni N, et al. Effect of corticosteroid use by dose on the risk of developing organ damage over time in systemic lupus erythematosus: the Hopkins Lupus Cohort. Lupus Sci Med 2015;2:e000066.

26. Bruce IN, O'Keeffe AG, Farewell V, Hanly JG, Manzi S, Su L, et al. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the Systemic Lupus International Collaborating Clinics (SLICC) Inception Cohort. Ann Rheum Dis 2015;74: 1706–13.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

B Cell Tetherin: A Flow Cytometric Cell-Specific Assay for Response to Type I Interferon Predicts Clinical Features and Flares in Systemic Lupus Erythematosus

Yasser M. El-Sherbiny,[1] iD   Md Yuzaiful Md Yusof,[2] iD   Antonios Psarras,[2]   Elizabeth M. A. Hensor,[2] iD
Kumba Z. Kabba,[3]   Katherine Dutton,[2]   Alaa A. A. Mohamed,[4]   Dirk Elewaut,[5] iD   Dennis McGonagle,[2]
Reuben Tooze,[3]   Gina Doody,[3]   Miriam Wittmann,[2]   Paul Emery,[2] iD   and Edward M. Vital[2]

Objective. Type I interferon (IFN) responses are broadly associated with autoimmune diseases, including systemic lupus erythematosus (SLE). Given the cardinal role of autoantibodies in SLE, this study was undertaken to investigate whether the findings of a B cell–specific IFN assay correlate with SLE activity.

Methods. B cells and peripheral blood mononuclear cells (PBMCs) were stimulated with type I IFN and type II IFN. Gene expression was analyzed, and the expression of pathway-related membrane proteins was determined. A flow cytometry assay for tetherin (CD317), an IFN-induced protein ubiquitously expressed on leukocytes, was validated in vitro and then clinically against SLE diagnosis, plasmablast expansion, and the British Isles Lupus Assessment Group (BILAG) 2004 score in a discovery cohort (n = 156 SLE patients, 30 rheumatoid arthritis [RA] patients, and 25 healthy controls). A second, longitudinal validation cohort of 80 SLE patients was also evaluated for flare prediction.

Results. In vitro, a close cell-specific and dose-response relationship between type I IFN–responsive genes and cell surface tetherin was observed in all immune cell subsets. Tetherin expression on multiple cell subsets was selectively responsive to stimulation with type I IFN compared to types II and III IFNs. In patient samples from the discovery cohort, memory B cell tetherin showed the strongest associations with diagnosis (SLE:healthy control effect size 0.11 [$P = 0.003$]; SLE:RA effect size 0.17 [$P < 0.001$]), plasmablast numbers in rituximab-treated patients (R = 0.38, $P = 0.047$), and BILAG 2004. These associations were equivalent to or stronger than those for IFN score or monocyte tetherin. Memory B cell tetherin was found to be predictive of future clinical flares in the validation cohort (hazard ratio 2.29 [95% confidence interval 1.01–4.64]; $P = 0.022$).

Conclusion. Our findings indicate that memory B cell surface tetherin, a B cell–specific IFN assay, is associated with SLE diagnosis and disease activity, and predicts flares better than tetherin on other cell subsets or whole blood assays, as determined in an independent validation cohort.

INTRODUCTION

Type I interferons (IFNs) are a highly pleiotropic group of cytokines that link the innate and adaptive immune systems and play a pivotal role in autoimmune disease (1–3). All nucleated cells express type I IFN receptors and express a set of IFN-stimulated genes (ISGs) after exposure to type I IFN (4,5). Hundreds of effects of type I IFN on various cellular processes, interactions, and disease processes have been described. A challenge in the assessment of type I IFN response in an individual disease is therefore ensuring that the appropriate cellular response can be detected within this complex system.

Type I IFN proteins are unstable in blood and not easily detected even in monogenic interferonopathies with known high type I IFN production, possibly due to their efficient binding to the abundant IFN receptor (6). Type I IFN activity is therefore usually measured using the expression of ISGs in whole blood. We previously analyzed ISG expression in sorted cells from patients with systemic lupus erythematosus (SLE), a prototypic IFN-mediated disease, and healthy individuals and showed that in both groups, ISG expression was markedly higher in monocytes than in other circulating immune cells. ISG expression in monocytes therefore dominates ISG assays that use unsorted blood (7).

These differing levels of ISG expression in different cell populations may be due to the rate of turnover in each population, their trafficking to sites of higher type I IFN production in inflamed tissues, or priming for type I IFN response by other inflammatory mediators. In autoimmunity, type I IFN assays may have value to predict flares and response to a range of different targeted therapies (8). However, existing whole blood IFN biomarkers show poor or uncertain correlation with disease activity (9–11).

The measurement of type I IFN using ISG expression in whole blood has 2 key weaknesses with regard to interpreting pathogenic processes. First, changes in expression may reflect the expansion or contraction of certain circulating leukocyte populations (12,13) that differ in their level of ISG expression. This characteristically occurs in inflammatory diseases. In the case of SLE, lymphopenia is almost universally seen (14). So, any difference in whole blood gene expression may not necessarily indicate a change in the production of or exposure to type I IFN. Second, analyzing whole blood ISG expression does not allow the detection of key pathogenic processes among the noise of other, less relevant, effects of type I IFN on biology. For example, B cells are a key mediator in SLE (15,16). Type I IFN stimulates B cells to differentiate into plasmablasts, which are expanded in SLE and correlate with disease activity (17,18). We previously demonstrated that the rate of plasmablast regeneration after rituximab treatment predicts clinical outcome (19). We also previously showed that type I IFN imprints plasma cells for the secretion of the proinflammatory molecule ISG-15 (17). Assessment of type I IFN activity in unsorted blood gives limited information about the degree to which B cells have

specifically been stimulated by type I IFN. Further, gene expression assays do not prove that a phenotypic change in target cells has occurred—there has been no widely used biomarker for IFN response at the protein level. This may be one reason why some patients classified as having a low IFN signature have responded well to IFN-blocking therapy (20).

In order to resolve these problems, we developed a flow cytometry assay that allows measurement of type I IFN response in individual cells without the need for cell sorting. We measured the expression of tetherin (also known as bone marrow stromal antigen 2 [BST-2]; CD317), a glycosyl phosphatidylinositol–anchored protein with a unique topology that is ubiquitously expressed on the surface of nucleated cells. This molecule is prominent in viral immunology and encoded by a commonly measured ISG expressed in all leukocytes (4,5,21–23). Unlike most ISGs, *BST2* encodes a cell surface protein and can be easily measured in patient samples by flow cytometry. Sialic acid–binding Ig-like lectin 1 (Siglec-1) is another flow cytometry type I IFN biomarker that has been described previously (24,25). However, Siglec-1 is only expressed on monocytes so resolves the issue of changes in the size of cell populations but does not allow interrogation of type I IFN responses in individual cells subsets, including the key B cell populations that are strongly linked to clinical and experimental disease (26–28).

We hypothesized that a dominant pathogenic role of type I IFN in SLE is its effect on B cells, promoting plasmablast differentiation and clinical disease. Our reasons for addressing B cells as a particular cell of interest in SLE were: 1) SLE is associated with autoantibodies, which are made by B cells; 2) there are a number of susceptibility loci for SLE in genes with important roles in B cell signaling and function, such as *LYN*, *BLK*, *BANK1*, *PTPN22*, *TNFAIP3*, and *TNIP1* (29); and 3) the only targeted therapy licensed for SLE targets B cells specifically. Using in vitro stimulation and sorted cells from SLE patients and healthy individuals, we showed that tetherin accurately captures cell-specific responses to type I IFN. A crucial issue in biomarker research is demonstrating that biomarkers are predictive, correlate with a range of outcomes, and can be reproduced in validation studies. In our study, longitudinal analysis of discovery and validation cohorts showed that memory B cell tetherin levels more accurately correlated with plasmablast expansion and clinical features of disease, and predicted flares better, compared to monocyte tetherin or whole blood ISG expression.

PATIENTS AND METHODS

The discovery cohort included 156 consecutive SLE patients, 25 age-matched healthy controls, and 30 patients with active rheumatoid arthritis (RA) as non-SLE inflammatory disease controls. The RA patients were positive for anti–citrullinated protein antibody, negative for antinuclear antibody (ANA), and had a mean Disease Activity Score in 28 joints of 3.9 (95% confidence interval [95% CI] 3.23–4.56). An independent validation cohort consisted of 80 SLE patients recruited and studied longitudinally

(n = 236 SLE patients total). SLE disease activity was assessed at the time of sampling using the British Isles Lupus Assessment Group 2004 (BILAG 2004) score (30). Patients in the validation cohort were also followed up for subsequent flare (a BILAG score of A or B). SLE patient demographics and disease activity are shown in Supplementary Table 1, available on the *Arthritis & Rheu-matology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract. Patients with acute or chronic viral infection at the time of blood sampling were excluded from this study.

All individuals provided informed written consent, and the study was carried out in compliance with the Declaration of Helsinki. The patient blood samples used for this study were
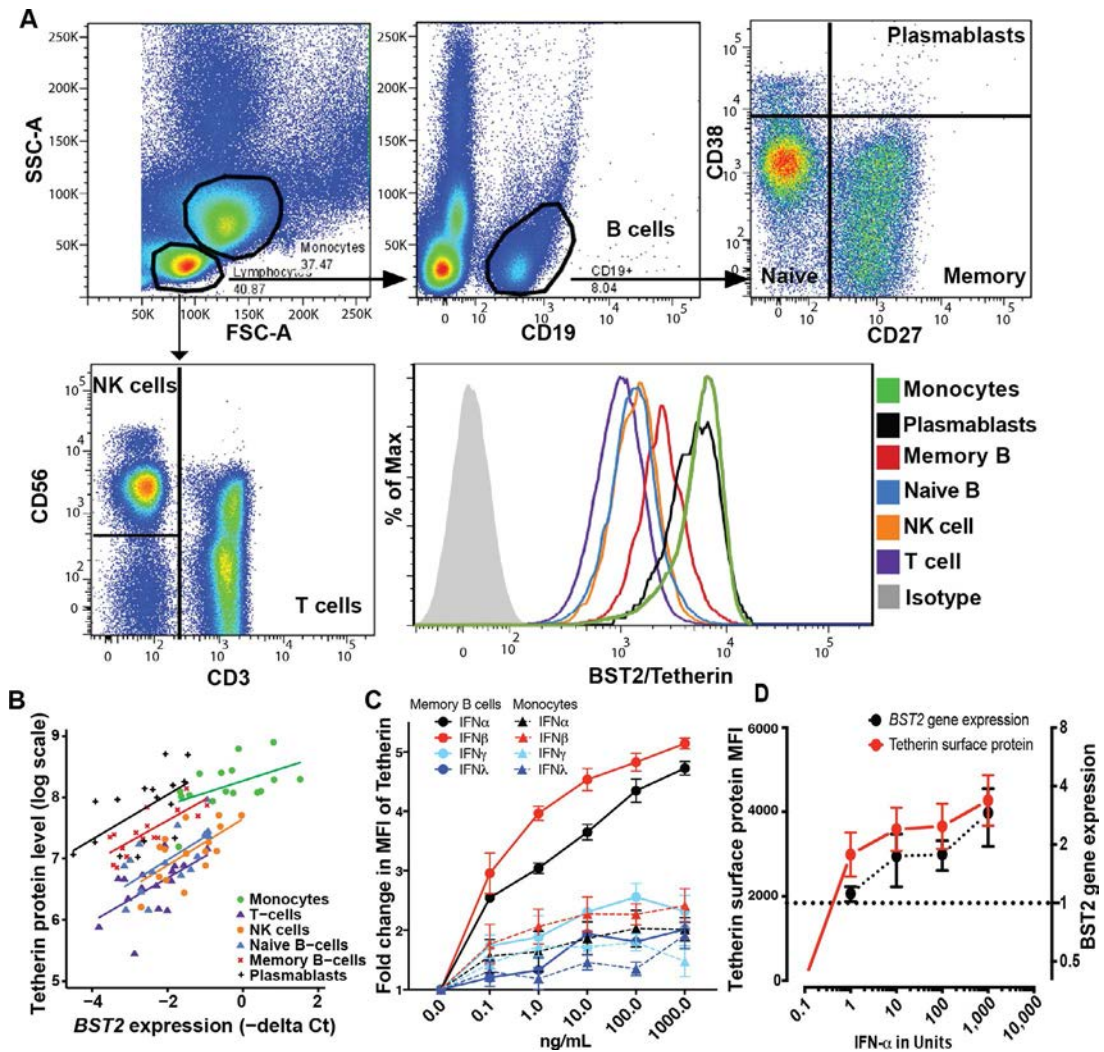
Figure 1. Tetherin is a scalable cell-specific measure of type I interferon (IFN) response. **A**, Gating strategy for flow cytometric assessment of tetherin on immune cell subsets. A representative flow cytometry plot of tetherin protein expression on individual immune cell subsets is shown. FSC-A and SSC-A were used to define lymphocytes and monocytes. B cells were defined as CD19+ lymphocytes and subdivided into naive, memory, and plasmablast subsets using CD27 and CD38. T cells were defined as CD3+, and natural killer (NK) cells were defined as CD3–CD56+ lymphocytes. The mean fluorescence intensity (MFI) of bone marrow stromal antigen 2/tetherin for each cell subset compared to isotype control is shown. **B**, Correlation of tetherin protein level with *BST*2 gene expression for the indicated immune cell subsets. In order to validate tetherin as a cell-specific marker, tetherin protein expression was compared with expression of its gene *BST*2 in various immune cell subsets in systemic lupus erythematosus patients and healthy controls. Cell surface tetherin protein levels were determined in unsorted peripheral blood mononuclear cells (PBMCs) by flow cytometry, and *BST*2 gene expression data were determined by quantitative polymerase chain reaction of cells sorted by fluorescence-activated cell sorting. There was a strong correlation between gene expression and protein level within each subset, allowing differences in IFN-stimulated gene expression between cell subsets to be measured without cell sorting (for monocytes, R = 0.47, *P* = 0.064; for T cells, R = 0.61, *P* = 0.012; for NK cells, R = 0.63, *P* = 0.008; for naive B cells, R = 0.63, *P* = 0.009; for memory B cells, R = 0.78, *P* = 0.001; and for plasmablasts, R = 0.58, *P* = 0.018). **C**, Dose-dependent response of memory B cell tetherin and monocyte tetherin to IFN. Healthy control PBMCs (n = 3 samples) were stimulated with increasing doses of IFNα, IFNβ, IFNγ, and IFNλ, and tetherin MFI was determined by flow cytometry. **D**, Tetherin protein levels and *BST*2 gene expression levels in sorted B cells stimulated in vitro with increasing doses of IFNα and evaluated by flow cytometry. There was a parallel increase in each marker. Dotted line indicates a 1-fold increase in *BST*2 gene expression. In **C** and **D**, values are the mean ± SD.

obtained with ethics approval (REC 10/H1306/88, National Research Ethics Committee Yorkshire and Humber–Leeds East), and healthy control participant peripheral blood samples were obtained under study number 04/Q1206/107. All experiments were performed in accordance with relevant guidelines and regulations. The University of Leeds was contracted with administrative sponsorship. Additional details are included in the Supplementary Methods, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract, and in a previously published methodology article (7).

RESULTS

BST-2/tetherin as a cell-specific phenotypic biomarker of type I IFN response. Global gene expression profiles have shown that many ISGs are responsive to both IFNα and IFNγ, while other ISGs respond specifically to IFNα (7,17). We therefore tested the effect of IFNα (type I IFN) and IFNγ (type II IFN) on 31 of the most commonly reported ISGs. TaqMan quantitative polymerase chain reaction (qPCR) analysis of B cells in vitro was performed as previously described (17) (Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.41187/abstract). In vitro stimulation confirmed that *BST2* was in the group of ISGs predominantly responsive to type I IFN.

For this reason, we used multiparameter flow cytometry to detect and quantify tetherin on peripheral blood mononuclear cells (PBMCs), as described in the Supplementary Methods (available

on the *Arthritis & Rheumatology* web site at http://onlinelibrary. wiley.com/doi/10.1002/art.41187/abstract). We used a gating strategy that allowed us to define T cells, natural killer (NK) cells, and monocytes as well as the B cell subsets naive B cells, memory B cells, and plasmablasts. For each of these populations, the mean fluorescence intensity (MFI) of tetherin compared to isotype control was determined (Figure 1A). We compared cell surface tetherin protein levels, determined by flow cytometry, with *BST2* gene expression, determined by qPCR, for these 6 cell subsets sorted by fluorescence-activated cell sorting from 10 SLE patients and 6 healthy controls. *BST2* gene expression levels were substantially positively correlated with cell surface tetherin protein levels within each of the subsets (Figure 1B). These data confirm that varying levels of tetherin/*BST2* expression between cell subsets and differences between individuals may be captured using flow cytometry without the need for cell sorting. Furthermore, we compared tetherin and Siglec-1 MFI on monocytes, B cells, and T cells in samples from 25 SLE patients and 5 healthy controls. We confirmed that tetherin correlated with Siglec-1 on monocytes only because other cell subsets lacked Siglec-1 expression (Supplementary Figure 2, available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.41187/abstract).

Dose response of tetherin to type I IFN, type II IFN, and type III IFN. We tested the dose-response relationship of tetherin to IFNα, IFNβ (both type I IFN), IFNγ (type II IFN), and IFNλ (type III IFN) on all circulating cell subsets. Healthy control PBMCs were stimulated for 48 hours with doses of 0.1–1,000 ng/ml and then analyzed by flow cytometry. Interestingly, memory B cell

Table 1. Tetherin levels in cell subsets in SLE patients and healthy controls*

	Mean MFI tetherin protein levels in SLE patients (n = 113)†	Within SLE, between cell subset		Mean MFI tetherin protein levels in healthy controls (n = 17)†	Between group (SLE:healthy controls)		Between group, between cell subset	
		Ratio (90% CI)	P		Ratio (90% CI)	P	Ratio (90% CI)	P
All subjects								
Monocytes	3,388	Reference		2,837	1.19 (0.91–1.58)	0.293	Reference	
T cells	687	0.20 (0.19–0.22)	<0.001	475	1.45 (1.17–1.79)	0.005	1.21 (1.00–1.46)	0.092
NK cells	1,129	0.33 (0.31–0.36)	<0.001	824	1.37 (1.09–1.72)	0.024	1.15 (0.94–1.40)	0.258
Naive B cells	1,118	0.33 (0.30–0.36)	<0.001	712	1.57 (1.22–2.03)	0.004	1.32 (1.03–1.68)	0.062
Memory B cells	1,586	0.47 (0.43–0.50)	<0.001	1,033	1.53 (1.23–1.92)	0.002	1.29 (1.04–1.58)	0.046
Plasmablasts	2,650	0.78 (0.72–0.84)	<0.001	1,813	1.46 (1.15–1.86)	0.009	1.22 (0.99–1.51)	0.115
Rituximab-naive only‡								
Monocytes	3,206	Reference		2,949	1.09 (0.80–1.48)	0.657	Reference	
T cells	666	0.21 (0.19–0.23)	<0.001	494	1.35 (1.07–1.70)	0.034	1.24 (1.01–1.52)	0.080
NK cells	1,068	0.33 (0.31–0.36)	<0.001	857	1.25 (0.98–1.58)	0.129	1.15 (0.93–1.41)	0.271
Naive B cells	1,132	0.35 (0.32–0.39)	<0.001	740	1.53 (1.20–1.95)	0.004	1.41 (1.15–1.73)	0.006
Memory B cells	1,574	0.49 (0.45–0.53)	<0.001	1,074	1.47 (1.17–1.83)	0.005	1.35 (1.11–1.64)	0.013
Plasmablasts	2,597	0.81 (0.74–0.89)	<0.001	1,885	1.38 (1.08–1.76)	0.033	1.27 (1.02–1.57)	0.068

* Tetherin cell protein data were natural log–transformed prior to analysis. The back-transformed results represent the ratio of the value for each cell subset relative to the value for monocytes within the group of patients with systemic lupus erythematosus (SLE), and the ratio of the value for each cell subset in SLE patients relative to that in healthy controls. Interaction ratios (between group, between cell subset) are the ratio of the extent of the difference in the value for each cell subset relative to monocytes between SLE patients and healthy controls. MFI = mean fluorescence intensity; 90% CI = 90% confidence interval; NK = natural killer.
† Adjusted for age.
‡ n = 76 patients with SLE.

tetherin MFI was most responsive to increasing doses of IFNα and IFNβ, and showed more modest responses to IFNγ and IFNλ. Although monocytes had the highest expression of tetherin in patient samples and the highest basal expression in unstimulated healthy control PBMCs, they showed much lower fold change in tetherin response to type I IFN stimulation (Figure 1C). Furthermore, purified B cell response curves for *BST2* gene expression and tetherin protein MFI revealed a closely matched dose-response to IFNα (Figure 1D). We concluded that tetherin MFI determined by flow cytometry could accurately measure change in the expression of *BST2* in response to type I IFN and could be used to capture type I IFN exposure in a dose- and cell-specific manner.

B cell surface tetherin protein levels best demonstrate disease-associated IFN response in SLE. We next compared tetherin protein expression in immune cell subsets in SLE patients and healthy controls to determine which cell subset best demonstrates disease-associated change in IFN response. Using all discovery cohort data, tetherin protein levels were compared between the different cell subsets in SLE patients and healthy controls by flow cytometry (Table 1).

Tetherin levels differed significantly between cell subsets within the group of SLE patients and were highest in monocytes. Tetherin levels in T cells were 20% of the levels in monocytes, and tetherin levels in plasmablasts were 78% of the levels in monocytes (both $P < 0.001$). A comparison of the SLE and healthy control groups showed that monocyte tetherin MFI in SLE patients did not differ significantly from that in healthy controls (SLE:healthy control ratio 1.19; $P = 0.293$), whereas a significantly higher level of tetherin was seen in all other subsets in SLE patients versus healthy controls (ratios 1.37–1.57; all $P < 0.05$). When the between-group ratio for each of the other cell subsets was compared to that for monocytes, the greatest difference was seen for memory B cell tetherin levels, which were increased in SLE patients ($P = 0.046$).

Rituximab treatment could confound accurate measurement of the B cell phenotype in the SLE patients. We therefore repeated these analyses in rituximab-naive patients (Table 1). In rituximab-naive patients (n = 76), the largest disease-associated increases in tetherin expression versus healthy controls were seen for naive B cells (ratio 1.53) and memory B cells (ratio 1.47). These ratios for naive and memory B cells were significantly different from that for monocytes ($P = 0.006$ and $P = 0.013$, respectively). These results indicate that differences in IFN response between cell subsets at the protein level are clinically relevant, and that B cell tetherin is the most clinically relevant parameter.

Tetherin and IFN gene expression assays. Overall comparisons of tetherin levels measured on memory B cells and 2 validated IFN gene expression scores are shown in Supplementary Figures 3 and 4, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract. As expected, given the difference in cell populations analyzed, there was a significant correlation but a degree of disagreement between these assays.
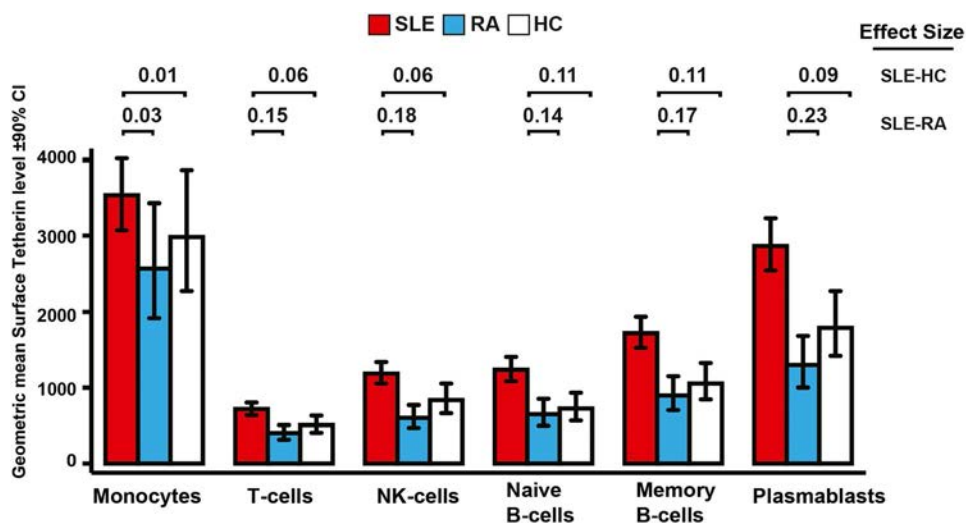
Figure 2. Performance of the tetherin interferon (IFN) flow cytometry assay in discriminating patients based on diagnosis. Age-adjusted differences in tetherin levels on the indicated cell subsets between patients with systemic lupus erythematosus (SLE), patients with active rheumatoid arthritis (RA) (with a Disease Activity Score in 28 joints of >3.2), and healthy controls (HCs) are shown. Cell surface bone marrow stromal antigen 2/tetherin protein levels were determined by flow cytometry of peripheral blood mononuclear cells. Effect sizes (partial η²) indicate the degree to which variables differed between groups. We considered an effect size of ≤0.01 to be small, ~0.06 to be medium, and ≥0.14 to be large (31). Bars show the mean and 90% confidence interval (90% CI). NK = natural killer. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract.

Clinical validation of the tetherin IFN assay: diagnosis. We evaluated the performance of the tetherin flow cytometry assay in distinguishing between patients diagnosed as having SLE, those diagnosed as having active RA, and healthy controls. Given our previous results, for this analysis we included only rituximab-naive patients controlled for age (Figure 2). The full statistical table is shown in Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract. An effect size of ≤0.01 was considered small, ~0.06 medium, and ≥0.14 large, as described by Cohen (31). Tetherin data revealed a marked difference between cell subsets. Monocyte tetherin levels did not differentiate SLE patients from healthy controls at all, with a ratio of 1.19 (90% CI 0.87–1.61) and an effect size of 0.01. T cells and NK cells had moderate effect sizes of 0.06 each. However, naive and memory B cell subsets had medium to large effect sizes of 0.11, with ratios of 1.63 (90% CI 1.26–2.11) and 1.59 (90% CI 1.21–2.09), respectively. For memory B cells, the SLE:healthy control effect size was 0.11 ($P = 0.003$), and the SLE:RA effect size was 0.17 ($P < 0.001$).

Tetherin levels differentiated SLE from other inflammatory disease when compared to active RA. Monocyte tetherin levels had no diagnostic function, with a ratio of 1.37 (90% CI 1.00–1.88) and an effect size of 0.03. However, all other cell subsets had large effect sizes, ranging from 0.14 to 0.23, the effect size for plasmablasts (ratio 2.20 [90% CI 1.66–2.93]).

Clinical validation of IFN assays: disease activity and autoantibodies. For disease activity, we investigated the association between the number of active organ systems (BILAG domains scored A, B, or C) per patient compared to tetherin levels on cell subsets as well as our recently described IFN score A, which comprises 12 type I IFN–selective ISGs (7). We controlled for age in all SLE patients (164 observations in 124 patients). The number of active domains was categorized as 0 (n = 22), 1 (n = 54), 2 (n = 57), or ≥3 (n = 31).

At the 10% level of significance, disease activity was associated with IFN score A ($R^2 = 0.08$, $P = 0.027$) and tetherin surface expression on T cells ($R^2 = 0.07$, $P = 0.007$), NK cells ($R^2 = 0.09$, $P = 0.001$), memory B cells ($R^2 = 0.09$, $P = 0.006$), and plasmablasts ($R^2 = 0.06$, $P = 0.020$). The degree of association was weaker, and hence not significant, for monocytes ($R^2 = 0.04$, $P = 0.179$) and naive B cells ($R^2 = 0.04$, $P = 0.103$).

For IFN score A, the relationship with disease activity was not linear. The only significant association between the score and disease activity was attributable to patients with the most severely active disease (≥3 domains). A similar, although not significant, pattern was observed for monocyte tetherin levels. In contrast, there was a linear relationship between memory B cell tetherin levels and disease activity, with a stepwise increase in tetherin expression for each increase in the number of active domains (Figure 3A). We did not expect a strong correlation

between memory B cell tetherin levels and IFN score in unsorted PBMCs. Since memory B cells are only ~2% of PBMCs, these assays do not measure the same biologic effect. We found a moderate correlation (Spearman's R = 0.356, $P < 0.0001$) (Supplementary Figure 3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract).

To investigate whether the difference between IFN assays was due to the type of organ system affected, we analyzed the 2 most commonly affected domains (mucocutaneous and musculoskeletal) in combination, excluding patients with activity (BILAG scores of A, B, or C) in any of the other domains. Although there was a significant relationship between each IFN assay and overall disease activity, the relationship with IFN assays varied between these 2 organ systems (Figure 3B). For IFN score A, increased expression was only seen with mucocutaneous disease activity. For monocyte tetherin levels, increases were observed only in patients with musculoskeletal disease activity. This finding may explain why this assay does not show a linear relationship with disease activity. However, memory B cell tetherin had a more consistent relationship with disease activity in both organ systems. Tetherin levels were lowest in patients in clinical remission, higher in patients with active disease in a single organ, and highest in patients with active disease in both organs.

The numbers of patients with other active organ domains were more limited. Of patients with no activity in other domains, 12 had active hematologic disease (BILAG score of A or B) (immune-mediated hemolysis or thrombocytopenia). Memory B cell tetherin MFI in the patients with active hematologic disease versus those with inactive disease was 1,954 versus 1,494, respectively ($P = 0.005$). Eight patients had active renal disease. Comparing these 8 patients with active renal disease to patients with inactive disease also revealed a significant increase in memory B cell tetherin levels (MFI 2,625 versus 1,562; $P = 0.005$). Tetherin levels were not associated with glucocorticoid use (Supplementary Figure 5, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract). In our rituximab-naive patients, there was a positive correlation between memory B cell tetherin expression and autoantibodies summarized as the number of extractable nuclear antigen subtypes (R = 0.412, $P = 0.0001$) (Supplementary Figure 6, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract).

For additional comparison to alternative IFN assays, we classified patients as type I IFN positive or type I IFN negative according to a 5-gene IFN signature, as described by Higgs et al (32). Results are shown in Supplementary Figure 7, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract. The majority of the SLE patients were in the type I IFN–positive subgroup. As expected, this subgroup had worse BILAG disease activity ($P = 0.016$). To determine whether tetherin expression gave additional informa-
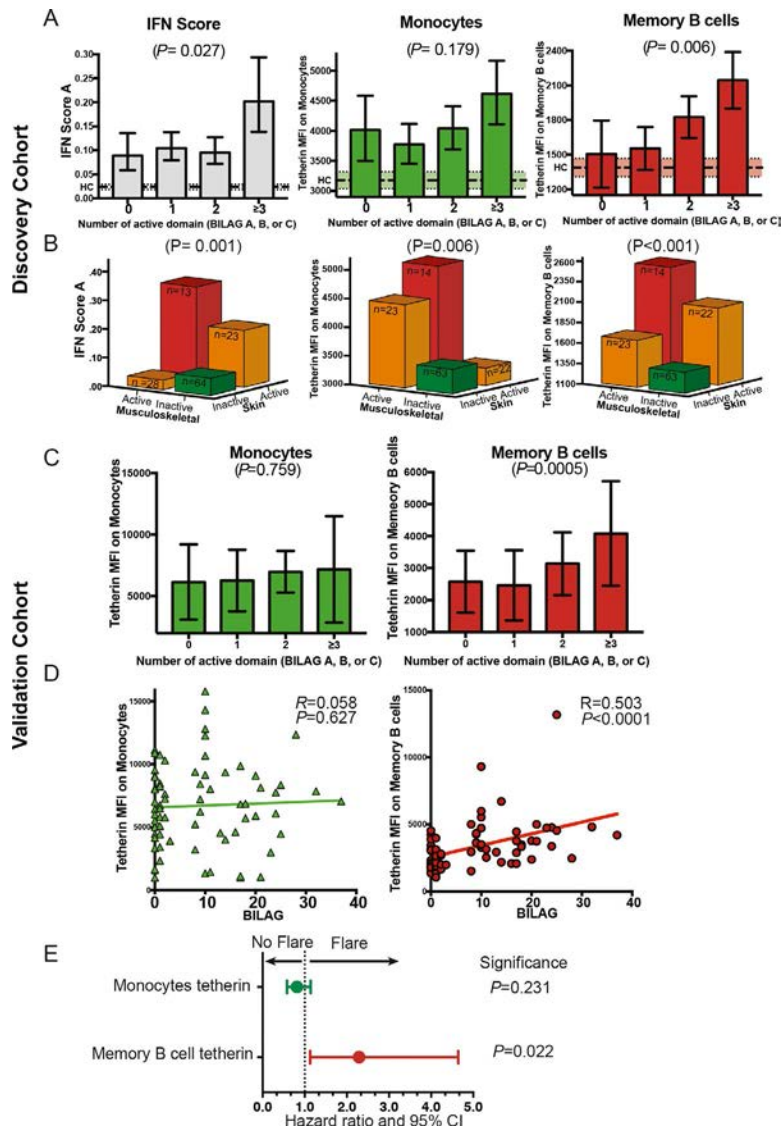
Figure 3. Association between interferon (IFN) assays and disease activity in systemic lupus erythematosus (SLE). **A**, Association between different IFN assays (IFN score A, monocyte tetherin levels, and memory B cell tetherin levels) and the number of organ systems (domains) with active disease in the discovery cohort (164 observations in 124 SLE patients). IFN score A was increased in patients with ≥3 active domains, but not in patients with 1 or 2 active domains, compared to those with 0 active domains (remission). Tetherin mean fluorescence intensity (MFI) measured on memory B cells demonstrated a more consistent stepwise increase with increasing disease activity. Bars show the mean and 90% confidence interval (90% CI), calculated using the $2^{-\Delta Ct}$ method (i.e., taller bars represent higher expression). Broken lines and shaded areas represent the mean and 90% CI in healthy controls (HCs; n = 23). **B**, Association between different IFN assays and musculoskeletal and mucocutaneous disease activity. Disease activity was defined as active (British Isles Lupus Assessment Group [BILAG] score of A or B) or inactive (BILAG score of D or E). Patients with activity in other organs were excluded. For IFN score A, there were inconsistent relationships with disease activity, with an increase with skin involvement, but not musculoskeletal involvement alone. For monocyte tetherin levels, increased protein expression was seen with musculoskeletal disease activity but not with skin activity alone. Tetherin levels measured on memory B cells demonstrated a consistent relationship with both common types of clinical disease. Bars show the median. **C**, Association between different IFN assays (monocyte tetherin levels and memory B cell tetherin levels) and the number of organ systems with active disease in the validation cohort. Results were similar to those for the discovery cohort, shown in **A**. Bars show the mean ± SD (n = 80 patients). **D**, Scatterplots showing association between overall disease activity (BILAG global score) and tetherin levels. There was a significant association between BILAG global score and memory B cell tetherin levels but not monocyte tetherin levels. **E**, Relationship between tetherin levels and SLE disease flare. Memory B cell tetherin levels were significantly predictive of subsequent clinical flare (hazard ratio [HR] 2.290 [95% CI 1.013–4.644]; P = 0.022), while monocyte tetherin levels were not (HR 0.814 [95% CI 0.580–1.141]; P = 0.231).

tion in comparison to the gene expression status, we retested the association of tetherin levels with BILAG scores within the type I IFN–positive subgroup. We still found a significant association

between tetherin expression and disease activity (Spearman's R = 0.321, P = 0.038), which could not be measured using the more standard assay. This finding indicates that memory B cell tetherin

gives additional clinically relevant information compared to the IFN signature alone.

Clinical validation of IFN assays: plasmablasts. Last, in the discovery cohort, we used the plasmablast count to represent current B cell activity and differentiation. Type I IFN is known to promote the differentiation of memory B cells into plasmablasts (33). We have previously shown that an early rapid population of plasmablasts after rituximab treatment led to an early clinical relapse (19,34). We hypothesized that the memory B cell tetherin level would correlate with circulating plasmablast numbers after rituximab treatment, reflecting an increased rate of differentiation secondary to type I IFN. The results are shown in Table 2. In rituximab-naive patients, no relationship was found between any tetherin IFN assay and plasmablast count. In patients who had received rituximab treatment there was no correlation between monocyte, NK, or T cell tetherin levels and plasmablast numbers, but memory B cell tetherin levels were significantly correlated with plasmablast numbers (Spearman's R = 0.38, $P = 0.047$) as well as inversely correlated with time to clinical relapse (R = 0.623, $P = 0.022$). To further explore whether tetherin surface protein expression was associated with the induction of relevant pathogenic pathways in B cells, we evaluated 2 transcripts for downstream plasmablast function: *IgJ* for antibody synthesis in all samples, and *ISG15* for ISG-15 protein secretion in sorted memory B cells. Both of these transcripts showed a significant correlation with flow cytometric measurement of memory B cell tetherin MFI (Supplementary Figure 8, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41187/abstract).

Independent validation cohort. The independent validation cohort consisted of an additional 80 patients with SLE who were recruited and studied prospectively. Memory B cell and monocyte tetherin levels were measured using fresh lysed whole

Table 2. Association between candidate IFN assays and plasmablast level following B cell depletion therapy in SLE patients*

	Plasmablast count before rituximab treatment (n = 50)	Plasmablast count after rituximab treatment (n = 28)
IFN score A	−0.11	0.24
Tetherin protein level		
Monocytes	−0.08	0.20
T cells	−0.16	0.32
NK cells	−0.14	0.05
Naive B cells	−0.04	0.30
Memory B cells	0.07	0.38†

* Values are Spearman's rank correlation coefficient. Interferon (IFN) score A was measured on unsorted peripheral blood mononuclear cells, and tetherin mean fluorescence intensity (MFI) on each cell subset was analyzed by flow cytometry. SLE = systemic lupus erythematosus; NK = natural killer.
† $P = 0.047$ for correlation between plasmablast count and tetherin MFI.

blood in an independent accredited diagnostic laboratory. Disease activity was measured at the time of sampling using the BILAG 2004. Patients were followed up for subsequent flare (a BILAG score of A or B).

We found a similar relationship between tetherin levels and disease activity as in our discovery cohort. There was a significant relationship between the number of organ domains with active disease and memory B cell tetherin levels ($P = 0.0005$) but no relationship with monocyte tetherin levels ($P = 0.759$) (Figure 3C). There was a significant association between global BILAG score and memory B cell tetherin levels (Spearman's R = 0.503, $P < 0.0001$) but no association with monocyte tetherin levels (R = 0.058, $P = 0.627$) (Figure 3D). Additionally, in this cohort we demonstrated that in patients in clinical remission at the time of sampling (n = 36), memory B cell tetherin levels predicted time to clinical flare. In a multivariable Cox regression analysis including memory B cell tetherin level, monocyte tetherin level, and age, memory B cell tetherin level was a significant predictor of subsequent BILAG A/B flare (hazard ratio [HR] 2.290 [95% CI 1.013–4.644]; $P = 0.022$). Monocyte tetherin level did not significantly predict flare (HR 0.814 [95% CI 0.580–1.141]; $P = 0.231$) (Figure 3E). In conclusion, we independently confirmed that disease activity is related to type I IFN response in memory B cells measured using tetherin, and further, that this is predictive of clinical outcome.

DISCUSSION

In this study we demonstrated the value of a novel cell-specific biomarker based on the IFN-inducible protein tetherin, using in vitro methods and human clinical studies. We showed that flow cytometric measurement of memory B cell surface tetherin levels captured cell-specific type I IFN response, was responsive to increasing doses of type I IFN, and had a strong and consistent relationship with disease activity, B cell activity, and time to flare in 2 cohorts of SLE patients. These results are important because type I IFN and B cells play a role in many autoimmune diseases, and their measurement has the potential to stratify outcomes and use of therapies, though previous studies have yielded conflicting results (35).

Better biomarkers are needed in SLE. European League Against Rheumatism treat-to-target recommendations advise treating to a target of low disease activity, while minimizing exposure to glucocorticoids (36). Predictors of a severe disease trajectory or flares are needed to achieve this goal. Response to conventional and targeted therapies in SLE and related diseases is variable, and reclassification of autoimmune diseases according to pathogenic mechanisms instead of clinical features has been proposed (35).

The crucial role of type I IFN in the pathogenesis of SLE and related diseases is indicated by genetic susceptibility and monogenic interferonopathies as well as evidence of overexpression (35). As such, it has face validity as a stratification biomarker.

Existing studies indicate the potential value of measuring type I IFN for diagnosis and prediction of flares. Type I IFN biomarkers may also predict clinical response to tumor necrosis factor blockade, B cell depletion, and type I IFN blockade in RA and SLE (35).

Nevertheless, there are limitations to previous approaches to measuring type I IFN activity, and some previous results have been contradictory. Direct measurement of type I IFN protein is limited by the number of different ligands and instability in serum, with most cell types expressing the type I IFN receptor. A recent improvement was the use of single-molecule arrays (Simoa). The higher sensitivity of Simoa allows for reliable measurement of IFNα (6). However, this assay is currently expensive and limited in availability and has not been validated against clinical outcomes. For ISG expression–based methods, another issue is the effect of other IFN subtypes or other inflammatory mediators. ISGs are known to fall into distinct subsets, which may be due to the effect of type II IFN (7,9). We previously showed that there are different patterns of ISG expression in different autoimmune diseases. In the present study we confirmed that tetherin is selectively responsive to type I IFN, and we included ANA-negative RA patients as inflammatory disease controls. (We did not see any elevation of tetherin levels in our RA patients as others have reported for an IFN signature, but this difference may be due to our selection of only ANA-negative cases rather than differences in the biomarkers).

While candidate biomarker discoveries in autoimmunity are numerous, a significant challenge is validation in clinically relevant contexts (37). An important aspect of our work is the degree of preclinical and clinical validation. We used 2 methods of validation to demonstrate that tetherin reflects cellular response to type I IFN. We demonstrated a correlation with existing validated IFN assays. However, such concurrent validity studies are limited by the potential imprecision of the IFN scores. These scores may be affected by changes in the cellular composition of the sample or other subtypes of IFNs. Moreover, tetherin assesses the response to IFN of a specific cell subset (we have shown memory B cells), while IFN scores assess a mixed population of cells and will be influenced by other cell types. For these reasons, the more important method of demonstrating that tetherin reflects cellular response to type I IFN is through in vitro stimulation assays. We showed that tetherin has a dose-dependent response to type I IFN in multiple cell subsets, far exceeding response to type II IFN. Our data therefore demonstrate good face and construct validity, as well as concurrent and prospective criterion validation and feasibility in a routine clinical setting. We also present validation against a range of different clinical and longitudinal end points.

Cell-specific measurement based on flow cytometry has been demonstrated previously using expression of Siglec-1, another cell surface protein convenient for flow cytometry that is expressed by monocytes. Monocyte Siglec-1 expression has been shown to correlate with disease activity as well as predict autoimmune congenital heart block (25,38,39). This was a significant advance in analysis of IFN status. In the present study we advanced this principle further by using a marker expressed on all circulating cells. Tetherin captures the same information as Siglec-1 on monocytes, but also evaluates other cell subsets. We have shown that results from these different subsets vary, with the strongest clinical correlation for memory B cells. This method has distinct advantages when there is particular interest in a specific cell population, such as with the B cell–directed therapies rituximab and belimumab in SLE. B cell response to type I IFN is crucial in SLE.

While there were many associations between tetherin protein expression and clinical features of SLE, memory B cell tetherin levels seemed to be particularly important. This marker correlated best with clinical features, and was the only marker to be associated with plasmablast number. After B cell depletion with rituximab, there is a highly variable rate of plasmablast repopulation that predicts clinical relapse. Understanding the determinants of these repopulation patterns may reveal upstream factors controlling B cell autoreactivity. One previous study showed a relationship between serum BAFF titers and the numbers of plasmablasts at relapse (40). However, BAFF may not be the only factor. Type I IFN also promotes B cell activation and differentiation into plasmablasts and plasma cells (28,41). This may include direct influences; for example, in animal models type I IFN influences B cell receptor– and Toll-like receptor–mediated response to self nuclear antigen. Our work provides data from human disease to support this observation (42,43). Additionally, type I IFN induces a plasma cell phenotype that secretes ISG-15 with additional proinflammatory effects (17). In the present study, we found that memory B cell tetherin levels correlated with plasmablast expansion after rituximab treatment. A plasmablast signature was recently shown to be a strong biomarker for SLE, and we and others previously showed that plasmablast expansion after rituximab was strongly predictive of clinical relapse (19,44,45). This was further supported by a correlation between memory B cell tetherin levels and transcripts representing disease-relevant B cell dysfunction.

The tetherin biomarker has some limitations. First, although this flow cytometry assay avoids confounders that may affect ISG expression scores, analyzing a single IFN-inducible transcript may be more susceptible to the influence of other inflammatory stimuli, which we cannot exclude based on these results. However, our data comparing SLE to the RA disease control are very consistent with those we observed using IFN scores, with a clear difference in IFN score A and tetherin expression between SLE and RA. Tetherin, like all type I IFN biomarkers, may be influenced by acute or chronic viral infections, which were excluded from this study. It may be more difficult to perform flow cytometry in some situations. However, with widespread use of flow cytometry in cell-targeted therapies in autoimmunity and oncology, as well as in routine monitoring of HIV, addition of tetherin cell surface staining is a highly cost-effective test. Tetherin may be analyzed in combination with B cell and plasmablast flow cytometry to stratify both B cell– and type I IFN–blocking therapy.

In summary, we describe measurement of the IFN-inducible protein tetherin on B cells as a cell-specific biomarker with a number of advantages and widespread applications in clinical and laboratory research in this rapidly expanding area of immunology.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Vital 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. El-Sherbiny, Emery, Vital.
Acquisition of data. El-Sherbiny, Md Yusof, Psarras, Hensor, Kabba, Dutton, Mohamed, Elewaut, McGonagle, Tooze, Doody, Wittmann.
Analysis and interpretation of data. El-Sherbiny, Md Yusof, Psarras, Hensor, Kabba, Dutton, Mohamed, Elewaut, McGonagle, Tooze, Doody, Wittmann, Emery, Vital.

REFERENCES

1. Crow MK. Type I interferon in the pathogenesis of lupus. J Immunol 2014;192:5459–68.

2. Rönnblom L, Eloranta ML. The interferon signature in autoimmune diseases. Curr Opin Rheumatol 2013;25:248–53.

3. Niewold TB. Interferon α as a primary pathogenic factor in human lupus. J Interferon Cytokine Res 2011;31:887–92.

4. Vidal-Laliena M, Romero X, March S, Requena V, Petriz J, Engel P. Characterization of antibodies submitted to the B cell section of the 8th Human Leukocyte Differentiation Antigens Workshop by flow cytometry and immunohistochemistry. Cell Immunol 2005;236:6–16.

5. Evans EJ, Hene L, Sparks LM, Dong T, Retiere C, Fennelly JA, et al. The T cell surface—how well do we know it? Immunity 2003;19:213–23.

6. Rodero MP, Decalf J, Bondet V, Hunt D, Rice GI, Werneke S, et al. Detection of interferon α protein reveals differential levels and cellular sources in disease. J Exp Med 2017;214:1547–55.

7. El-Sherbiny YM, Psarras A, Yusof MY, Hensor EM, Tooze R, Doody G, et al. A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. Sci Rep 2018;8:5793.

8. Raterman HG, Vosslamber S, de Ridder S, Nurmohamed MT, Lems WF, Boers M, et al. The interferon type I signature towards prediction of non-response to rituximab in rheumatoid arthritis patients. Arthritis Res Ther 2012;14:R95.

9. Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. Arthritis Rheumatol 2014;66:1583–95.

10. Kawasaki M, Fujishiro M, Yamaguchi A, Nozawa K, Kaneko H, Takasaki Y, et al. Fluctuations in the gene expression of peripheral blood mononuclear cells between the active and inactive phases of systemic lupus erythematosus. Clin Exp Rheumatol 2010;28:311–7.

11. Landolt-Marticorena C, Bonventi G, Lubovich A, Ferguson C, Unnithan T, Su J, et al. Lack of association between the interferon-α signature and longitudinal changes in disease activity in systemic lupus erythematosus. Ann Rheum Dis 2009;68:1440–6.

12. Becker AM, Dao KH, Han BK, Kornu R, Lakhanpal S, Mobley AB, et al. SLE peripheral blood B cell, T cell and myeloid cell transcriptomes display unique profiles and each subset contributes to the interferon signature. PLoS One 2013;8:e67003.

13. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA, et al. Individuality and variation in gene expression patterns in human blood. Proc Natl Acad Sci U S A 2003;100:1896–901.

14. Fayyaz A, Igoe A, Kurien BT, Danda D, James JA, Stafford HA, et al. Haematological manifestations of lupus [review]. Lupus Sci Med 2015;2:e000078.

15. Taylor KE, Chung SA, Graham RR, Ortmann WA, Lee AT, Langefeld CD, et al. Risk alleles for systemic lupus erythematosus in a large case-control collection and associations with clinical subphenotypes. PLoS Genet 2011;7:e1001311.

16. Md Yusof MY, Vital EM, Emery P. Biologics in systemic lupus erythematosus: current options and future perspectives. Br J Hosp Med (Lond) 2014;75:440–7.

17. Care MA, Stephenson SJ, Barnes NA, Fan I, Zougman A, El-Sherbiny YM, et al. Network analysis identifies proinflammatory plasma cell polarization for secretion of ISG15 in human autoimmunity. J Immunol 2016;197:1447–59.

18. Jacobi AM, Mei H, Hoyer BF, Mumtaz IM, Thiele K, Radbruch A, et al. HLA-DRhigh/CD27high plasmablasts indicate active disease in patients with systemic lupus erythematosus. Ann Rheum Dis 2010;69:305–8.

19. Vital EM, Dass S, Buch MH, Henshaw K, Pease CT, Martin MF, et al. B cell biomarkers of rituximab responses in systemic lupus erythematosus. Arthritis Rheum 2011;63:3038–47.

20. Kalunian KC, Merrill JT, Maciuca R, McBride JM, Townsend MJ, Wei X, et al. A phase II study of the efficacy and safety of rontalizumab (rhuMAb interferon-α) in patients with systemic lupus erythematosus (ROSE). Ann Rheum Dis 2016;75:196–202.

21. Blasius AL, Giurisato E, Cella M, Schreiber RD, Shaw AS, Colonna M. Bone marrow stromal cell antigen 2 is a specific marker of type I IFN-producing cells in the naive mouse, but a promiscuous cell surface antigen following IFN stimulation. J Immunol 2006;177:3260–5.

22. Neil SJ, Zang T, Bieniasz PD. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature 2008;451:425–30.

23. El-Sherbiny YM, El-Jawhari JJ, Moseley TA, McGonagle D, Jones E. T cell immunomodulation by clinically used allogeneic human cancellous bone fragments: a potential novel immunotherapy tool. Sci Rep 2018;8:13535.

24. Rose T, Grützkau A, Hirseland H, Huscher D, Dähnrich C, Dzionek A, et al. IFNα and its response proteins, IP-10 and SIGLEC-1, are biomarkers of disease activity in systemic lupus erythematosus. Ann Rheum Dis 2013;72:1639–45.

25. Biesen R, Demir C, Barkhudarova F, Grün JR, Steinbrich-Zöllner M, Backhaus M, et al. Sialic acid–binding Ig-like lectin 1 expression in inflammatory and resident monocytes is a potential biomarker for monitoring disease activity and success of therapy in systemic lupus erythematosus. Arthritis Rheum 2008;58:1136–45.

26. Morawski PA, Bolland S. Expanding the B cell-centric view of systemic lupus erythematosus. Trends Immunol 2017;38:373–82.

27. Nashi E, Wang Y, Diamond B. The role of B cells in lupus pathogenesis [review]. Int J Biochem Cell Biol 2010;42:543–50.

28. Malkiel S, Barlev AN, Atisha-Fregoso Y, Suurmond J, Diamond B. Plasma cell differentiation pathways in systemic lupus erythematosus [review]. Front Immunol 2018;9:427.

29. Vaughn SE, Kottyan LC, Munroe ME, Harley JB. Genetic susceptibility to lupus: the biological basis of genetic risk found in B cell signaling pathways. J Leukoc Biol 2012;92:577–91.

30. Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004: development and initial validation of an updated

version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford) 2005;44:902–6.

31. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.

32. Higgs BW, Liu Z, White B, Zhu W, White WI, Morehouse C, et al. Patients with systemic lupus erythematosus, myositis, rheumatoid arthritis and scleroderma share activation of a common type I interferon pathway. Ann Rheum Dis 2011;70:2029–36.

33. Cocco M, Stephenson S, Care MA, Newton D, Barnes NA, Davison A, et al. In vitro generation of long-lived human plasma cells. J Immunology 2012;189:5773–85.

34. Md Yusof MY, Shaw D, El-Sherbiny YM, Dunn E, Rawstron AC, Emery P, et al. Predicting and managing primary and secondary non-response to rituximab using B-cell biomarkers in systemic lupus erythematosus. Ann Rheum Dis 2017;76:1829–36.

35. Psarras A, Emery P, Vital EM. Type I interferon-mediated autoimmune diseases: pathogenesis, diagnosis and targeted therapy. Rheumatology (Oxford) 2017;56:1662–75.

36. Van Vollenhoven RF, Mosca M, Bertsias G, Isenberg D, Kuhn A, Lerstrøm K, et al. Treat-to-target in systemic lupus erythematosus: recommendations from an international task force. Ann Rheum Dis 2014;73:958–67.

37. Paulovich AG, Whiteaker JR, Hoofnagle AN, Wang P. The interface between biomarker discovery and clinical validation: the tar pit of the protein biomarker pipeline. Proteomics Clin Appl 2008;2:1386–402.

38. Kyogoku C, Smiljanovic B, Grün JR, Biesen R, Schulte-Wrede U, Häupl T, et al. Cell-specific type I IFN signatures in autoimmunity and viral infection: what makes the difference? PLoS One 2013;8:e83776.

39. Lisney AR, Szelinski F, Reiter K, Burmester GR, Rose T, Dörner T. High maternal expression of SIGLEC1 on monocytes as a surrogate marker of a type I interferon signature is a risk factor for the development of autoimmune congenital heart block. Ann Rheum Dis 2017;76:1476–80.

40. Carter LM, Isenberg DA, Ehrenstein MR. Elevated serum BAFF levels are associated with rising anti–double-stranded DNA antibody levels and disease flare following B cell depletion therapy in systemic lupus erythematosus. Arthritis Rheum 2013;65:2672–9.

41. Jego G, Palucka AK, Blanck JP, Chalouni C, Pascual V, Banchereau J. Plasmacytoid dendritic cells induce plasma cell differentiation through type I interferon and interleukin 6. Immunity 2003;19:225–34.

42. Lau CM, Broughton C, Tabor AS, Akira S, Flavell RA, Mamula MJ, et al. RNA-associated autoantigens activate B cells by combined B cell antigen receptor/Toll-like receptor 7 engagement. J Exp Med 2005;202:1171–7.

43. Chaturvedi A, Dorward D, Pierce SK. The B cell receptor governs the subcellular location of Toll-like receptor 9 leading to hyperresponses to DNA-containing antigens. Immunity 2008;28:799–809.

44. Anolik JH, Barnard J, Owen T, Zheng B, Kemshetti S, Looney RJ, et al. Delayed memory B cell recovery in peripheral blood and lymphoid tissue in systemic lupus erythematosus after B cell depletion therapy. Arthritis Rheum 2007;56:3044–56.

45. Lazarus MN, Turner-Stokes T, Chavele KM, Isenberg DA, Ehrenstein MR. B-cell numbers and phenotype at clinical relapse following rituximab therapy differ in SLE patients according to anti-dsDNA antibody levels. Rheumatology (Oxford) 2012;51:1208–15.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Role of Systemic Lupus Erythematosus Risk Variants With Opposing Functional Effects as a Driver of Hypomorphic Expression of *TNIP1* and Other Genes Within a Three-Dimensional Chromatin Network

Satish Pasula,[1] Kandice L. Tessneer,[1] Yao Fu,[1] Jaanam Gopalakrishnan,[2] Richard C. Pelikan,[1] Jennifer A. Kelly,[1] Graham B. Wiley,[1] Mandi M. Wiley,[1] and Patrick M. Gaffney[2] (iD)

Objective. Genetic variants in the region of tumor necrosis factor–induced protein 3–interacting protein 1 (*TNIP1*) are associated with autoimmune disease and reduced *TNIP1* gene expression. The aim of this study was to define the functional genetic mechanisms driving *TNIP1* hypomorphic expression imparted by the systemic lupus erythematosus–associated *TNIP1* H1 risk haplotype.

Methods. Dual luciferase expression and electrophoretic mobility shift assays were used to evaluate the allelic effects of 11 risk variants on enhancer function and nuclear protein binding in immune cell line models (Epstein-Barr virus [EBV]–transformed human B cells, Jurkat cells, and THP-1 cells), left in a resting state or stimulated with phorbol 12-myristate 13-acetate/ionomycin. HiChIP was used to define the regulatory 3-dimensional (3-D) chromatin network of the *TNIP1* haplotype by detecting in situ long-range DNA contacts associated with H3K27ac-marked chromatin in EBV B cells. Then, quantitative reverse transcription–polymerase chain reaction (qRT-PCR) was used to determine the expression of genes within the 3-D chromatin network.

Results. Bioinformatics analyses of 50 single-nucleotide polymorphisms on the *TNIP1* H1 risk haplotype identified 11 non–protein-coding variants with a high likelihood of influencing *TNIP1* gene expression. Eight variants in EBV B cells, 5 in THP-1 cells, and 2 in Jurkat cells exhibited various allelic effects on enhancer activation, resulting in a cumulative suppressive effect on *TNIP1* expression (net effect of risk variants −7.14 fold, −6.80 fold, and −2.44 fold, respectively; n > 3). Specifically, in EBV B cells, only 2 variants (rs10057690 and rs13180950) exhibited allele-specific loss of both enhancer activity and nuclear protein binding (each $P < 0.01$ relative to nonrisk alleles). In contrast, the rs10036748 risk allele reduced binding affinities of the transcriptional repressors basic helix-loop-helix family member 40/differentially expressed in chondrocytes 1 (bHLHe40/DEC1) ($P < 0.05$ relative to nonrisk alleles) and CREB-1 (P not significant) in EBV B cells, resulting in a gain of enhancer activity ($P < 0.05$). HiChIP and qRT-PCR analyses revealed that overall transcriptional repression of the *TNIP1* haplotype extended to the neighboring genes *DCTN4* and *GMA2*, both of which also showed decreased expression in the presence of the *TNIP1* risk haplotype ($P < 0.001$ and $P < 0.01$, respectively, relative to the nonrisk haplotype); notably, it was found that these genes share a 3-D chromatin network.

Conclusion. Hypomorphic *TNIP1* expression results from the combined concordant and opposing effects of multiple risk variants carried on the *TNIP1* risk haplotype, with the strongest regulatory effect in B lymphoid lineage cells. Furthermore, the *TNIP1* risk haplotype effect extends to neighboring genes within a shared chromatin network.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a challenging autoimmune disease characterized by a loss of immune self tolerance and widespread dysfunction of the innate and adaptive immune systems, which can lead to systemic end-organ damage, severe morbidity, and early mortality. Genome-wide association studies (GWAS) and targeted genetic scans of SLE case–control populations have identified more than 100 genetic loci associated with increased SLE susceptibility (1,2). Tumor necrosis factor (TNF)–induced protein 3–interacting protein 1 (*TNIP1*), located on 5q32–33.1, has been identified by multiple GWAS, large-scale replication, and genetic fine-mapping studies as a strong SLE susceptibly locus shared across multiple racial groups, including European American, African American, Gullah, Hispanic, Hispanic Dominican Republic, Chinese, and Japanese (3–9). Multiple independent and race-specific SLE risk haplotypes spanning the *TNIP1* locus have been identified, including independent H1 and H2 haplotypes that we previously identified in European Americans (3,6). Genetic variability within the *TNIP1* locus has also been associated with several other autoimmune diseases, including Sjögren's syndrome (10), systemic sclerosis (11,12), psoriatic arthritis (13), and psoriasis (13–16), as well as asthma (17) and some cancers (18–20).

TNIP1 encodes the polyubiquitin binding protein A20-binding inhibitor of NF-κB (ABIN-1). *TNIP1*/ABIN-1 is a critical negative regulator of the proinflammatory NF-κB signaling pathway (21), and has been implicated in regulation of Toll-like receptors (TLRs) (22), peroxisome proliferator-activated receptor (23), retinoic acid receptor (24), and CCAAT/enhancer binding protein β (C/EBPβ) (25) signaling pathways. Loss of *TNIP1*/ABIN-1 in mice exacerbates NF-κB and C/EBPβ signaling, resulting in chronic inflammation and the progressive development of lupus-like inflammatory phenotypes, including immune cell expansion and activation, circulating autoantibodies, and renal dysfunction (21,25,26). These findings strongly imply that the hypomorphic *TNIP1* expression that is associated with SLE *TNIP1* risk haplotypes is an important contributing factor in the pathogenesis of SLE (3,27).

Despite its strong genetic association with SLE and the identification of several risk variants, the mechanisms that modulate the cell type– and state-specific transcriptional regulation of *TNIP1* in the context of SLE is not well understood. Using a combination of in vitro assays across different immune cell lineages, in a resting state or after stimulation, we functionally characterized SLE risk variants located in enhancers spanning the *TNIP1* H1 risk haplotype. Our findings suggest that multiple variants with concordant and opposing allelic effects in vitro collectively function to suppress *TNIP1* gene expression, most potently in B lymphoid lineage cells. Moreover, we also demonstrate that the hypomorphic expression induced by the *TNIP1* risk haplotype extends beyond *TNIP1* to other genes that share 3-dimensional (3-D) chromatin contact domains with the risk haplotype.

MATERIALS AND METHODS

Antibodies and cell lines. Jurkat and THP-1 cells were purchased from ATCC. Epstein-Barr virus (EBV)–transformed human B cell lines were obtained from the Lupus Family Registry and Repository at the Oklahoma Medical Research Foundation, with Institutional Review Board approval (28). EBV B cell lines were selected using genotype data corresponding to the different *TNIP1* variants. Genotypes were verified by Sanger sequencing. Cell lines were maintained in RPMI 1640 medium supplemented with 10% fetal bovine serum, 1X penicillin/streptomycin antibiotic mixture (Atlanta Biologicals), and 2 mM L-glutamine (Lonza). THP-1 cell medium was also supplemented with 50 μM β-mercaptoethanol.

Cells were left in a resting state or stimulated with phorbol 12-myristate 13-acetate (PMA) and ionomycin (P/I) (50 ng/ml PMA and 500 ng/ml ionomycin) for 2 hours. The following antibodies were used: anti–early growth response 1 (Egr-1) and anti–CREB-1 (product nos. 44D5 and 48H2, respectively; Cell Signaling Technology), antibodies to basic helix-loop-helix family member 40/differentially expressed in chondrocytes 1 (bHLHe40/DEC1) (product no. NB100-1800; Novus Biologicals), anti–histone H3 (acetyl K27) and anti–β-actin (product nos. ab4279 and ab8226, respectively; Abcam), anti-FLAG (F1804; Sigma-Aldrich), and normal rabbit IgG (Millipore). All stock laboratory chemicals were from Sigma-Aldrich.

Dual luciferase reporter assay. DNA sequences, ~350 bp in length, surrounding selected nonrisk or risk *TNIP1* variants were amplified by polymerase chain reaction (PCR) (the oligonucleotide sequences for each probe are listed in Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract) and cloned into a minimal promoter luciferase plasmid, pGL4.23 (Promega). *CREB1* (OHu22955D) and *BHLHE40*/DEC1 (OHu17520) were purchased from GenScript and cloned into a Flag-tagged overexpression cloning vector, pcDNA3.1/C-(K)DYK. Each plasmid was transiently cotransfected with the pRL-TK plasmid into Jurkat cells (V4XC-1032, Amaxa Nucleofector SE kit), THP-1 cells (V4XC-3032, Amaxa Nucleofector SG kit), or EBV B cells (V4XC-2032, Amaxa Nucleofector SF kit) (all from Lonza). The pRL-TK plasmid was used for normalization and to calculate transfection efficiency. Twenty-four hours after transfection, cells were treated with P/I (50 ng/ml PMA and 500 ng/ml ionomycin) for 2 hours, and then enhancer activity was measured using a dual luciferase reporter assay (Promega) as previously described (29).

Electrophoretic mobility shift assay (EMSA). Complementary pairs of 40-bp nonrisk and risk probes (see Supplementary Table 1 at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract), chemically synthesized and 5′-end biotinylated (IDT), were annealed by heating at 95°C for 5 minutes,

and then cooled to room temperature. Ten micrograms of nuclear protein, which was extracted from Jurkat, EBV B, or THP-1 cells that were either left quiescent or stimulated with P/I (50 ng/ml PMA and 500 ng/ml ionomycin), was incubated with 20 fmoles of biotin–end-labeled probes in binding buffer (1 μg poly [dI-dC], 20 m*M* HEPES, 10% glycerol, 100 mM KCl, and 0.2 m*M* EDTA, pH 7.9) for 20 minutes at room temperature. DNA–protein complexes were resolved on a nondenaturing 5% acrylamide gel for ~60 minutes at 100V in 0.5X Tris borate/EDTA before being transferred onto a positively charged nylon membrane (AM10104; Thermo Fisher Scientific) in 0.5X Tris borate/EDTA at 300 mA for 30 minutes.

Membrane-bound DNA–protein complexes were ultraviolet cross-linked at 120 mJ/cm^2 using a UV Stratalinker 1800 (Stratagene), and then detected by horseradish peroxidase–conjugated streptavidin chemiluminescence (LightShift chemiluminescent EMSA kit 89880; Thermo Fisher Scientific) in accordance with the manufacturer's instructions. Chemiluminescence was captured on radiographic films that were exposed for 1–5 minutes and developed using a Mini-Med 90 X-ray Film Processor (AFP Manufacturing). For competition assays, 10-, 50-, and 200-fold excess of unlabeled nonrisk or risk probes was added to the EMSA binding reactions (for representative images, see Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract). Semiquantitative densitometry was performed using NIH ImageJ software.

DNA-affinity pulldown of nuclear proteins. Streptavidin magnetic beads (200 μg Dynabeads M-280 Streptavidin, 112-06D; Invitrogen) were subjected to 2 rounds of blocking with 1% bovine serum albumin (BSA) in phosphate buffered saline (PBS) for 15 minutes, followed by washing with PBS containing 1*M* NaCl and Tris–EDTA buffer. Biotinylated nonrisk or risk oligonucleotides were bound to 100 μl of the BSA-blocked streptavidin beads by incubating for 30 minutes at room temperature in Tris–EDTA buffer, followed by washing with Tris–EDTA buffer. A biotinylated scrambled oligonucleotide served as a negative control (see Supplementary Table 1 at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract). Fifty micrograms of nuclear extract was precleared by incubating with 100 μl of the BSA-blocked beads in binding buffer (250 m*M* NaCl, 50 m*M* Tris HCl, 50% glycerol, 2.5 m*M* dithiothreitol, 2.5 m*M* EDTA, pH 7.6) containing 15 ng/μl poly(-dI-dC) (product no. 81349-500UG; Sigma-Aldrich), 0.5 μg/ml BSA, and 0.1% Nonidet P40 for 30 minutes on ice. Precleared nuclear extracts were incubated with the oligonucleotide-linked streptavidin beads for 30 minutes in a 37°C water bath. Samples were gently shaken every 5 minutes. Beads were subsequently washed 3 times with binding buffer containing 0.1% Nonidet P40. The proteins were eluted in 50 μl of 0.2% sodium dodecyl sulfate (SDS) sample buffer by boiling for 5 minutes, and then resolved

on a Criterion XT 4%–12% Bis-Tris gel (product no. 3450124; Bio-Rad), followed by Western blotting.

Western blotting. Cells were pelleted and washed in cold PBS and lysed with radioimmunoprecipitation assay lysis buffer (25 m*M* Tris HCl, 150 m*M* NaCl, 5 m*M* EDTA, pH 8, 1% Triton X-100, 0.1% SDS, 0.5% sodium deoxycholate) containing a protease inhibitor cocktail (product no. 539132; EMD Millipore) and Halt phosphatase inhibitors (product no. 1862495; Thermo Fisher Scientific). Cells were lysed for 15 minutes on ice, followed by syringe lysis with a 27-gauge needle, and then cleared by centrifugation at maximum speed for 20 minutes.

Protein concentrations were determined using a Qubit protein assay kit (product no. Q33212; Thermo Fisher Scientific). Proteins were denatured in 2X SDS loading buffer by heating to 95°C for 5 minutes, and then separated by SDS–polyacrylamide gel electrophoresis, transferred to a PVDF membrane (product no. 1620177; Bio-Rad), blocked with 5% nonfat dairy milk, and analyzed by Western blotting using the indicated antibodies. Proteins were detected using Pierce enhanced chemiluminescence Western blotting substrate (product no. 32106; Thermo Fisher Scientific) and visualized using a ChemiDocMP imaging system (Bio-Rad).

Chromatin immunoprecipitation (ChIP) and real-time quantitative reverse transcription–PCR (qRT-PCR). ChIP assays were performed using a Covaris truChIP chromatin shearing kit and Magna ChIP protein A/G beads (MilliporeSigma), in accordance with the manufacturer's recommendations. In brief, 1×10^7 EBV B cells carrying the *TNIP1* rs10036748 nonrisk or risk alleles were treated with P/I (50 ng/ml PMA and 500 ng/ml ionomycin) in growth medium for 2 hours, and then crosslinked with 1% formaldehyde. Nuclei were isolated and sonicated in 1 ml of lysis buffer with a Covaris S1 sonicator. Chromatin–protein complexes (500 μl) were immunoprecipitated overnight at 4°C by mild agitation with antibodies specific for CREB-1, bHLHe40/DEC1, or normal rabbit IgG (as a negative control). DNA was eluted from the immunoprecipitated chromatin complexes, reverse-crosslinked, purified using Agencourt AMPure XP beads (Beckman Coulter), and subjected to qRT-PCR analysis using RT2 SYBR Green (Qiagen) and primers neighboring the rs10036748 variant (see Supplementary Table 1 at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract).

HiChIP. H3K27ac-mediated chromatin interactions were measured for the whole genome of EBV B cells as part of a previously published study (30). HiChIP raw reads (fastq files) were aligned to those of the hg19 human reference genome, using HiC-Pro (31). Aligned data were processed and analyzed through the hichipper data preprocessing pipeline (32). Loops were derived from the linked paired-end reads that overlap with anchors, and

findings were analyzed and visualized using the R package diffloop (33). HiChIP sequencing data have been made available in the Gene Expression Omnibus database (accession no. GSE116193).

RNA extraction and real-time qRT-PCR. Total RNA from EBV B cells carrying the nonrisk or risk *TNIP1* H1 haplotype was isolated using a Direct-zol RNA MiniPrep Plus kit (Zymo Research), in accordance with the manufacturer's protocol (3). Synthesis of complementary DNA was performed using a QuantiTect reverse transcriptase kit (Qiagen) in accordance with the manufacturer's recommendations. Gene expression was measured by real-time PCR analysis using RT2 SYBR Green (Qiagen). Gene expression primers for human *ANXA6* (QT00066941), human *DCTN4* (QT00038766), human *GM2A* (QT00071967), human *SMIM3* (QT01028090), human *TNIP1* (QT00044072), and human *GAPDH* (PPH00150F) were purchased from Qiagen.

RESULTS

Role of SLE risk variants in collectively suppressing the expression of *TNIP1*. We used RegulomeDB, which collates bioinformatics data from ENCODE and other sources, to prioritize variants on the SLE-associated risk haplotypes that are most likely to be functional regulators of *TNIP1* expression (defined as those with a RegulomeDB score ≤3a) (34). Of the 50 variants on the risk haplotype (3), RegulomeDB identified 11 that were located in non–protein-coding regions enriched for H3K4me1 (marker of a poised enhancer) and H3K27ac (marker of an active enhancer) and ChIP-seq transcription factor binding sites in several cell lines,

as well as peripheral blood mononuclear cells and primary CD19+ B cells (Figure 1; see also Supplementary Figure 2 and Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract).

To begin functionally characterizing the allelic and cell type–specific regulatory potential of these variants, we created a series of luciferase expression vectors by PCR amplification of an ~350-bp sequence containing the respective nonrisk or risk alleles from EBV B cells carrying the nonrisk or SLE risk haplotype, and then cloning it into a minimal promoter luciferase vector designed to measure the activity of upstream DNA elements using a dual luciferase assay (29). We transiently transfected the luciferase constructs into B lymphoid (EBV B), monocytoid (THP-1), or T lymphoid (Jurkat) lineage cells for 24 hours. Thereafter, we measured luciferase activity in the cells in a resting state or after 2 hours of stimulation with P/I. After normalization to the vector-only control, an increase in luciferase activity over that of the vector-only control was interpreted as increased enhancer activity. Cell type–specific increases in enhancer activity were observed for 10 variants in EBV B cells (Figure 2A), 9 variants in THP-1 cells (Figure 2B), and 5 variants in Jurkat cells (Figure 2C).

Of the 10 variants demonstrating enhancer activity in EBV B cells, 8 showed significant allele-specific effects (Figure 2A). Three variants (rs10057690, rs13180950, and rs10036748) retained the allelic effect irrespective of whether the cells remained quiescent or were stimulated with P/I. Only 1 variant, rs2042234, produced a significant allelic effect in stimulated EBV B cells, while 3 others (rs4958879, rs4958435, and rs2233287) lost evidence for allele-specific activity with stimulation. The rs62383767 variant flipped
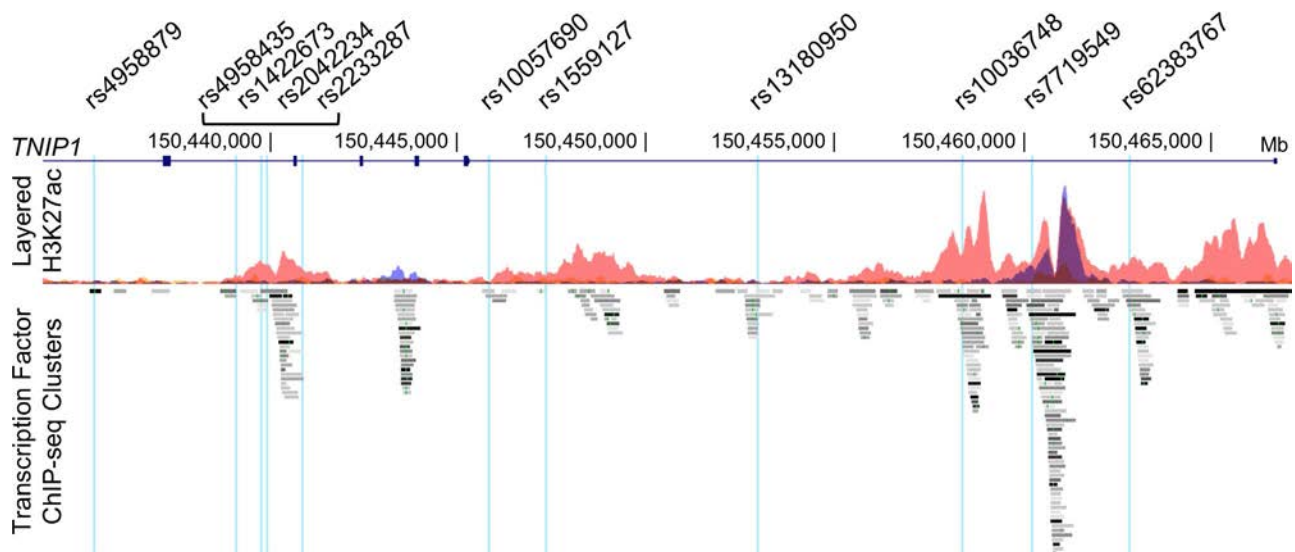
Figure 1. Bioinformatics analysis of the 11 systemic lupus erythematosus (SLE) risk variants located in regulatory elements of the *TNIP1* locus. RegulomeDB was used to identify the 11 non–protein-coding SLE risk variants on the *TNIP1* locus that were predicted to have regulatory functions, and that are positioned in enhancer regions identified by enrichment of H3K27ac marks and transcription factor binding according to chromatin immunoprecipitation sequencing (ChIP-seq). Variant positions are indicated on the UCSC Genome Browser tracks: Gene Symbol, custom track of the *TNIP1* SLE risk single-nucleotide polymorphisms, ENCODE H3K27Ac chromatin marks for GM12878, H1-hESC, and K562 cell lines, and ENCODE ChIP-seq transcription factor enrichment.
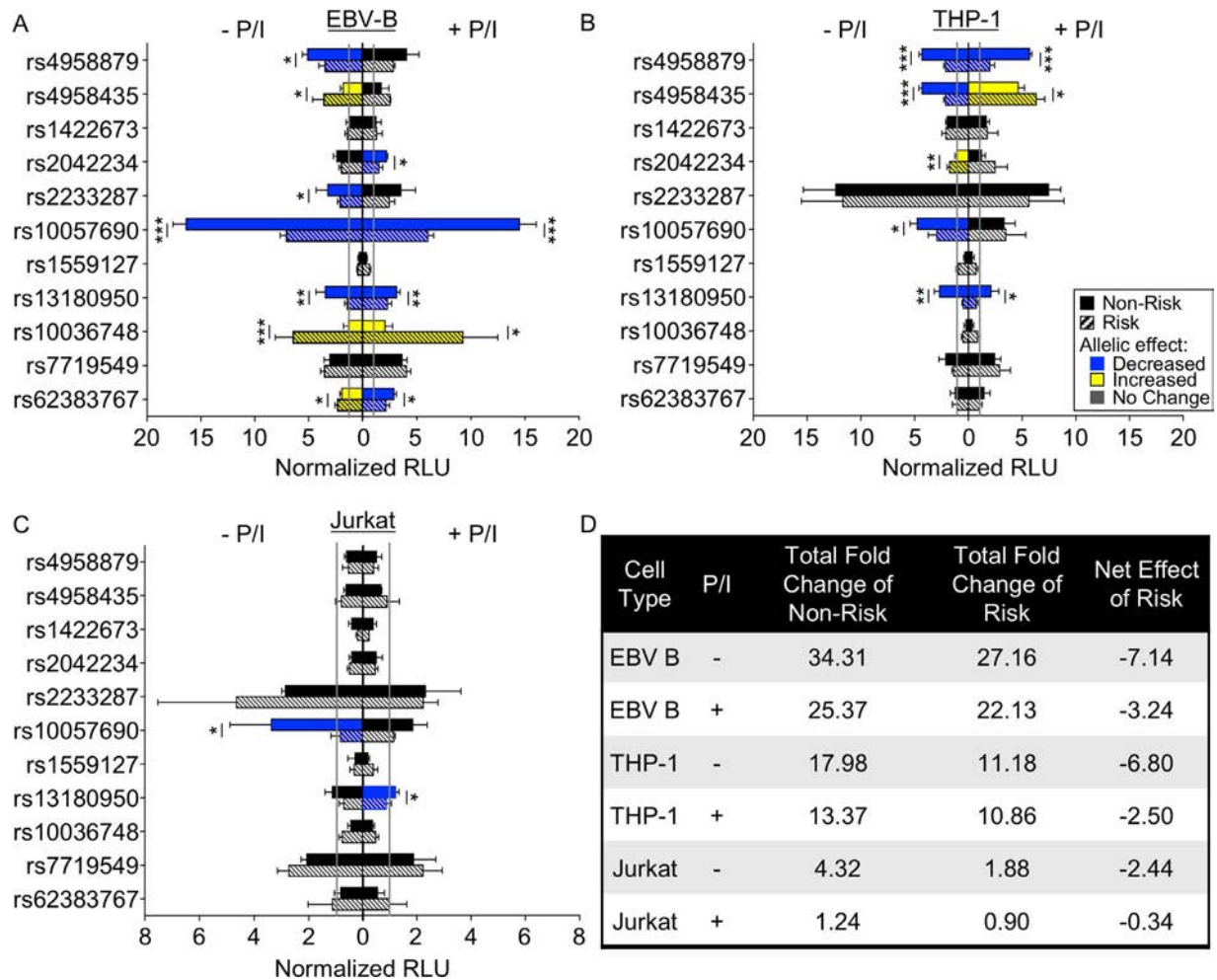
Figure 2. The *TNIP1* locus has a complex regulatory mechanism with cell type–specific and stimulation-dependent regulatory elements. **A–C**, Sequences carrying the nonrisk alleles (solid bars) or risk alleles (hatched bars) of the indicated variants were cloned into a luciferase vector with a minimal promoter. Luciferase activity was measured after transient transfection in Epstein-Barr virus (EBV)–transformed B cells (**A**), THP-1 cells (**B**), or Jurkat cells (**C**), either at rest (left) or after stimulation with phorbol 12-myristate 13-acetate/ionomycin (P/I) (right). Luciferase activity was normalized to the values for the vector-only control (vertical gray lines), with values presented as the normalized relative luciferase units (RLU). Luciferase activity in the risk allele relative to the nonrisk allele was assessed as either a significant decrease (blue) or significant increase (yellow) or a lack of allelic effect (black). Bars show the mean ± SEM. * = *P* < 0.05; ** = *P* < 0.01; *** = *P* < 0.001 for risk versus nonrisk, by Student's *t*-test (n > 3). **D**, Summary analysis of the *TNIP1* haplotype effect on luciferase activity in EBV B, THP-1, and Jurkat cells with or without stimulation with P/I. The total fold change in luciferase activity for nonrisk or risk alleles over vector-only control were calculated for all variants exhibiting significant allelic effects, and then the cumulative effect was estimated by subtracting the total nonrisk effect from the total risk effect, to yield the net effect of risk.

from increased to decreased luciferase activity from the risk allele following stimulation. Interestingly, the 2 variants with the strongest enhancer effect exhibited opposing allelic effects, in which the risk allele of rs10057690 significantly reduced enhancer activity, and the risk allele of rs10036748 significantly increased enhancer activity.

Of the 9 variants demonstrating enhancer activity in THP-1 cells (Figure 2B), 5 demonstrated significant allelic effects on luciferase activity. Risk alleles of rs4958879 and rs13180950 significantly impaired enhancer activity irrespective of whether the cells remained quiescent or were stimulated with P/I. The increased enhancer activity of the rs2042234 risk allele and the decreased

enhancer activity of the rs10057690 risk allele were observed only in unstimulated cells. Stimulation of THP-1 cells flipped the allelic effect of rs4958435, increasing the luciferase activity from the risk allele over that from the nonrisk allele. Only 5 variants in Jurkat cells exhibited enhancer effects, with 2 variants, rs10057690 and rs13180950, showing reduced luciferase activity for the risk alleles, with, however, opposing stimulation dependence (Figure 2C).

In order to summarize the cumulative effect of these variants in each cell line, either at rest or following stimulation, we calculated the difference in the sum of the fold change in luciferase activity between the risk and nonrisk variants in the cells under each condition (Figure 2D). EBV B and THP-1 cells demonstrated

the strongest net effect of the risk alleles in attenuating luciferase activity (net effect of risk −7.14 fold and −6.80 fold, respectively) compared with that found in Jurkat cells (net effect of risk −2.44 fold), suggesting that the SLE risk haplotype is most potent in suppressing *TNIP1* expression in B lymphoid and monocytoid cell types. Stimulation with P/I augmented to some degree the luciferase expression from the risk allele in each cell type, but not enough to reverse the cumulative suppression of luciferase activity. Overall, these results are consistent with the observations of hypomorphic *TNIP1* gene expression in vivo in carriers of the *TNIP1* risk haplotype.

Involvement of SLE risk variants in altering nuclear protein complex binding at the *TNIP1* locus. We performed EMSAs using extracts from Jurkat, THP-1, or EBV B cells, either at rest or stimulated with P/I, to determine whether the risk alleles of the 11 variants altered the binding affinities of nuclear protein complexes. In all 3 cell types, nearly all of the variants that exhibited binding demonstrated a qualitative loss of nuclear protein complex binding to the risk allele (see Supplemen-

tary Figures 3–5, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract).

Densitometry was used to obtain semiquantitative estimates of the nuclear protein complex binding for each cell type, in either resting or stimulated conditions. Any variants that did not demonstrate detectable binding were omitted from the semiquantitative analysis. Only 2 variants (rs4958879 and rs10036748) exhibited a significant loss of binding to the risk allele in unstimulated Jurkat cells (Figure 3A and Supplementary Figure 3 [http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract]). The allelic effect of rs4958879 was lost with stimulation of Jurkat cells (Figure 3B and Supplementary Figure 3 [http://online library.wiley.com/doi/10.1002/art.41188/abstract]). In THP-1 cells, no significant allelic effects were observed (Figures 3C and D and Supplementary Figure 4 [http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract]).

The nuclear extracts from EBV B cells demonstrated the most robust allele-specific effects, with 3 risk alleles (rs10057690, rs13180950, and rs10036748) significantly reducing binding of nuclear proteins from unstimulated cells (Figure 3E and
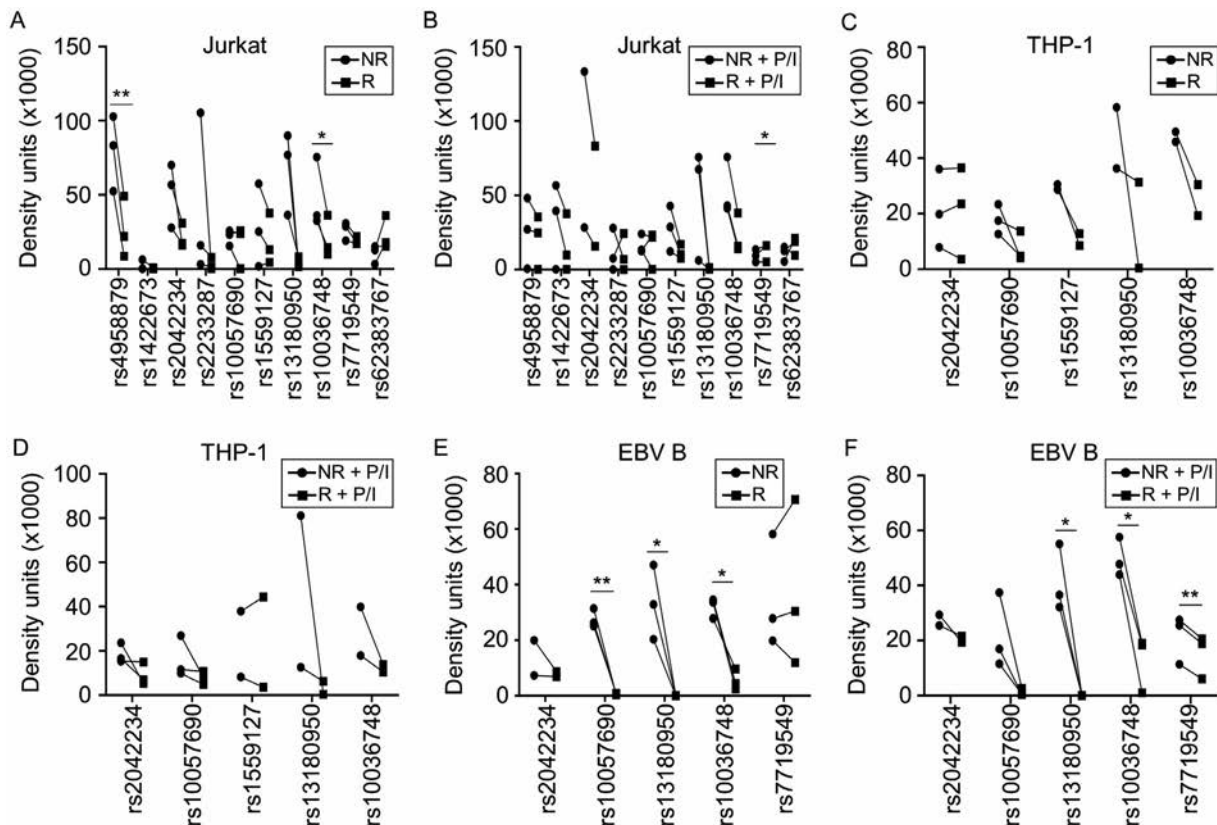
Figure 3. *TNIP1* systemic lupus erythematosus (SLE) risk alleles differentially affect nuclear complex binding in immune cells. Electrophoretic mobility shift assays were performed using biotinylated oligonucleotides containing the nonrisk (NR) or risk (R) alleles of the indicated variants. Nuclear extracts were from Jurkat cells (**A** and **B**), THP-1 cells (**C** and **D**), or Epstein-Barr virus (EBV)–transformed B cells (**E** and **F**), either at rest (**A**, **C**, and **E**) or after stimulation with phorbol 12-myristate 13-acetate/ionomycin (P/I) (**B**, **D**, and **F**). Representative images are shown in Supplementary Figures 3–5 (http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract). * = *P* < 0.05; ** = *P* < 0.01 for risk versus nonrisk, by paired *t*-test.

Supplementary Figure 5 [http://onlinelibrary.wiley.com/doi/10.10 02/art.41188/abstract]). Stimulation of EBV B cells resulted in a loss of allele-specific significance for the rs10057690 risk variant, but produced a gain of allele-specific significance for the risk allele of rs7719549 (Figure 3F and Supplementary Figure 5 [http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract]). Taken together, the combined results of the luciferase and EMSA assays suggest that the functional potency of the *TNIP1* risk haplotype is most active in B lymphoid lineage cells.

Impairment of nuclear factor binding to the SLE risk allele rs10036748. In EBV B cells, 2 variants, rs10057690 and rs13180950, showed concordant loss of nuclear protein binding and reduced luciferase activity for the risk alleles relative to the nonrisk alleles (Figures 2A and 3E and F). In contrast, rs10036748 showed discordant effects, with the risk allele showing loss of binding in the EMSAs, but increased luciferase activity (Figure 2A and Figures 3E and F), suggesting that this allele may oppose the overall reduction in *TNIP1* gene expression imparted by the risk haplotype in vivo.

We hypothesized that the loss of binding of a transcription inhibitor complex was responsible for the enhanced allele-specific luciferase activity conferred by the rs10036748 risk (A) allele. To test this, we used bioinformatics analyses to identify nuclear proteins that could be predicted to bind to the rs10036748 variant. Based on bioinformatics data from the ENCODE and Genomatix databases, we identified Egr-1, bHLHe40/DEC1, and CREB-1 as relevant nuclear factor binding proteins. All 3 of these are transcription factors reportedly involved in the activation, proliferation, differentiation, and/or function of different immune cell subtypes (35–40). In addition, both bHLHe40/DEC1 and CREB-1 have been reported to function as transcriptional repressors in specific cell types and states. The bHLHe40/DEC1 transcription factor is an essential transcriptional repressor of interleukin-10 production both in Th1 cells and in mice infected with *Mycobacterium tuberculosis* (41–43), and CREB-1 reportedly competes against CBP/p300 to impair NF-κB signaling in germinal B cells (40,44).

We performed DNA-affinity pulldowns using nuclear extracts from Jurkat, EBV B, and THP-1 cells to determine the allele-specific binding of Egr-1, bHLHe40/DEC1, and CREB-1 to the rs10036748 variant. Egr-1 from Jurkat cell nuclear extracts exhibited reduced binding to the risk (A) allele, relative to the nonrisk (G) allele, of rs10036748 (Figure 4A). We did not observe allele-specific binding by Egr-1 in EBV B cells, and THP-1 nuclear extracts did not pull down detectable levels of Egr-1 (Figure 4A). In contrast, bHLHe40/DEC1 and CREB-1 exhibited reduced binding to the risk (A) allele in all 3 cell types (Figure 4A).

To confirm the allele-specific binding of bHLHe40/DEC1 and CREB-1 in situ, we performed ChIP–qPCR in EBV B cells carrying the nonrisk or risk *TNIP1* haplotype. Both bHLHe40/DEC1 and CREB-1 showed decreased binding to the risk (A) allele relative to
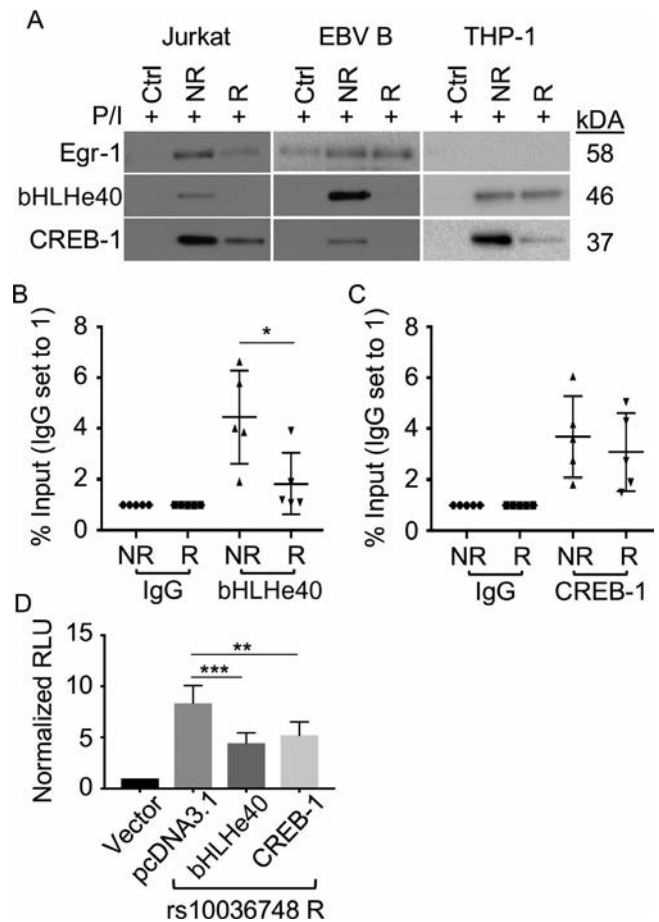
Figure 4. The systemic lupus erythematosus (SLE) risk allele rs10036748 reduces binding of early growth response 1 (Egr-1), basic helix-loop-helix family member 40/differentially expressed in chondrocytes 1 (bHLHe40/DEC1), and CREB-1. **A**, DNA-affinity pulldown assays using annealed biotinylated oligonucleotides containing the nonrisk (NR) or risk (R) allele of rs10036748 were performed with nuclear extracts from Jurkat cells, Epstein-Barr virus (EBV)–transformed B cells, or THP-1 cells that had been stimulated with phorbol 12-myristate 13-acetate/ionomycin (P/I) for 2 hours. Eluted proteins were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and analyzed by Western blotting using specific antibodies against each protein. A biotinylated scrambled oligonucleotide served as a negative control. Representative images from 1 of 3 experiments are shown. **B** and **C**, Chromatin immunoprecipitation followed by quantitative reverse transcription–polymerase chain reaction (ChIP–qPCR) was performed in P/I-stimulated homozygous EBV B cell lines carrying the nonrisk or risk allele of rs10036748. The cells were analyzed by ChIP–qPCR using specific antibodies against bHLHe40/DEC1 (**B**) or CREB-1 (**C**) (n = 5 cell lines per genotype). Rabbit IgG was used as an isotype control. **D**, Luciferase activity was measured in EBV B cells transiently transfected with Firefly luciferase vector-only control or cotransfected with luciferase vector carrying the rs10036748 risk allele and pcDNA3.1, pcDNA3.1–bHLHe40/DEC1, or pcDNA3.1–CREB-1. Luciferase activity was normalized to the values for the vector-only control, with values presented as the normalized relative luciferase units (RLU). In **B–D**, symbols represent individual samples; horizontal lines with bars show the mean ± SEM. * = P < 0.05; ** = P < 0.01; *** = P < 0.001 by Student's t-test.

the nonrisk (G) allele (Figures 4B and C), but only the reduction in bHLHe40/DEC1 binding reached significance (Figure 4B). Collectively, these results suggest that the risk (A) allele of rs10036748 impairs nuclear protein binding, particularly bHLHe40/DEC1, in immune-relevant cell lines.

In DNA-affinity pulldowns and ChIP–qPCR assays, both bHLHe40/DEC1 and CREB-1 demonstrated reduced binding to the risk allele, and both are reported to function as transcription repressors. Therefore, we reasoned that the increased luciferase activity observed for the rs10036478 risk allele may result from an allele-dependent loss of these transcriptional inhibitors. Furthermore, we hypothesized that the overexpression of either bHLHe40/DEC1 or CREB-1 could overcome the lower affinity binding to the rs10036748 risk allele, thereby suppressing the allelic increase in luciferase activity. Indeed, transient overexpression of either bHLHe40/DEC1 or CREB-1 in EBV B cells

expressing the luciferase vector carrying the rs10036748 risk allele significantly reduced the allele-specific increase in luciferase activity (Figure 4D and Supplementary Figure 6, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract).

Influence of the *TNIP1* locus on gene expression locally and at a distance. Enhancers can regulate gene expression locally or at a distance through long-range DNA looping (45). We therefore hypothesized that the hypomorphic effect of the *TNIP1* risk haplotype might also manifest an effect on other genes within its 3-D chromatin network. To define the 3-D chromatin network of the *TNIP1* locus, we performed HiChIP in EBV B cells with immunoprecipitation of chromatin using antibodies for H3K27ac (30). Several long-range interactions were observed between the enhancer spanning the *TNIP1* risk haplotype and
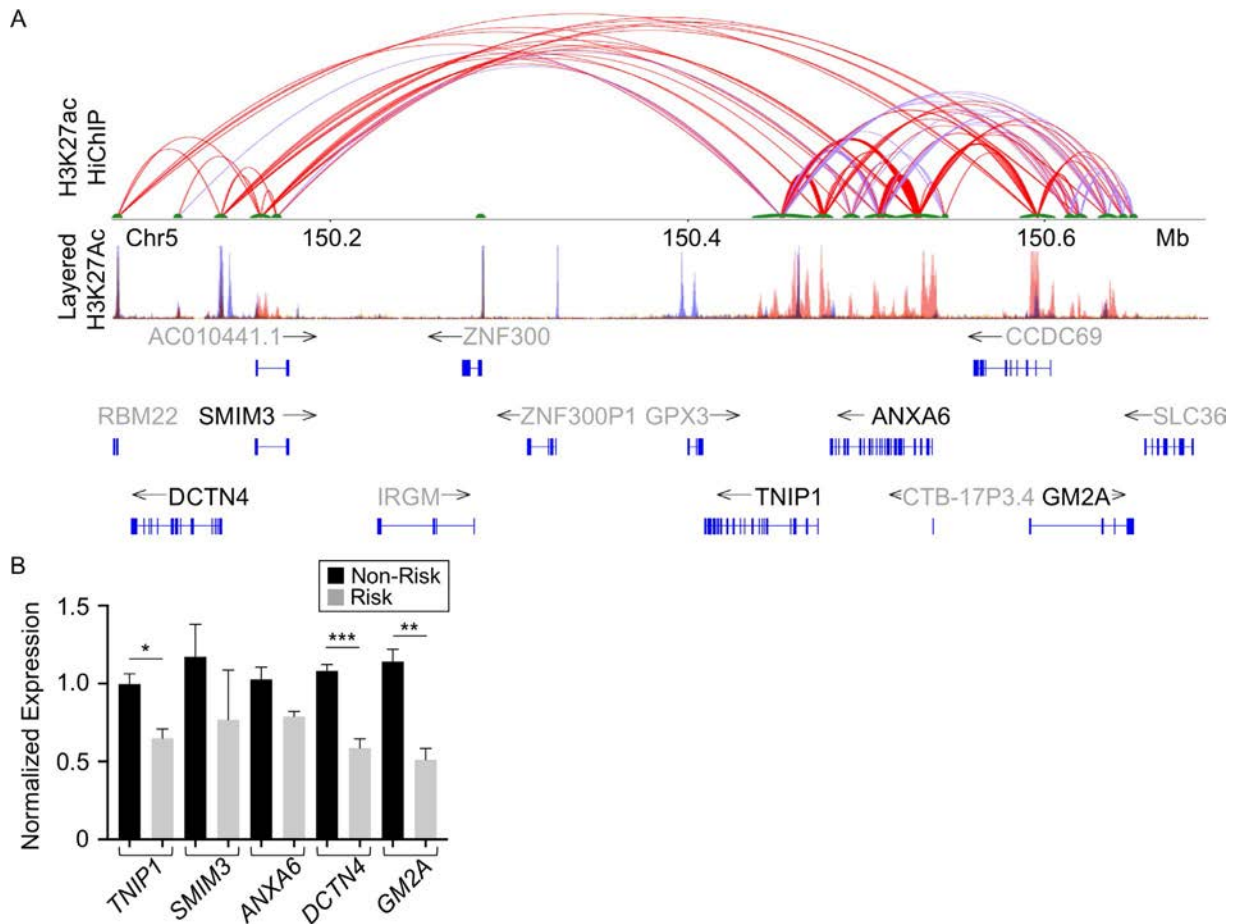
Figure 5. The *TNIP1* locus has a complex regulatory structure that influences gene expression locally and at a distance. **A,** H3K27ac HiChIP looping interactions within the *TNIP1* region were visualized as a 2-dimensional looping diagram. Thickness of the arc is proportional to the frequency of observed paired-end tags (PET scan threshold of 6). The enhancer with a strong H3K27ac chromatin immunoprecipitation (ChIP) peak spanning the *TNIP1* gene locus forms long-range interactions with multiple H3K27ac anchors proximal to distant genes. H3K27Ac Peak Track was adapted from the UCSC Genome Browser ENCODE H3K27Ac chromatin marks for GM12878, H1-hESC, and K562 cell lines. **B,** Expression of distant genes in the H3K27ac regulatory network with *TNIP1* was measured in resting Epstein-Barr virus (EBV)–transformed B cell lines carrying the nonrisk or risk H1 *TNIP1* haplotype using quantitative reverse transcription–polymerase chain reaction. Expression values were normalized to those for *GAPDH*, and then to *TNIP1* expression in nonrisk EBV B cell controls. Results are the mean ± SEM (n > 3). * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$ by Student's *t*-test.

flanking genes known to be expressed in immune cell types (see www.gtexportal.org) (46), including *SMIM3*, *ANXA6*, *DCTN4*, and *GM2A* (Figure 5A). Similar chromatin looping events were also observed in publicly available Hi-C data from the GM12878 EBV B cell line (see Supplementary Figure 7, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41188/abstract) (47,48).

To determine whether the *TNIP1* risk haplotype influenced the expression of these distant genes, we performed qRT-PCR using RNA isolated from resting EBV B cells carrying the *TNIP1* nonrisk or risk H1 haplotype. As expected, the *TNIP1* SLE risk haplotype significantly reduced *TNIP1* expression (Figure 5B). EBV B cells carrying the *TNIP1* risk haplotype also exhibited a significant loss of *DCTN4* and *GM2A* expression. Although showing a similar trend, no significant differences in the expression of *SMIM3* and *ANXA6* in the presence of the *TNIP1* risk haplotype, relative to the nonrisk haplotype, were observed. Collectively, these data suggest that the effect of the *TNIP1* risk haplotype extends beyond *TNIP1* to include other genes whose promoters are contacted by the *TNIP1* haplotype through long-range DNA interactions.

DISCUSSION

Despite the fact that associations between the *TNIP1* SLE risk haplotype, hypomorphic *TNIP1* expression, and dysregulation of key inflammatory signaling pathways have been reported, few studies have investigated the regulatory mechanisms controlling *TNIP1* expression in the context of the genetic risk haplotype. Using a combination of in vitro assays aimed at characterizing the *TNIP1* risk variants that are predicted to have functional effects, we discovered a complex regulatory mechanism whereby the cumulative loss of nuclear protein complex binding and loss of enhancer activity at specific risk alleles drives the hypomorphic expression of *TNIP1*. Moreover, both EMSA and luciferase assays revealed the most significant and reproducible effects in EBV B cells, suggesting that the *TNIP1* risk haplotype may be most potent in cells of B lymphoid lineage.

With respect to the effects of stimulation on enhancer activity and allele-specific effects, the luciferase assays provided the most definitive information. Our data demonstrated that stimulation with P/I across all 3 cell types alleviated the suppressive effects of the risk alleles on luciferase activity (Figure 2D), but not enough to completely reverse the suppressive effect of the *TNIP1* haplotype. This suggests that the magnitude of suppression of *TNIP1* expression by the *TNIP1* risk haplotype can be modulated by the overall activation of cellular pathways. Since we used a cell surface receptor–independent activation method, we cannot reach a conclusion as to how the activation of pathways downstream of cell surface receptors, such as the TNF receptor and TLRs, might alter the suppressive effect of the *TNIP1* risk alleles.

We chose to further study the *TNIP1* risk variant, rs10036748, which is an index single-nucleotide polymorphism for this locus in multiple SLE GWAS and is shared among several risk haplotypes

and across multiple racial groups, including the H1 and H2 haplotypes in European Americans and the Block 2 haplotype in the Han Chinese (3–9), because it exhibited an allele-specific increase in enhancer activity, but a discordant reduction in nuclear protein complex binding. This illustrates how genetic variants modulate gene expression through not only loss of binding of transcription-activating proteins, but also loss of binding of transcriptional repressor proteins. Our affinity pulldown and ChIP–qPCR data suggest that the rs10036748 risk allele promotes enhancer activity by reducing the binding affinity of bHLHe40/DEC1 and CREB-1. Forced expression of bHLHe40/DEC1 and CREB-1 can overcome the low affinity binding of the risk allele, returning the luciferase activity to the level of the nonrisk allele. Taken together, these results suggest that the risk allele of rs10036748 opposes the hypomorphic expression of the *TNIP1* risk haplotype by lowering the binding affinity of bHLHe40/DEC1 and CREB-1 transcriptional repressors. Further studies using genetic engineering methods that swap the risk and nonrisk alleles at rs10036748 will be necessary to confirm our observations.

Another interesting observation from our study is that the suppression of gene transcription induced by the *TNIP1* risk haplotype is not limited to *TNIP1*. Within the 3-D chromatin network defined by H3K27ac HiChIP data, we observed significant reductions in the transcription of *DCTN4* and *GM2A*. Similar trends were also observed for *SMIM3* and *ANXA6,* but these reductions did not reach statistical significance. These results suggest that the functional impact of the *TNIP1* risk haplotype likely extends beyond its effect on *TNIP1* expression toward pathways involving suppressed expression of *DCTN4* and *GMA2*. For example, *DCTN4* missense variants have recently been shown to be associated with shorter time to development of chronic *Pseudomonas aeruginosa* infection, a condition associated with worse long-term pulmonary disease and survival in patients with cystic fibrosis (49). Patients with cystic fibrosis who are biallelic for *DCTN4* missense variants and the *TNIP1* risk haplotype could be at risk of even earlier infection, due to the reduced expression of wild-type *DCTN4* in the presence of the *TNIP1* risk haplotype.

Our study has some limitations that must be considered in the interpretation of our data. First, the EMSA and luciferase assays do not assess function of the risk alleles in their natural genomic context. Second, binding of proteins to the probes used in both assays is dependent, to some extent, on the length of the probe and the sequences contained therein. Third, we measured the function of risk variants in immortalized cell lines. Both the Jurkat and THP-1 cell lines are derived from spontaneous malignant transformations, whereas EBV B cells are virally transformed; the molecular composition of these cell lines is not expected to fully reflect the precise molecular composition of the primary cells from which they were originally derived.

Despite these limitations, the data presented herein demonstrate that the majority of the SLE risk alleles reduced nuclear protein complex binding with variable allelic effects on enhancer

activity (Figure 2D), resulting in an overall affect that is consistent with the observations of hypomorphic *TNIP1* expression. Therefore, we are confident that our data provide a reasonable approximation for how the *TNIP1* risk haplotype operates in primary immune cells in patients with 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 published. Dr. Gaffney 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. Pasula, Tessneer, Fu, Gaffney.
Acquisition of data. Pasula, Fu, Gopalakrishnan, Pelikan, Kelly, Gaffney.
Analysis and interpretation of data. Pasula, Tessneer, Fu, Gopalakrishnan, Pelikan, Kelly, G. B. Wiley, M. M. Wiley, Gaffney.

REFERENCES

1. Chen L, Morris DL, Vyse TJ. Genetic advances in systemic lupus erythematosus. Curr Opin Rheumatol 2017;29:423–33.

2. Deng Y, Tsao BP. Updates in lupus genetics. Curr Rheumatol Rep 2017;19:68.

3. Adrianto I, Wang S, Wiley GB, Lessard CJ, Kelly JA, Adler AJ, et al. Association of two independent functional risk haplotypes in TNIP1 with systemic lupus erythematosus. Arthritis Rheum 2012;64:3695–705.

4. Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, Hu Z, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat Genet 2009;41:1234–7.

5. He CF, Liu YS, Cheng YL, Gao JP, Pan TM, Han JW, et al. TNIP1, SLC15A4, ETS1, RasGRP3 and IKZF1 are associated with clinical features of systemic lupus erythematosus in a Chinese Han population. Lupus 2010;19:1181–6.

6. Zhang DM, Cheng LQ, Zhai ZF, Feng L, Zhong BY, You Y, et al. Single-nucleotide polymorphism and haplotypes of TNIP1 associated with systemic lupus erythematosus in a Chinese Han population. J Rheumatol 2013;40:1535–44.

7. Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, Sun X, et al. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat Genet 2009;41:1228–33.

8. Kawasaki A, Ito S, Furukawa H, Hayashi T, Goto D, Matsumoto I, et al. Association of TNFAIP3 interacting protein 1, TNIP1 with systemic lupus erythematosus in a Japanese population: a case-control association study. Arthritis Res Ther 2010;12:R174.

9. Liu Z, Yu Y, Yue Y, Hearth-Holmes M, Lopez PD, Tineo C, et al. Genetic alleles associated with SLE susceptibility and clinical manifestations in Hispanic patients from the Dominican Republic. Curr Mol Med 2019;19:164–71.

10. Lessard CJ, Li H, Adrianto I, Ice JA, Rasmussen A, Grundahl KM, et al. Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjögren's syndrome. Nat Genet 2013;45:1284–92.

11. Allanore Y, Saad M, Dieudé P, Avouac J, Distler JH, Amouyel P, et al. Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. PLoS Genet 2011;7:e1002091.

12. Bossini-Castillo L, Martin JE, Broen J, Simeon CP, Beretta L, Gorlova OY, et al. Confirmation of TNIP1 but not RHOB and PSORS1C1 as systemic sclerosis risk factors in a large independent replication study. Ann Rheum Dis 2013;72:602–7.

13. Bowes J, Orozco G, Flynn E, Ho P, Brier R, Marzo-Ortega H, et al. Confirmation of TNIP1 and IL23A as susceptibility loci for psoriatic arthritis. Ann Rheum Dis 2011;70:1641–4.

14. Chen Y, Yan H, Song Z, Chen F, Wang H, Niu J, et al. Down-regulation of TNIP1 expression leads to increased proliferation of human keratinocytes and severer psoriasis-like conditions in an imiquimod-induced mouse model of dermatitis. PLoS One 2015;10:e0127957.

15. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, et al, for the Collaborative Association Study of Psoriasis. Genome-wide scan reveals association of psoriasis with IL-23 and NF-κB pathways [letter]. Nat Genet 2009;41:199–204.

16. Yan KX, Zhang YJ, Han L, Huang Q, Zhang ZH, Fang X, et al. TT genotype of rs10036748 in TNIP1 shows better response to methotrexate in a Chinese population: a prospective cohort study. Br J Dermatol 2019;181:778–85.

17. Li X, Ampleford EJ, Howard TD, Moore WC, Torgerson DG, Li H, et al. Genome-wide association studies of asthma indicate opposite immunopathogenesis direction from autoimmune diseases. J Allergy Clin Immunol 2012;130:861–8.

18. Li C, Zhao Z, Zhou J, Liu Y, Wang H, Zhao X. Relationship between the TERT, TNIP1 and OBFC1 genetic polymorphisms and susceptibility to colorectal cancer in Chinese Han population. Oncotarget 2017;8:56932–41.

19. Cheng Y, Jiang X, Jin J, Luo X, Chen W, Li Q, et al. TNIP1 polymorphisms with the risk of hepatocellular carcinoma based on chronic hepatitis B infection in Chinese Han population. Biochem Genet 2019;57:117–28.

20. Liu Z, Shi Y, Na Y, Zhang Q, Cao S, Duan X, et al. Genetic polymorphisms in TNIP1 increase the risk of gastric carcinoma. Oncotarget 2016;7:40500–7.

21. G'Sell RT, Gaffney PM, Powell DW. A20-binding inhibitor of NF-κB activation 1 is a physiologic inhibitor of NF-κB: a molecular switch for inflammation and autoimmunity [review]. Arthritis Rheumatol 2015;67:2292–302.

22. Shamilov R, Aneskievich BJ. TNIP1 in autoimmune diseases: regulation of Toll-like receptor signaling [review]. J Immunol Res 2018;2018:3491269.

23. Flores AM, Gurevich I, Zhang C, Ramirez VP, Devens TR, Aneskievich BJ. TNIP1 is a corepressor of agonist-bound PPARs. Arch Biochem Biophys 2011;516:58–66.

24. Gurevich I, Zhang C, Francis N, Struzynsky CP, Livings SE, Aneskievich BJ. Human TNFα-induced protein 3-interacting protein 1 (TNIP1) promoter activation is regulated by retinoic acid receptors. Gene 2013;515:42–8.

25. Qian T, Chen F, Shi X, Li J, Li M, Chen Y, et al. Upregulation of the C/EBP β LAP isoform could be due to decreased TNFAIP3/TNIP1 expression in the peripheral blood mononuclear cells of patients with systemic lupus erythematosus. Mod Rheumatol 2017;27:657–63.

26. Kuriakose J, Redecke V, Guy C, Zhou J, Wu R, Ippagunta SK, et al. Patrolling monocytes promote the pathogenesis of early lupus-like glomerulonephritis. J Clin Invest 2019;129:2251–65.

27. Caster DJ, Korte EA, Nanda SK, McLeish KR, Oliver RK, G'Sell RT, et al. ABIN1 dysfunction as a genetic basis for lupus nephritis. J Am Soc Nephrol 2013;24:1743–54.

28. Rasmussen A, Sevier S, Kelly JA, Glenn SB, Aberle T, Cooney CM, et al. The Lupus Family Registry and Repository. Rheumatology (Oxford) 2011;50:47–59.

29. Wang S, Wen F, Wiley GB, Kinter MT, Gaffney PM. An enhancer element harboring variants associated with systemic lupus erythematosus engages the TNFAIP3 promoter to influence A20 expression. PLoS Genet 2013;9:e1003750.

30. Pelikan RC, Kelly JA, Fu Y, Lareau CA, Tessneer KL, Wiley GB, et al. Enhancer histone-QTLs are enriched on autoimmune risk haplotypes and influence gene expression within chromatin networks. Nat Commun 2018;9:2905.

31. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol 2015;16:259.

32. Lareau CA, Aryee MJ. Hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data [letter]. Nat Methods 2018;15:155–6.

33. Lareau CA, Aryee MJ. Diffloop: a computational framework for identifying and analyzing differential DNA loops from sequencing data. Bioinformatics 2018;34:672–4.

34. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012;22:1790–7.

35. Bencheikh L, Diop MK, Rivière J, Imanci A, Pierron G, Souquere S, et al. Dynamic gene regulation by nuclear colony-stimulating factor 1 receptor in human monocytes and macrophages. Nat Commun 2019;10:1935.

36. Kharbanda S, Nakamura T, Stone R, Hass R, Bernstein S, Datta R, et al. Expression of the early growth response 1 and 2 zinc finger genes during induction of monocytic differentiation. J Clin Invest 1991;88:571–7.

37. Ye Y, Liu M, Tang L, Du F, Liu Y, Hao P, et al. Iguratimod represses B cell terminal differentiation linked with the inhibition of PKC/EGR1 axis. Arthritis Res Ther 2019;21:92.

38. Mora-López F, Pedreño-Horrillo N, Delgado-Pérez L, Brieva JA, Campos-Caro A. Transcription of PRDM1, the master regulator for plasma cell differentiation, depends on an SP1/SP3/EGR-1 GC-box. Eur J Immunol 2008;38:2316–24.

39. Dinkel A, Warnatz K, Ledermann B, Rolink A, Zipfel PF, Bürki K, et al. The transcription factor early growth response 1 (Egr-1) ad-

40. Wen AY, Sakamoto KM, Miller LS. The role of the transcription factor CREB in immune function. J Immunol 2010;185:6413–9.

41. Yu F, Sharma S, Jankovic D, Gurram RK, Su P, Hu G, et al. The transcription factor Bhlhe40 is a switch of inflammatory versus antiinflammatory Th1 cell fate determination. J Exp Med 2018;215:1813–21.

42. Huynh JP, Lin CC, Kimmey JM, Jarjour NN, Schwarzkopf EA, Bradstreet TR, et al. Bhlhe40 is an essential repressor of IL-10 during Mycobacterium tuberculosis infection. J Exp Med 2018;215:1823–38.

43. Zhao M, Xiaofei L, Gang C, Wei L, Jing X, Gang H, et al. DEC1 binding to the proximal promoter of CYP3A4 ascribes to the downregulation of CYP3A4 expression by IL-6 in primary human hepatocytes. Biochem Pharmacol 2012;84:701–11.

44. Meyer SN, Scuoppo C, Vlasevska S, Bal E, Holmes AB, Holloman M, et al. Unique and shared epigenetic programs of the CREB-BP and EP300 acetyltransferases in germinal center B cells reveal targetable dependencies in lymphoma. Immunity 2019;51:535–47.

45. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 2015;518:337–43.

46. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580–5.

47. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014;159:1665–80.

48. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, et al. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biol 2018;19:151.

49. Emond MJ, Louie T, Emerson J, Zhao W, Mathias RA, Knowles MR, et al. Exome sequencing of extreme phenotypes identifies DCTN4 as a modifier of chronic Pseudomonas aeruginosa infection in cystic fibrosis [letter]. Nat Genet 2012;44:886–9.

vances differentiation of pre-B and immature B cells. J Exp Med 1998;188:2215–24.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Systemic Sclerosis Dermal Fibroblasts Induce Cutaneous Fibrosis Through Lysyl Oxidase–like 4: New Evidence From Three-Dimensional Skin-like Tissues

Mengqi Huang,[1] [iD] Guoshuai Cai,[2] Lauren M. Baugh,[3] Zhiyi Liu,[4] Avi Smith,[5] Matthew Watson,[3] Dillon Popovich,[6] Tianyue Zhang,[5] Lukasz S. Stawski,[7] Maria Trojanowska,[7] Irene Georgakoudi,[3] Lauren D. Black III,[8] Patricia A. Pioli,[6] Michael L. Whitfield,[6] and Jonathan Garlick[5]

Objective. Systemic sclerosis (SSc) is a clinically heterogeneous disease characterized by increased collagen accumulation and skin stiffness. Our previous work has demonstrated that transforming growth factor β (TGFβ) induces extracellular matrix (ECM) modifications through lysyl oxidase–like 4 (LOXL-4), a collagen crosslinking enzyme, in bioengineered human skin equivalents (HSEs) and self-assembled stromal tissues (SAS). We undertook this study to investigate cutaneous fibrosis and the role of LOXL-4 in SSc pathogenesis using HSEs and SAS.

Methods. SSc-derived dermal fibroblasts (SScDFs; n = 8) and normal dermal fibroblasts (NDFs; n = 6) were incorporated into HSEs and SAS. These 3-dimensional skin-like microenvironments were used to study the effects of dysregulated LOXL-4 on ECM remodeling, fibroblast activation, and response to TGFβ stimulation.

Results. SScDF-containing SAS showed increased stromal thickness, collagen deposition, and interleukin-6 secretion compared to NDF-containing SAS ($P < 0.05$). In HSE, SScDFs altered collagen as seen by a more mature and aligned fibrillar structure ($P < 0.05$). With SScDFs, enhanced stromal rigidity with increased collagen crosslinking ($P < 0.05$), up-regulation of *LOXL4* expression ($P < 0.01$), and innate immune signaling genes were observed in both tissue models. Conversely, knockdown of *LOXL4* suppressed rigidity, contraction, and α-smooth muscle actin expression in SScDFs in HSE, and TGFβ-induced ECM aggregation and collagen crosslinking in SAS.

Conclusion. A limitation to the development of effective therapeutics in SSc is the lack of in vitro human model systems that replicate human skin. Our findings demonstrate that SAS and HSE can serve as complementary in vitro skin-like models for investigation of the mechanisms and mediators that drive fibrosis in SSc and implicate a pivotal role for LOXL-4 in SSc pathogenesis.

INTRODUCTION

Skin fibrosis is a common feature of systemic sclerosis (SSc). Pathologic hallmarks of dermal fibrosis include an increased proportion of α-smooth muscle actin (α-SMA)–expressing fibroblasts and excess production and altered remodeling of collagenous extracellular matrix (ECM) (1).

This leads to abnormal tissue contraction and stiffness, with decreased mobility and other complications (2). In SSc, this becomes even more critical as the severity of skin involvement, which inversely correlates with patient survival (3,4), is a key indicator of disease progression and patient prognosis. Unfortunately, no effective therapy has been developed to treat fibrotic progression in skin.

Current methods of studying SSc include animal models of fibrosis and 2-dimensional (2-D) monolayer culture systems that use SSc patient–derived fibroblasts. However, both of these biological systems have drawbacks. Although animal models have been widely used to study fibrosis, mouse models of SSc may not fully recapitulate the range of clinical features associated with disease (5). Five commonly used models, including tight skin 1 (TSK1/+), TSK2/+, sclerodermatous graft-versus-host disease (GVHD), and bleomycin-induced fibrosis models, were analyzed in a comparative genomic study (6). Findings demonstrated that sclerodermatous GVHD (7), bleomycin-induced fibrosis (6), and TSK2/+ (8) represented different subsets of SSc patients, and surprisingly, TSK1/+ did not resemble SSc on the molecular level. Alternatively, fibroblasts derived from SSc biopsy samples might be more representative of SSc; however, traditional 2-D monolayer cultures have been shown to poorly reproduce gene expression changes in skin and do not recreate in vivo skin phenotypes when grown on plastic in vitro (9,10).

In order to fill in the gaps between these existing models and human skin, a critical goal is to create a biological system that can model the complex intercellular interactions and tissue microenvironment present in SSc patients. We have previously shown that 2 well-characterized skin fabrication systems, self-assembled stromal tissues (SAS) and human skin equivalents (HSEs), hold promise for in vitro studies of SSc that cannot be performed in either 2-D culture or by using existing animal models (11). In the present study, SAS were used to analyze endogenous ECM produced by dermal fibroblasts that were grown from the skin of SSc patients (SSc-derived dermal fibroblasts [SScDFs]). Additionally, we used HSEs that are morphologically similar to full-thickness human skin in order to study SScDFs in a realistic 3-D tissue microenvironment in which epithelial–dermal cross-talk is indispensable for recreating cell phenotypes observed in vivo. Together, SAS and HSE serve as complementary skin-like tissue models which may provide valuable insights when investigating SSc pathogenesis by studying mediators and mechanisms that drive fibrotic phenotypes.

Type I collagen plays a major role in providing mechanical stability to tissues and structures and is known to be altered in the connective tissue of patients with SSc (12,13). The lysyl oxidase (LOX) family, made up of the LOX and LOX-like (LOXL) 1–4 enzymes, is responsible for initiating covalent crosslinking in collagen fibrils by deamination on lysine residues (14). This crosslinking reaction makes the collagen more resistant to degradation and also provides additional mechanical strength to the ECM (15,16). LOXs have been evaluated in the skin of SSc patients using immunohistochemical staining and, while LOXs were increased in interstitial fibroblasts compared to normal skin, LOX expression was not augmented in SSc patients with skin atrophy (17,18). Researchers have demonstrated that serum LOX levels are elevated in SSc patients and are directly correlated with the Modified Rodnan Skin Thickness Score (MRSS) (3), suggesting that LOX levels may be a potential biomarker of fibrosis in SSc (19,20), although expression

of LOXL enzymes (LOXL-1–4) was not evaluated in those studies. Additionally, it has previously been shown that LOXL-4–mediated crosslinking may signal through the transforming growth factor β (TGFβ) pathway to induce fibrosis by modifying ECM organization and tissue stiffness (11,21). Based on these findings, we conducted the present study in order to investigate the potential role of LOXL-4 in the pathogenesis of SSc, using 3-D human SSc tissue–based models of skin fibrosis.

MATERIALS AND METHODS

Human dermal fibroblasts. Using a protocol approved by the Tufts Medical Center Institutional Review Board (IRB), deidentified skin biopsy specimens (4 mm in diameter, 3 mm in depth) were obtained from the forearm/bicep of scleroderma patients and sex-matched healthy controls. Dermal fibroblasts were isolated as previously described (22). Briefly, skin biopsy samples were incubated overnight at 4ºC in Dispase (Invitrogen) to remove epidermis from skin. The connective tissues were minced and incubated in a mixture of collagenase (Invitrogen), hyaluronidase (Sigma-Aldrich), and Dulbecco's modified Eagle's medium/F-12 (Invitrogen) for 1 hour at 37ºC with constant stirring. Cells and tissues were collected by centrifugation and cultured in 10% fetal bovine serum (FBS) medium (HyClone).

Human peripheral blood plasma. Each SSc patient and a sex-matched control donor provided 20 ml of blood after approval by the Dartmouth-Hitchcock Medical Center IRB. Peripheral blood was transferred to a tube with Ficoll and centrifuged at 1,600 revolutions per minute for 30 minutes. Plasma sited in the top layer was collected from the buffy coat layer beneath.

Three-dimensional tissue constructs. For HSEs, SScDFs or normal dermal fibroblasts (NDFs) were mixed with bovine CTX-I (Organogenesis) to a final concentration of 3×10^5 fibroblasts/ml and cultured with 10% FBS. After the collagen matrix was remodeled by the embedded fibroblasts for 1 week, 5×10^5 human neonatal keratinocytes were added onto the collagen. Next, HSEs were submerged in 0.3% FBS for 5 days and lifted into a liquid–air interface for the subsequent 8 days to be fed with 2% FBS and treated with 8 μ*M* β-aminopropionitrile (β-APN). For SAS, SScDFs or NDFs were seeded into Transwell inserts (Millipore) at 1.6×10^5 fibroblasts/insert and cultured with 5% FBS and 10 μg/ml ascorbic acid (Sigma-Aldrich) for 5 weeks. HSEs and SAS were constructed using NDFs and SScDFs at passage 4 or 5. For studies using short hairpin RNA (shRNA) lentivirus transduction experiments, fibroblasts were used at passage 7 or 8.

Atomic force microscopy (AFM). Fresh 3-D tissues were mechanically characterized by performing microindentation with a 5-μm borosilicate spherical probe (Novascan) at a nominal spring constant of 0.06 N/m. Calibration of the microscopy

system (3100 Veeco) was performed according to the instructions of the manufacturer (Diagnostic Instruments) before each indentation measurement. Force-contact profiles were acquired at an indentation velocity of 20 μm/second separated by 5 μm in a 16 × 16 μm sample grid to map an area of 80 × 80 μm. Elastic moduli at each point at the grid were calculated by fitting force-contact data using the Hertz model (23).

Multiphoton microscopy. HSE was maintained in a humidified chamber to prevent dehydration during imaging. Images (512 × 512 pixels; 600.0 × 600.0 μm) were acquired using a 3-μm Z-step (70 optical sections total) with a Leica TCS SP2 confocal microscope. The second harmonic generation (SHG) images (excitation 800 nm, emission 400 ± 10 nm) and two-photon excited fluorescence (TPEF) images (excitation 755 nm, emission 525 ± 25 nm) were processed in MatLab to obtain quantitative metrics of collagen microstructure (24). Next, a weighted summation algorithm was used to determine the fiber 3-D orientations, based on representative collagen spatial organization. The cumulative TPEF signal over the collagen-containing regions was computed in order to determine the collagen crosslinking, as conducted in our previous studies (11,25).

Quantification of hydroxyproline. The total amount of de novo collagen in SAS was determined using a hydroxyproline assay kit (Cell Biolabs) as previously described (11). Briefly, SAS were homogenized in phosphate buffered saline by sonication on ice. Samples were hydrolyzed using 12N hydrochloric acid at 120°C for 3 hours and then incubated with chloramine T mixture for 30 minutes at room temperature followed by the 4-dimethylaminobenzaldehyde mixture for 90 minutes at 60°C. The absorbance was read at 540–560 nm, and hydroxyproline content was normalized to the DNA amount measured with a Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific).

Collagen crosslinking. To assess the collagen crosslinking level in SAS, a human CTX-I kit was used according to the instructions of the manufacturer (MyBioSource). Briefly, 50 μl aliquot of homogenized SAS sample that was used for hydroxyproline assay was incubated in microtiter plate precoated with horseradish peroxidase–conjugated CTX-I antibody at 37°C for 1 hour. The enzyme–substrate reaction was terminated by adding 50 μl of sulfuric acid solution, and the color change was measured spectrophotometrically at a wavelength of 450 nm.

Short hairpin RNA lentivirus–mediated knockdown. Briefly, pLKO.1 plasmid encoding shRNA (Sigma-Aldrich) against human *LOXL4* or *SMAD3* gene sequences (see Supplementary Table 1 on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.41163/abstract) was transfected with lentiviral packaging plasmids, pMD2.G and psPAX2 (both from Addgene), into HEK 293T cells using FuGENE6 (Promega)

to generate lentiviral particles. The shRNA nontarget vector was used as negative control. For transduction, human fibroblasts were seeded at 1 × 10^5 fibroblasts/well in the 6-well plate for 24 hours and incubated with 2 ml lentiviral supernatant, supplemented with 1 ng/ml Polybrene (Sigma-Aldrich) for 7 hours. Subsequently, complete medium was added, followed by 2 μg/ml puromycin (Sigma-Aldrich) for 2 weeks.

DNA microarray analysis. RNA was isolated using an RNeasy Fibrous Tissue Mini Kit (Qiagen) followed by amplification and labeling using a Low Input Quick Amp Labeling Kit (Agilent). Signal density was log$_2$ lowess normalized and filtered for probes with an intensity of ≥2-fold over the local background in Cy3 or Cy5 channels. Probes with >20% missing data were excluded.

Differential expression analysis was performed using R, version 3.4.4. False discovery rate–adjusted P values less than 0.05 were considered significant. The bias between 2 individual patient sets was adjusted by ComBat (26), and expression profiles of differentially expressed genes were hierarchically clustered and presented using a heatmap. Gene set enrichment analysis (27) was applied in order to identify the enrichment of hallmark gene sets in the Molecular Signatures Database (28).

Statistical analysis. Values are reported as the mean ± SD, with a sample size (n ≥ 4) specified for each experiment. Experimental data were analyzed by Student's unpaired t-test for differences between each of the groups, with two-way analysis of variance for overall condition effects using GraphPad Prism 5.0 software. P values less than 0.05 were considered significant for differences between groups. More details can be found in Supplementary Methods (http://onlinelibrary.wiley.com/doi/10.1002/art.41163/abstract).

RESULTS

Cutaneous fibrosis–associated features retained in 3-D in vitro tissues populated with SScDFs. The 2 complementary skin-like 3-D tissue platforms containing HDFs accurately recapitulated key features of the in vivo SSc skin environment, as indicated by endogenous stroma deposition (SAS) and interactions between keratinocytes and fibroblasts embedded in an ECM (HSEs) (Figure 1A). To assess whether SScDFs exhibit excessive ECM production compared to NDFs, both were stimulated with ascorbic acid for 5 weeks to form SAS. As demonstrated in Figures 1B–D, SScDF-constructed SAS deposited a cellular, stromal tissue that demonstrated notable increases in both tissue thickness and rigidity compared to NDF-constructed SAS. This tissue phenotype aligns well with the clinical diagnosis of skin pathogenesis in SSc patients.

Since collagen makes up ~75–80% of human skin and forms fibrils that provide the structural and mechanical properties in fibrous tissue (29), we next investigated collagen content and crosslinking in these skin-like tissue models harboring SScDFs.
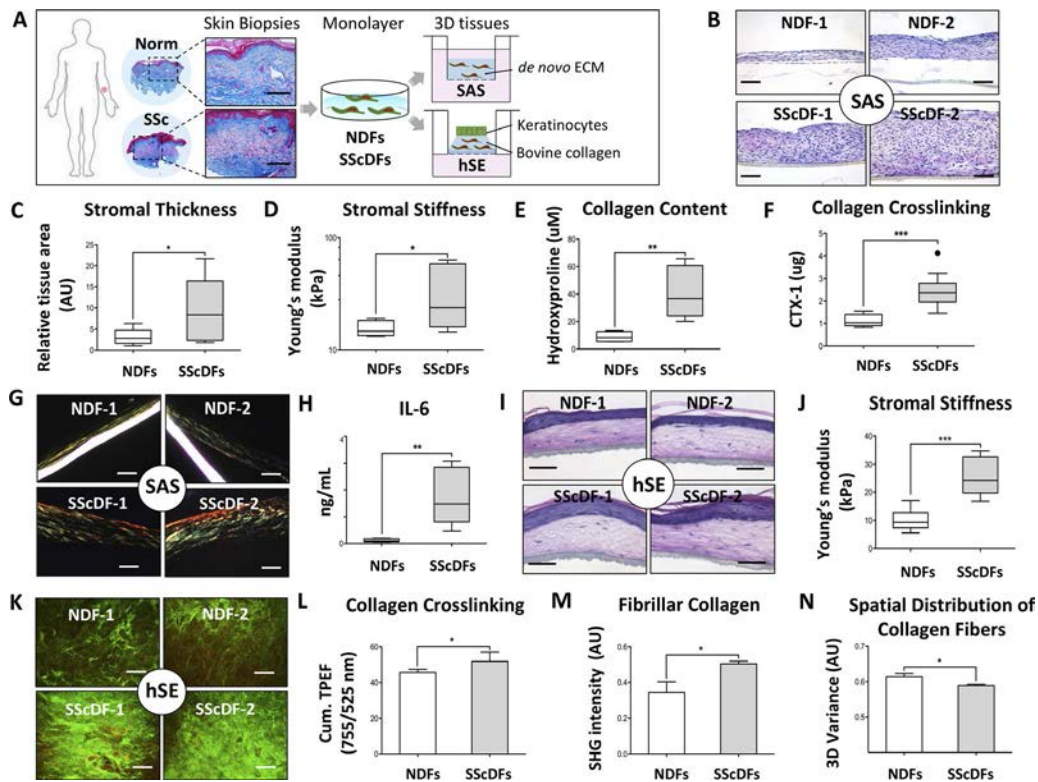
Figure 1. Altered extracellular matrix (ECM) properties of systemic sclerosis (SSc) recapitulated in models using 2 complementary skin-like tissues: self-assembled stromal (SAS) and human skin equivalent (HSE). **A**, Schematic representation of the incorporation of dermal fibroblasts derived from normal (Norm) skin biopsy samples (NDFs; n = 6) or SSc skin biopsy samples (SScDFs; n = 8) into SAS and HSE. Bars = 200 μm. **B**, Representative hematoxylin and eosin (H&E) staining of SAS. Bars = 60 μm. **C**, Stromal tissue thickness of SAS assessed by quantifying the tissue surface area in **B**. **D**, Stromal stiffness of SAS measured by atomic force microscopy (AFM). **E**, Total collagen content of SAS lysate analyzed by hydroxyproline assay. **F**, Collagen crosslinking level of SAS measured with a CTX-I kit. **G**, Representative images with picrosirius red staining. Yellow/orange birefringence represents type I collagen; white represents polycarbonate membrane. Bars = 30 μm. **H**, Interleukin-6 (IL-6) measured from SAS supernatant. **I**, Representative H&E staining of HSE. Bars = 100 μm. **J**, Stromal stiffness of HSE measured by AFM. **K**, Representative images of second harmonic generation (SHG) microscopy (green) and two-photon excited fluorescence (TPEF) microscopy (red). Bars = 100 μm. In **C**–**F**, **H**, and **J**, data are shown as box plots, where each box represents the interquartile range (IQR). Lines inside the boxes represent the median, and lines outside the boxes represent 1.5 times the upper and lower IQRs. **L**–**N**, Collagen microstructures quantified by multiphoton (SHG/TPEF) microscopy. Levels of collagen crosslinking (cumulated TPEF) (**L**), total collagen (SHG intensity) (**M**), and spatial distribution (3-dimensional [3-D] variance) (**N**) are shown. Bars show the mean ± SD. * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$.

Our results showed that a more collagenous matrix and higher crosslinking levels were found in SScDF-constructed SAS compared to NDF-constructed SAS (Figures 1E and F). Additionally, very few well-developed fibril bundles were detected in NDF-constructed SAS (Figure 1G), suggesting that the excessive collagen deposition in fibrotic skin may result from collagen fibers with a greater degree of crosslinking, which renders them more resistant to degradation.

Because interleukin-6 (IL-6) is secreted by dermal fibroblasts and has been shown to mediate skin thickness in SSc patients (30), soluble IL-6 levels were measured in SAS supernatant. As demonstrated in Figure 1H, secreted IL-6 expression was significantly higher in SScDFs compared to NDFs. To investigate fibroblasts in a relevant microenvironment when grown in the presence of a fully differentiated epithelium, NDFs and SScDFs were incorporated into HSE, and their ability to modify the surrounding bovine

collagen that formed the dermis of HSE was evaluated (Figure 1I). Consistent with findings in SAS, stromal stiffness measured in the dermal surface of HSE with SScDFs was 2.5-fold higher than in tissue with NDFs (Figure 1J). Noninvasive TPEF and SHG microscopy of collagenous ECM (31,32) was assessed to confirm this alteration of the mechanical properties of skin. Elevated crosslinking, with higher amounts of collagen fibers in the ECM embedded with SScDFs compared to NDFs, was observed (Figures 1K–M). A recently developed quantitative analysis of 3-D directional variance (33) was used to measure the spatial organization of fibrillar collagen in HSE, and SScDF-modified collagen fibers were more aligned compared to those modified by NDFs (Figure 1N). Of note, a similar microstructure alteration was previously observed in skin biopsy samples from both bleomycin-treated mice and SSc patients (34,35). Taken together, these findings demonstrate that both SAS and HSE provide rigorous skin-like tissue platforms to

address the critical role of fibroblasts during the development of fibrosis in the connective tissue of SSc patients.

Increased expression of *LOXL4* in 3-D skin-like tissue harboring SScDFs. To identify potentially dysregulated genes associated with the fibrotic phenotypes observed in 3-D skin-like tissues, we performed genome-wide analyses of gene expression in SAS that were constructed using SScDFs from 3 patients with limited SSc and 5 patients with diffuse SSc and NDFs from 6 sex-matched controls (Table 1 and Supplementary Table 2, http://onlinelibrary.wiley.com/doi/10.1002/art.41163/abstract). A total of 895 genes were found to be differentially expressed between SScDFs and NDFs in SAS. Signaling pathway analysis of 463 up-regulated genes in SScDFs demonstrated an increase in genes that are hallmarks of SSc and are functionally enriched for TGFβ signaling, as well as increases in interferon and inflammatory responses (Figure 2A and Supplementary Table 2, http://onlinelibrary.wiley.com/doi/10.1002/art.41163/abstract).

Consistent with our previously published findings (11), *LOXL4* showed increased expression in tissues containing SScDFs compared to those with NDFs and confirmed at the protein level (Figures 2A and B). Other genes with increased expression in SScDFs included *TGFBR1*, *THBS2*, *MMP2*, and *PLOD* (Figure 2A), which are well recognized for their concerted role in the regulation of ECM homeostasis. Moreover, up-regulation of *LOXL4* in the skin biopsy samples from SSc patients was confirmed using a previously published clinical data set (Figure 2C). We performed an analysis by intrinsic molecular subset using the originally published sample assignments (36). Increased expression of *LOXL4* was identified in proliferative and inflammatory SSc subsets but not in limited or normal-like SSc subsets (Figure 2D), which further elucidates the heterogeneity of SSc at the molecular level.

To provide soluble mediators present in the local in vivo microenvironment, SAS containing NDFs or SScDFs were treated with plasma collected from either a healthy donor or an SSc patient during the early stage of tissue development, when the ECM is deposited and organized. As presented in Figures 2E and F, the difference in *LOXL4* expression between SScDFs and NDFs increased from ~2.5-fold in HSE without plasma to more than 8-fold in SAS with plasma. Additionally, the difference in stromal stiffness between SSc plasma–incubated SAS containing SScDFs and control plasma–incubated SAS containing NDFs increased from 2-fold to 3-fold compared to that of SScDFs and NDFs without plasma in the HSE (Figures 1J and 2G). These data indicate that 3-D tissue models are useful for monitoring the effects of altered expression of *LOXL4* on fibrosis in SScDFs compared to control NDFs.

LOXL-4–mediated collagen crosslinking maintains fibrogenic phenotypes in HSE embedded with SScDFs. To elucidate the role of LOXL-4 in modulating the ECM of skin-like tissues, SScDFs and NDFs were transduced with a lentivirus carrying shRNA sequences that impair *LOXL4* activity (Figure 3A). Following verification of knockdown of *LOXL4* expression or controls, SScDFs and NDFs were incorporated into HSE (Figure 3B). The level of contraction, mediated by the fibroblasts embedded within HSEs, was characterized by measuring the size of the plateau area in the center of HSEs. We found that the contraction was higher in SScDFs compared to NDFs expressing vector control (Figures 3B and C). In contrast, contraction was inhibited by suppression of *LOXL4* in both SScDFs and NDFs (Figures 3B and C).

To determine if LOXL-4–mediated collagen crosslinking could also affect the stromal stiffness of HSE, AFM was performed. As shown in Figure 3D, *LOXL4* knockdown resulted in decreased rigidity in both SScDFs and NDFs. Interestingly, levels of messenger RNA (mRNA) for *COL1A2* were not affected by the expression

Table 1. Demographic and clinical characteristics of donors of skin-derived NDFs (n = 6) and SScDFs (n = 8)*

	Ethnicity	Age/sex	Location	Type of SSc
NDF donor				
1	White	53/F	Forearm	–
2	White	62/F	Forearm	–
3	White	28/F	Forearm	–
4	White	60/M	Forearm	–
5	White	23/F	Forearm	–
6	White	34/M	Forearm	–
SScDF donor				
1	Hispanic	53/F	Forearm, uninvolved skin	Limited
2	White	72/F	Bicep, uninvolved skin	Diffuse
3	Hispanic	53/F	Forearm, uninvolved skin	Limited
4	White	62/F	Forearm, uninvolved skin	Limited
5	White	64/M	Forearm, involved skin	Diffuse
6	White	45/F	Forearm, involved skin	Diffuse
7	White	50/M	Forearm, involved skin	Diffuse
8	White	63/F	Forearm, involved skin	Diffuse
CP	White	46/F	Peripheral blood	–
SP	White	34/F	Peripheral blood	Diffuse

* Six normal dermal fibroblast (NDF) strains and 8 systemic sclerosis dermal fibroblast (SScDF) strains were established using skin biopsy samples. One control plasma (CP) and 1 SSc plasma (SP) were collected from peripheral blood.
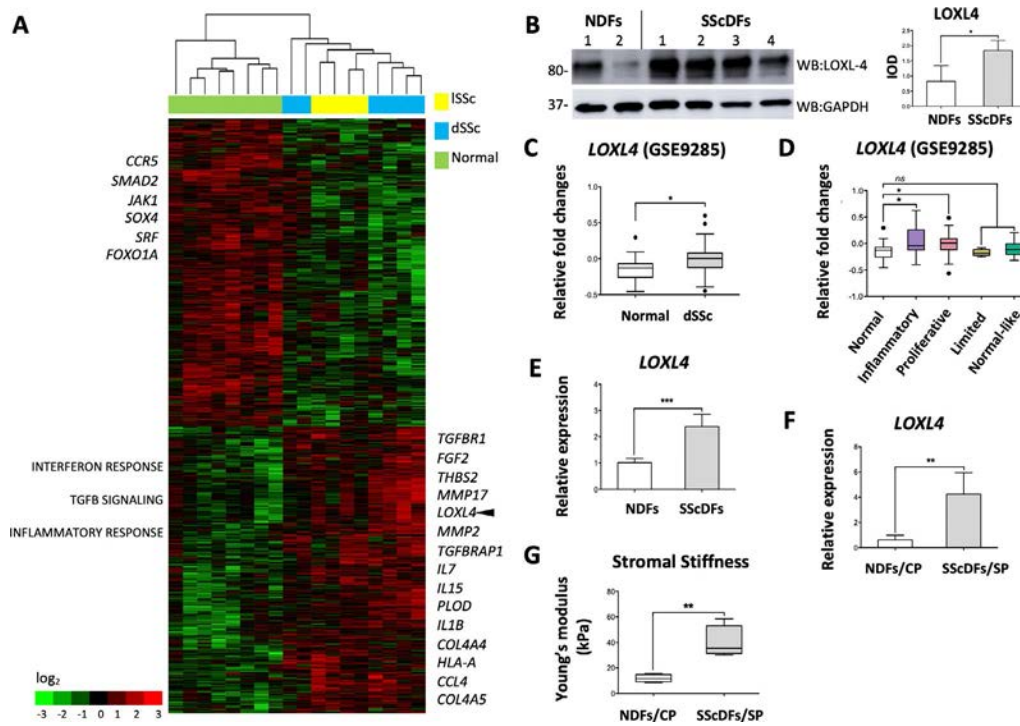
Figure 2. Expression of *LOXL4* is elevated in both SScDFs and SSc skin biopsy samples. **A**, Hierarchical clustering of 895 genes that were significantly dysregulated (adjusted $P < 0.05$) in SScDFs (n = 8) compared to NDFs (n = 6) in SAS. Microarray gene expression signal was $\log_2$ transformed and quantile normalized. Batch effects were corrected by ComBat analysis. Selected differentially expressed genes are displayed on both sides. Selected signaling pathways displayed on the left side were determined by gene set enrichment analysis on the MSigDB hallmark gene set (false discovery rate < 0.02) of 432 up-regulated genes in SScDFs. **B**, Western blot (WB) of lysyl oxidase–like 4 (LOXL-4) (80 kd) and GAPDH (37 kd) using SAS lysates of NDFs (n = 2) and SScDFs (n = 4), and quantified integrated optical density (IOD) results. **C** and **D**, Relative fold changes of *LOXL4* expression in GSE9285 by DNA microarray analysis of skin biopsy specimens. **E**, Relative expression of *LOXL4* mRNA in NDFs (n = 6) and SScDFs (n = 8) isolated from HSE, measured by quantitative polymerase chain reaction (qPCR) and normalized to GAPDH expression. **F** and **G**, Expression of *LOXL4* (**F**) and stromal stiffness (**G**) in SAS containing NDFs (n = 4) and SScDFs (n = 4) treated with control plasma (CP) or SSc plasma (SP), respectively, for 1 week. Relative expression of *LOXL4* mRNA was measured by qPCR and normalized to GAPDH. Stromal stiffness of SAS was measured by AFM. In **B**, **E**, and **F**, bars show the mean ± SD. In **C**, **D**, and **G**, data are shown as box plots, where each box represents the interquartile range (IQR). Lines inside the boxes represent the median, and lines outside the boxes represent 1.5 times the upper and lower IQRs. * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$. lSSc = limited SSc; dSSc = diffuse SSc (see Figure 1 for other definitions). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41163/abstract.

of *LOXL4*, regardless of the fibroblasts' origin (Figure 3E). Fibroblast activation indicated by expression level of *ACTA2* with α-SMA staining was significantly higher in SScDFs compared to NDFs, in which both were inhibited by *LOXL4* knockdown (Figures 3F–H). Additionally, there was a positive correlation between decreasing levels of *ACTA2* with α-SMA staining and decreased stromal stiffness by *LOXL4* knockdown in both NDFs and SScDFs (Figure 3D). These results suggest that inhibition of *LOXL4* in SScDFs was able to attenuate fibrotic phenotypes, including increased contractility and dermal stiffness via suppression of myofibroblast activation.

Next, we examined the effects of β-APN, a small molecular inhibitor of the LOX family, on HSE containing either SScDFs or NDFs. Consistent with previous findings using *LOXL4* knockdown, β-APN significantly decreased the stromal stiffness in HSE with SScDFs. However, we failed to observe a difference in NDFs (Figure 3I). Given the roles of *COL1A2* and *ACTA2* in fibrosis, we also tested the effect of β-APN treatment on expression

of these genes. While no significant differences in *COL1A2* were observed between NDFs and SScDFs with or without treatment with β-APN (Figure 3J), *ACTA2* was significantly down-regulated by β-APN in both NDFs and SScDFs (Figure 3K). This may represent part of a positive feedback loop via sensing of mechanical force that is related to altered stiffness. Using 3-D variance, we observed decreased alignment of collagen fibers in β-APN–treated HSE containing SScDFs (Figure 3L). Collectively, these results suggest that suppression of the collagen-modifying enzyme, LOXL-4, which is up-regulated in SScDFs, may attenuate excessive collagen production and altered ECM organization and structures characteristic of the fibrotic dermis.

Blocking of TGFβ-enhanced collagen synthesis by targeting LOXL-4 in SScDF-harboring SAS. To understand the role of LOXL-4 during the TGFβ-induced biosynthesis of collagen, shRNA against *SMAD3* were used to block
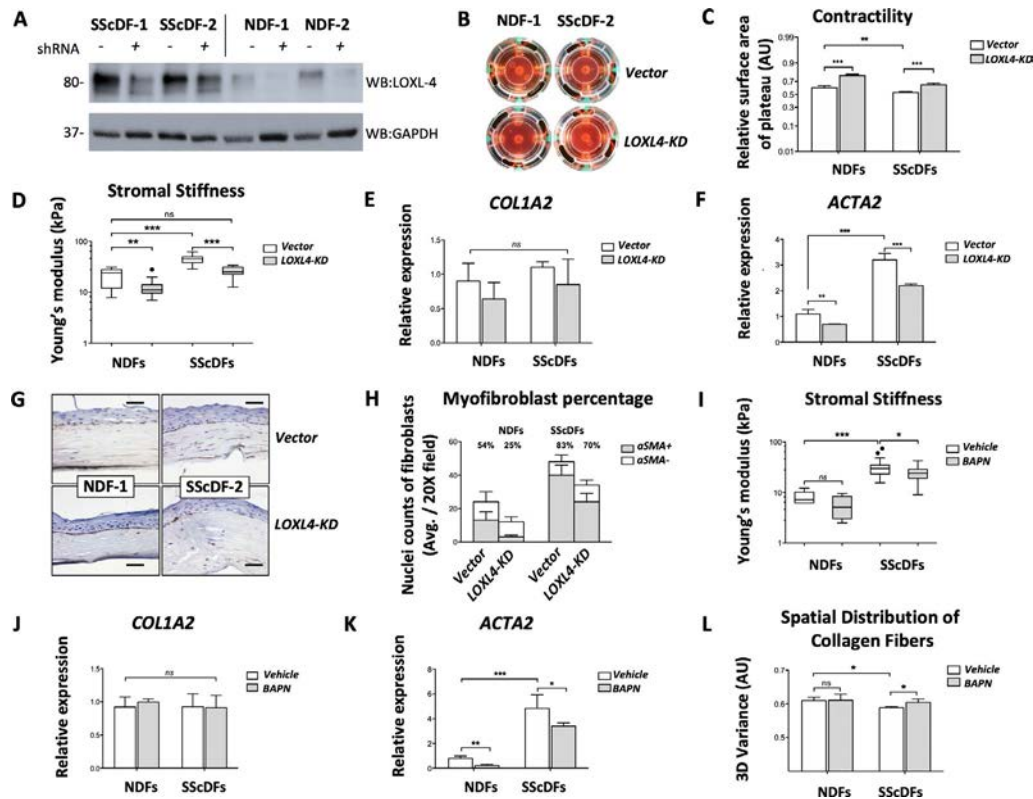
Figure 3. Suppression of lysyl oxidase–like 4 (LOXL-4) in SScDFs alleviates profibrogenic phenotypes in HSE. *LOXL*4 expression was knocked down by transduction of lentivirus carrying short hairpin RNA (shRNA) targeting *LOXL*4 or control sequence (vector) into NDF-1, NDF-2, SScDF-1, and SScDF-2 (n = 3 biologic replicates for each). HSE was incorporated with NDF-1, NDF-2, SScDF-1, and SScDF-2 and treated with 8 μg β-aminopropionitrile (β-APN) for 8 days. **A**, Western blot (WB) bands showing the protein levels of LOXL-4 (80 kd) and GAPDH (37 kd) in NDFs and SScDFs. **B**, Representative overhead images of HSE, with tissue plateau area circled by dashed white line. **C**, Contractile levels induced by fibroblasts as indicated by relative surface area, calculated by comparing surface of tissue plateau in **B** to Transwell membrane surface. **D** and **I**, Stromal stiffness of HSE measured by AFM, with either *LOXL*4 knockdown (KD) (**D**) or β-APN (**I**). Data are shown as box plots, where each box represents the interquartile range (IQR). Lines inside the boxes represent the median, and lines outside the boxes represent 1.5 times the upper and lower IQRs. **E**, **F**, **J**, and **K**, Relative expression of mRNA for *COL1A2* (**E** and **J**) and *ACTA2* (**F** and **K**) in fibroblasts dissociated from HSE, measured by quantitative polymerase chain reaction and normalized to GAPDH expression. **G**, Immunohistochemistry staining of α-smooth muscle actin (α-SMA). Bars = 100 μm. **H**, Quantification of myofibroblast percentage in **G** by identifying the number of α-SMA–positive fibroblasts among the total number of fibroblasts. **L**, Spatial distribution (3-dimensional [3-D] variance) of fibrillar collagen in HSE quantified by multiphoton (SHG/TPEF) microscopy. Bars show the mean ± SD. * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$. NS = not significant (see Figure 1 for other definitions). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41163/abstract.

canonical TGFβ-mediated responses in SScDFs and NDFs (Figures 4A and B). Indeed, suppression of *SMAD3* resulted in reduced production of type I collagen in SAS (Figures 4C and D), indicating that autocrine TGFβ signaling promotes ECM aggregation. In parallel, shRNA against *LOXL4* in SScDFs and NDFs decreased levels of type I collagen deposition in SAS, similar to the inhibitory effects of blocking canonical TGFβ via *SMAD3* knockdown (Figures 4C and D).

Next, exogenous TGFβ1 was added to SAS to evaluate paracrine TGFβ-mediated responses in SScDFs and NDFs. SAS treated with TGFβ1 formed nodules in both SScDFs and NDFs (Figure 4E), suggesting an increased accumulation of ECM in response to TGFβ. The differences between SScDFs and NDFs were not altered by TGFβ1 induction, as demonstrated by the increased mass of tissue cross-sections in SScDF-constructed

SAS (Figure 4E). One possible explanation for this is that collagen deposition was increased in SScDF-constructed SAS compared to NDF-constructed SAS in either the presence or absence of TGFβ1 (Figure 4F). Suppression of *LOXL4* significantly inhibited the nodular-like structures induced by TGFβ1, to an extent similar to that seen with *SMAD3*, in both SScDFs and NDFs (Figure 4E). TGFβ1-induced collagen accumulation was successfully blocked by *LOXL4* knockdown to a level comparable to that observed by inhibiting the canonical TGFβ pathway via *SMAD3* knockdown (Figures 4G and H). As expected, the collagen crosslinking level of SAS was found to be higher in SScDFs than NDFs and was significantly enhanced by TGFβ1 in both SScDFs and NDFs (Figure 4I). While suppression of *LOXL4* impaired TGFβ1-induced crosslinking of collagen, the initial differences between SScDFs and NDFs remained (Figure 4J). In contrast, inhibition of *SMAD3*
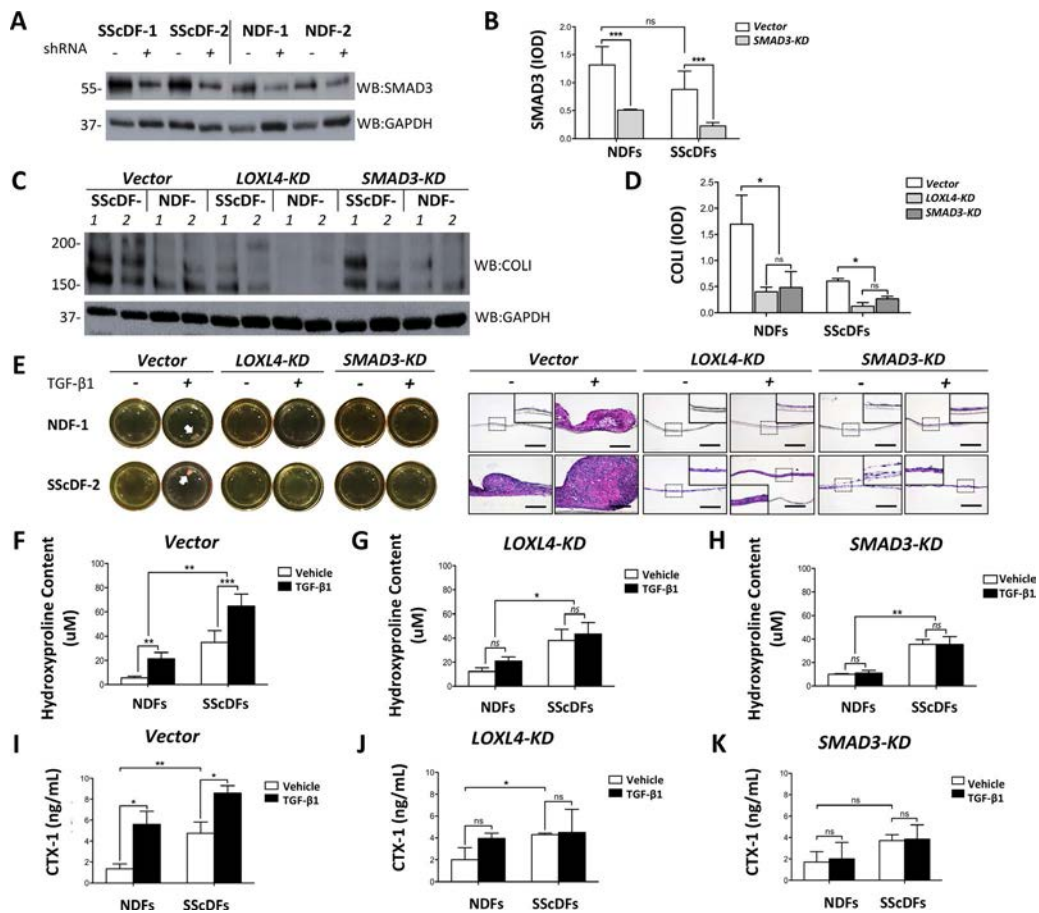
Figure 4. Blocking *LOXL4* in SScDFs tempers transforming growth factor β1 (TGFβ1)–induced collagen accumulation in SAS. Lentiviruses carrying short hairpin RNA (shRNA) targeting *LOXL4* knockdown (KD), *SMAD3* knockdown, or the control sequence (vector) were transduced into NDF-1, NDF-2, SScDF-1, and SScDF-2. **A**, Western blot (WB) bands showing protein levels of SMAD3 (55 kd) and GAPDH (37 kd) from the cell lysates of NDF-1, NDF-2, SScDF-1, and SScDF-2. **B**, Quantified integrated optical density (IOD) results from the SMAD3 Western blot, normalized to GAPDH. **C**, Western blot bands showing protein levels of Col1 (200–150 kd) and GAPDH (37 kd) from the tissue lysates of SAS. **D**, Quantified integrated optical density (IOD) results from the Col1 Western blot, normalized to GAPDH. **E**, Representative images of SAS treated with 2 ng/ml TGFβ1 for 2 weeks and corresponding H&E staining. **Arrows** show SAS nodule. Bars = 60 μm. **Insets** are enlarged images of the boxed areas showing thin tissue in SAS. **F**–**H**, Total collagen content of SAS lysate after treatment with vector (**F**), *LOXL4* knockdown lentivirus (**G**), or *SMAD3* knockdown lentivirus (**H**), analyzed by hydroxyproline assay and normalized to total cell number. **I**–**K**, Collagen crosslinking level of SAS after treatment with vector (**I**), *LOXL4* knockdown lentivirus (**J**), or *SMAD3* knockdown lentivirus (**K**), measured by CTX-I analysis and normalized to total collagen detected in **F**, **G**, and **H**. Bars show the mean ± SD (n = 6). * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$. NS = not significant (see Figure 1 for other definitions). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41163/abstract.

completely abrogated not only the TGFβ1 effects but also the differences in basal collagen crosslinking between SScDFs and NDFs (Figure 4K). Collectively, these results suggest that TGFβ-induced collagen aggregation is at least partially mediated by enhanced crosslinking in SScDFs in a *LOXL4*-dependent manner.

DISCUSSION

In recent years, significant progress has been made in understanding of site-specific, fibrogenic characteristics of fibroblasts from a variety of origins under pathologic conditions (37–40). The plasticity of fibroblasts may contribute to the reservoir of disease-causing myofibroblasts in SSc, which may in

part explain the functional cellular heterogeneity of this disease. The ability of SSc fibroblasts to promote autocrine recruitment and differentiation in SSc is actively being investigated. Here, we used 2 complementary skin-like tissue models to provide new insights into these mechanisms by studying mediators and regulators that drive the pathogenesis of SSc.

We previously found that LOXL-4 was elevated in SScDFs and may be partly responsible for mediating TGFβ-induced fibrosis (11). Increasing evidence has demonstrated that both canonical and noncanonical (JNK/activator protein 1 [AP-1]) pathways are involved in TGFβ-mediated *LOXL4* induction (41,42), contributing to ECM remodeling and fibrosis. Regulatory regions of *LOXL4* contain binding elements for the AP-1 transcription complex com-

posed of Jun/Fos and ATF (42), and direct interaction between *SMAD3* and AP-1 has been reported to regulate TGFβ responses via binding its promoter (43). Analysis using the Tfsitescan tool predicted multiple AP-1 binding sites in *LOXL4* promoter (Supplementary Table 3 and Supplementary Methods, http://onlinelibrary.wiley.com/doi/10.1002/art.41163/abstract). Moreover, genomic analysis of GSE12493 (44,45) demonstrates up-regulated expression of *LOXL4* and *JUNB* in SScDFs treated with TGFβ for up to 24 hours (Supplementary Figure 1, http://online library.wiley.com/doi/10.1002/art.41163/abstract). These results suggest that TGFβ induces transcription of *LOXL4*, and this regulation may occur through AP-1.

This study introduces a range of SSc fibroblast lines from patients with clinically diagnosed diffuse or limited SSc, allowing assessment of fibrosis in a heterogeneous patient population. It is the first time that 3-D skin-like tissues have been used to investigate altered expression of *LOXL4* in SScDFs and suggests that increased *LOXL4* signaling in SSc enhances fibrosis through impaired collagen crosslinking. To date, most studies have been carried out in LOXL-2, a member of the LOX family, which has a role in fibrotic disease and metastasis (46,47). A clinical trial of simtuzumab, a humanized antibody to LOXL-2, has been performed in patients with idiopathic pulmonary fibrosis (IPF) and was found not to be efficacious (48). However, a recent comprehensive study demonstrated the different roles of LOX and LOXL subtypes in IPF and determined that LOX plays a prominent role in fibroblast activation (47). Moreover, Vadasz et al reported elevated LOX serum levels in SSc (20), although enhanced immunostaining of LOX was not identified in patients' skin. Substantial evidence, including that in this study, implicates *LOXL4* as a target that contributes to both vascular and fibrotic processes associated with ECM synthesis and remodeling (41,49). Most recently, *LOXL4* was found to promote cancer growth and metastasis by regulating cell cycle mediators (50,51). Therefore, antibody therapies directed against *LOXL4* may not merely target antifibrotic pathways but may also modulate pathways associated with anti-vasculopathic phenotypes, which regulate the proliferation and migration of endothelial cells.

Previously, we observed that genes in dermal fibroblasts were differentially expressed when fibroblasts were grown in 3-D tissue microenvironments compared to conventional monolayer cultures (11). Here, we found increased tissue rigidity in vector-transduced dermal fibroblasts compared to primary dermal fibroblasts (Figure 3D). This may be due to altered cell properties as a result of monolayer cell culture for 2 weeks during puromycin selection. On the other hand, the difference of stromal rigidity in HSE treated with β-APN was not significant in NDFs, while *ACTA2* was still decreased by β-APN (Figures 3I and K). We therefore hypothesized that the basal stiffness of NDF-containing HSE (~8 kPa elastic modulus) was too low to be distinguished by AFM. However, once the basal stiffness level had increased up to 4-fold in stromal tissues harboring transduced NDFs (~30 kPa elastic modulus), the mechanical

property of the tissue captured by AFM could be consistent with decreasing expression of *ACTA2* (Figures 3D and F). It appears that the basal stromal stiffness was maintained in HSE with SScDFs compared to NDFs for both primary and transduced fibroblasts, and the 2.5-fold increased level was consistent with the observed change in bleomycin-treated mouse skin (52).

It is widely accepted that an aberrant immune response and vasculopathy work together to orchestrate fibroblast activation in SSc. Conversely, recent studies illustrate that SSc fibroblasts can affect the polarization state of inflammatory cells (53). Additionally, comprehensive transcriptional analysis of skin biopsy samples has demonstrated systematic differences in the gene expression profile of dermal fibroblasts from SSc patients into subsets that can be distinguished by inflammatory and TGFβ gene signatures (7,54). These concepts are consistent with our transcriptome analysis results which demonstrated that over-expressed genes in SScDF-containing SAS (*IL1B*, *HLA-A*, *IL7*, *IL15*, and *CCL4*) detected by global gene analysis are functionally enriched for inflammatory response and interferon responses (Figure 2A). Thus, it may be useful in future studies to selectively incorporate SSc peripheral blood cells into 3-D skin-like tissues in order to assess how SSc dermal fibroblasts interact with them to affect localized, tissue-based immune responses. For example, the clinical efficacy of tocilizumab suggests that activated macrophages may be a promising candidate in modifying the disease process of SSc (55). Consistent with this is the notion that the 3-D tissues can provide a skin-like microenvironment characterized by interactions between SSc dermal fibroblasts and macrophages that could enable the investigation of how intercellular cross-talk may regulate the transdifferentiation of macrophage activation.

Future studies will be needed to define the interplay of fibroblasts with immune cells in order to devise improved targeted therapeutics to treat SSc. Emerging data support the notion that pathways central to normal growth and development play a key role in the regulation of aberrant fibroblast activation in acquired diseases and fibrosis (56–58). The skin-like tissue platforms used in the present study may provide an innovative approach to test pathways that mediate fibroblast activation, which may ultimately provide personalized strategies for the treatment of patients with SSc.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Huang 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. Huang, Black, Whitfield, Garlick.
Acquisition of data. Huang, Cai, Baugh, Liu, Smith, Watson, Popovich, Zhang, Stawski, Trojanowska, Pioli, Whitfield, Garlick.
Analysis and interpretation of data. Huang, Cai, Liu, Georgakoudi, Black, Whitfield, Garlick.

ROLE OF THE STUDY SPONSOR

Celdara Medical had no role in the study design, collection, interpretation or analysis 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 Celdara Medical.

REFERENCES

1. Ehrlich HP, Allison GM, Leggett M. The myofibroblast, cadherin, α smooth muscle actin and the collagen effect. Cell Biochem Funct 2006;24:63–70.

2. Rosin NL, Agabalyan N, Olsen K, Martufi G, Gabriel V, Biernaskie J, et al. Collagen structural alterations contribute to stiffening of tissue after split-thickness skin grafting. Wound Repair Regen 2016;24:263–74.

3. Clements PJ, Hurwitz EL, Wong WK, Seibold JR, Mayes M, White B, et al. Skin thickness score as a predictor and correlate of outcome in systemic sclerosis: high-dose versus low-dose penicillamine trial. Arthritis Rheum 2000;43:2445–54.

4. Steen VD, Medsger TA Jr. Improvement in skin thickening in systemic sclerosis associated with improved survival. Arthritis Rheum 2001;44:2828–35.

5. Marangoni RG, Varga J, Tourtellotte WG. Animal models of scleroderma: recent progress. Curr Opin Rheumatol 2016;28:561–70.

6. Sargent JL, Li Z, Aliprantis AO, Greenblatt M, Lemaire R, Wu MH, et al. Identification of optimal mouse models of systemic sclerosis by interspecies comparative genomics. Arthritis Rheumatol 2016;68:2003–15.

7. Greenblatt MB, Sargent JL, Farina G, Tsang K, Lafyatis R, Glimcher LH, et al. Interspecies comparison of human and murine scleroderma reveals IL-13 and CCL2 as disease subset-specific targets. Am J Pathol 2012;180:1080–94.

8. Long KB, Li Z, Burgwin CM, Choe SG, Martyanov V, Sassi-Gaha S, et al. The Tsk2/+ mouse fibrotic phenotype is due to a gain-of-function mutation in the PIIINP segment of the Col3a1 gene. J Invest Dermatol 2015;135:718–27.

9. Gardner H, Shearstone JR, Bandaru R, Crowell T, Lynes M, Trojanowska M, et al. Gene profiling of scleroderma skin reveals robust signatures of disease that are imperfectly reflected in the transcript profiles of explanted fibroblasts. Arthritis Rheum 2006;54:1961–73.

10. Whitfield ML, Finlay DR, Murray JI, Troyanskaya OG, Chi JT, Pergamenschikov A, et al. Systemic and cell type-specific gene expression patterns in scleroderma skin. Proc Natl Acad Sci U S A 2003;100:12319–24.

11. Huang M, Liu Z, Baugh L, DeFuria J, Maione A, Smith A, et al. Lysyl oxidase enzymes mediate TGF-β1-induced fibrotic phenotypes in human skin-like tissues. Lab Invest 2019;99:514–27.

12. Silver FH, Freeman JW, DeVore D. Viscoelastic properties of human skin and processed dermis. Skin Res Technol 2001;7:18–23.

13. Wang R, Brewster LP, Gleason RL Jr. In-situ characterization of the uncrimping process of arterial collagen fibers using two-photon confocal microscopy and digital image correlation. J Biomech 2013;46:2726–9.

14. Clarke DL, Carruthers AM, Mustelin T, Murray LA. Matrix regulation of idiopathic pulmonary fibrosis: the role of enzymes [review]. Fibrogenesis Tissue Repair 2013;6:20.

15. Kagan HM, Li W. Lysyl oxidase: properties, specificity, and biological roles inside and outside of the cell. J Cell Biochem 2003;88:660–72.

16. Leask A. Matrix remodeling in systemic sclerosis. Semin Immunopathol 2015;37:559–63.

17. Chanoki M, Ishii M, Kobayashi H, Fushida H, Yashiro N, Hamada T, et al. Increased expression of lysyl oxidase in skin with scleroderma. Br J Dermatol 1995;133:710–5.

18. Rajkumar VS, Howell K, Csiszar K, Denton CP, Black CM, Abraham DJ. Shared expression of phenotypic markers in systemic sclerosis indicates a convergence of pericytes and fibroblasts to a myofibroblast lineage in fibrosis. Arthritis Res Ther 2005;7:R1113–23.

19. Rimar D, Rosner I, Nov Y, Slobodin G, Rozenbaum M, Halasz K, et al. Lysyl oxidase is a potential biomarker of fibrosis in systemic sclerosis. Arthritis Rheumatol 2014;66:726–30.

20. Vadasz Z, Balbir Gurman A, Meroni P, Farge D, Levi Y, Ingegnoli F, et al. Lysyl oxidase—a possible role in systemic sclerosis-associated pulmonary hypertension: a multicentre study. Rheumatology (Oxford) 2019;58:1547–55.

21. Sethi A, Wordinger RJ, Clark AF. Gremlin utilizes canonical and non-canonical TGFβ signaling to induce lysyl oxidase (LOX) genes in human trabecular meshwork cells. Exp Eye Res 2013;113:117–27.

22. Proia DA, Kuperwasser C. Reconstruction of human mammary tissues in a mouse model. Nat Protoc 2006;1:206–14.

23. Thomas G, Burnham NA, Camesano TA, Wen Q. Measuring the mechanical properties of living cells using atomic force microscopy. J Vis Exp 2013. URL: https://www.jove.com/video/50497/measuring-mechanical-properties-living-cells-using-atomic-force.

24. Liu Z, Quinn KP, Speroni L, Arendt L, Kuperwasser C, Sonnenschein C, et al. Rapid three-dimensional quantification of voxel-wise collagen fiber orientation. Biomed Opt Express 2015;6:2294–310.

25. Quinn KP, Sullivan KE, Liu Z, Ballard Z, Siokatas C, Georgakoudi I, et al. Optical metrics of the extracellular matrix predict compositional and mechanical changes after myocardial infarction. Sci Rep 2016;6:35823.

26. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007;8:118–27.

27. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.

28. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1:417–25.

29. Lovell CR, Smolenski KA, Duance VC, Light ND, Young S, Dyson M. Type I and III collagen content and fibre distribution in normal human skin during ageing. Br J Dermatol 1987;117:419–28.

30. O'Reilly S, Cant R, Ciechomska M, van Laar JM. Interleukin-6: a new therapeutic target in systemic sclerosis? Clin Transl Immunol 2013;2:e4.

31. Baugh LM, Liu Z, Quinn KP, Osseiran S, Evans CL, Huggins GS, et al. Non-destructive two-photon excited fluorescence imaging identifies early nodules in calcific aortic-valve disease. Nat Biomed Eng 2017;1:914–24.

32. Sood D, Chwalek K, Stuntz E, Pouli D, Du C, Tang-Schomer M, et al. Fetal brain extracellular matrix boosts neuronal network formation in

3D bioengineered model of cortical brain tissue. ACS Biomater Sci Eng 2016;2:131–40.

33. Liu Z, Pouli D, Sood D, Sundarakrishnan A, Hui Mingalone CK, Arendt LM, et al. Automated quantification of three-dimensional organization of fiber-like structures in biological tissues. Biomaterials 2017;116:34–47.

34. Cao L, Lafyatis R, Burkly LC. Increased dermal collagen bundle alignment in systemic sclerosis is associated with a cell migration signature and role of Arhgdib in directed fibroblast migration on aligned ECMs. PLoS One 2017;12:e0180751.

35. Lee R, Perry B, Heywood J, Reese C, Bonner M, Hatfield CM, et al. Caveolin-1 regulates chemokine receptor 5-mediated contribution of bone marrow-derived cells to dermal fibrosis. Front Pharmacol 2014;5:140.

36. Milano A, Pendergrass SA, Sargent JL, George LK, McCalmont TH, Connolly MK, et al. Molecular subsets in the gene expression signatures of scleroderma skin. PLoS One 2008;3:e2696.

37. Arciniegas E, Neves CY, Carrillo LM, Zambrano EA, Ramírez R. Endothelial-mesenchymal transition occurs during embryonic pulmonary artery development. Endothelium 2005;12:193–200.

38. Frid MG, Kale VA, Stenmark KR. Mature vascular endothelium can give rise to smooth muscle cells via endothelial-mesenchymal trans-differentiation: in vitro analysis. Circ Res 2002;90:1189–96.

39. Willis BC, Borok Z. TGF-β-induced EMT: mechanisms and implications for fibrotic lung disease. Am J Physiol Lung Cell Mol Physiol 2007;293:L525–34.

40. Hashimoto N, Phan SH, Imaizumi K, Matsuo M, Nakashima H, Kawabe T, et al. Endothelial-mesenchymal transition in bleomycin-induced pulmonary fibrosis. Am J Respir Cell Mol Biol 2010;43:161–72.

41. Busnadiego O, Gonzalez-Santamaría J, Lagares D, Guinea-Viniegra J, Pichol-Thievend C, Muller L, et al. LOXL4 is induced by transforming growth factor β1 through Smad and JunB/Fra2 and contributes to vascular matrix remodeling. Mol Cell Biol 2013; 33:2388–401.

42. Sethi A, Mao W, Wordinger RJ, Clark AF. Transforming growth factor-β induces extracellular matrix protein cross-linking lysyl oxidase (LOX) genes in human trabecular meshwork cells. Invest Ophthalmol Vis Sci 2011;52:5240–50.

43. Verrecchia F, Vindevoghel L, Lechleider RJ, Uitto J, Roberts AB, Mauviel A. Smad3/AP-1 interactions control transcriptional responses to TGF-β in a promoter-specific manner. Oncogene 2001;20:3332–40.

44. Sargent JL, Milano A, Bhattacharyya S, Varga J, Connolly MK, Chang HY, et al. A TGFβ-responsive gene signature is associated with a subset of diffuse scleroderma with increased disease severity. J Invest Dermatol 2010;130:694–705.

45. Johnson ME, Mahoney JM, Taroni J, Sargent JL, Marmarelis E, Wu MR, et al. Experimentally-derived fibroblast gene signatures identify molecular pathways associated with distinct subsets of systemic sclerosis patients in three independent cohorts. PLoS One 2015;10:e0114017.

46. Tadmor T, Bejar J, Attias D, Mischenko E, Sabo E, Neufeld G, et al. The expression of lysyl-oxidase gene family members in myeloproliferative neoplasms. Am J Hematol 2013;88:355–8.

47. Vadasz Z, Kessler O, Akiri G, Gengrinovitch S, Kagan HM, Baruch Y, et al. Abnormal deposition of collagen around hepatocytes in Wilson's disease is associated with hepatocyte specific expression of lysyl oxidase and lysyl oxidase like protein-2. J Hepatol 2005;43:499–507.

48. Raghu G, Brown KK, Collard HR, Cottin V, Gibson KF, Kaner RJ, et al. Efficacy of simtuzumab versus placebo in patients with idiopathic pulmonary fibrosis: a randomised, double-blind, controlled, phase 2 trial. Lancet Respir Med 2017;5:22–32.

49. Jiang WP, Sima ZH, Wang HC, Zhang JY, Sun LS, Chen F, et al. Identification of the involvement of LOXL4 in generation of keratocystic odontogenic tumors by RNA-Seq analysis. Int J Oral Sci 2014;6:31–8.

50. Li R, Wang Y, Zhang X, Feng M, Ma J, Li J, et al. Exosome-mediated secretion of LOXL4 promotes hepatocellular carcinoma cell invasion and metastasis. Mol Cancer 2019;18:18.

51. Shao J, Lu J, Zhu W, Yu H, Jing X, Wang YL, et al. Derepression of LOXL4 inhibits liver cancer growth by reactivating compromised p53. Cell Death Differ 2019;26:2237–52.

52. Lagares D, Santos A, Grasberger PE, Liu F, Probst CK, Rahimi RA, et al. Targeted apoptosis of myofibroblasts with the BH3 mimetic ABT-263 reverses established fibrosis. Sci Transl Med 2017;9:eaal3765.

53. Asano Y. Systemic sclerosis [review]. J Dermatol 2018;45:128–38.

54. Kumar M, Makonchuk DY, Li H, Mittal A, Kumar A. TNF-like weak inducer of apoptosis (TWEAK) activates proinflammatory signaling pathways and gene expression through the activation of TGF-β-activated kinase 1. J Immunol 2009;182:2439–48.

55. Khanna D, Denton CP, Jahreis A, van Laar JM, Frech TM, Anderson ME, et al. Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial. Lancet 2016;387:2630–40.

56. Gilbane AJ, Denton CP, Holmes AM. Scleroderma pathogenesis: a pivotal role for fibroblasts as effector cells [review]. Arthritis Res Ther 2013;15:215.

57. McAnulty RJ. Fibroblasts and myofibroblasts: their source, function and role in disease. Int J Biochem Cell Biol 2007;39:666–71.

58. Kendall RT, Feghali-Bostwick CA. Fibroblasts in fibrosis: novel roles and mediators [review]. Front Pharmacol 2014;5:123.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Differential DNA Methylation of Networked Signaling, Transcriptional, Innate and Adaptive Immunity, and Osteoclastogenesis Genes and Pathways in Gout

Zengmiao Wang,[1] Ying Zhao,[1] Amanda Phipps-Green,[2] Ru Liu-Bryan,[3] Arnoldas Ceponis,[1] David L. Boyle,[1] Jun Wang,[1] Tony R. Merriman,[2] Wei Wang,[1] and Robert Terkeltaub[3] (iD)

Objective. In gout, autoinflammatory responses to urate crystals promote acute arthritis flares, but the pathogeneses of tophi, chronic synovitis, and erosion are less well understood. Defining the pathways of epigenomic immunity training can reveal novel pathogenetic factors and biomarkers. The present study was undertaken to seminally probe differential DNA methylation patterns utilizing epigenome-wide analyses in patients with gout.

Methods. Peripheral blood mononuclear cells (PBMCs) were obtained from a San Diego cohort of patients with gout (n = 16) and individually matched healthy controls (n = 14). PBMC methylome data were processed with ChAMP package in R. ENCODE data and Taiji data analysis software were used to analyze transcription factor (TF)–gene networks. As an independent validation cohort, whole blood DNA samples from New Zealand Māori subjects (n = 13 patients with gout, n = 16 control subjects without gout) were analyzed.

Results. Differentially methylated loci clearly separated gout patients from controls, as determined by hierarchical clustering and principal components analyses. *IL23R*, which mediates granuloma formation and cell invasion, was identified as one of the multiple differentially methylated gout risk genes. Epigenome-wide analyses revealed differential methylome pathway enrichment for B and T cell receptor signaling, Th17 cell differentiation and interleukin-17 signaling, convergent longevity regulation, circadian entrainment, and AMP-activated protein kinase signaling, which are pathways that impact inflammation via insulin-like growth factor 1 receptor, phosphatidylinositol 3-kinase/Akt, NF-κB, mechanistic target of rapamycin signaling, and autophagy. The gout cohorts overlapped for 37 (52.9%) of the 70 TFs with hypomethylated sequence enrichment and for 30 (78.9%) of the 38 enriched KEGG pathways identified via TFs. Evidence of shared differentially methylated gout TF-gene networks, including the NF-κB activation–limiting TFs *MEF2C* and *NFATC2*, pointed to osteoclast differentiation as the most strongly weighted differentially methylated pathway that overlapped in both gout cohorts.

Conclusion. These findings of differential DNA methylation of networked signaling, transcriptional, innate and adaptive immunity, and osteoclastogenesis genes and pathways suggest that they could serve as novel therapeutic targets in the management of flares, tophi, chronic synovitis, and bone erosion in patients with gout.

INTRODUCTION

Acute gouty arthritis is promoted by NLRP3 inflammasome–driven autoinflammatory responses to urate crystals (1,2). Gout flares are characteristically superimposed on urate crystalline macroaggregates, including in tophi, the majority of whose volume is composed of variable foci of fibrosis, chronic synovitis, and granulomatous changes (3,4). B and T lymphocytes and plasma cells are among the cells surrounding crystals in tophi (5). Furthermore, osteoclastogenesis is prominent at the tophus–bone interface, and is linked with bone erosion (3). Observations of enhanced in vitro osteoclastogenesis of RANKL- and macrophage colony-stimulating factor (M-CSF)–treated peripheral blood mononuclear cells (PBMCs) in patients with severe, tophaceous disease have suggested that molecular programming in PBMCs is a feature of advanced gout (6). However, major knowledge gaps remain with regard to the mechanisms promoting chronic synovitis, tophus development and associated granulomatous changes, and bone erosion in gout (1–3,5,6).

Altered innate and adaptive immunity in gout could reflect dietary and genetic factors (1), including the effects of common genetic variants linked to ion transport and the effects of purine, carbohydrate, or lipoprotein metabolism (7–9). However, gout-associated gene variants generally have weak effects on gene expression and transcript processing or stability (7). Epigenetic changes regulate gene expression and also cell differentiation and function in response to diet, lifestyle, and other factors (10,11). Epigenetic training of both adaptive and innate immunity, including through reversible histone modifications (12,13), can regulate inflammation phenotypes. Notably, hyperuricemia modulates the inflammatory responses of monocytes, and increases murine urate crystal–induced inflammation in vivo (13). Repeated exposure of leukocytes to urate crystals, interleukin-1β (IL-1β), and other mediators of inflammation can also modify the epigenome (12,13).

Epigenetic changes include DNA methylation, defined by covalent methyl group binding to a 5′ carbon cytosine nucleotide, typically followed by guanine (termed CpG) (11). This relatively stable epigenetic change regulates both myeloid and lymphoid lineage development (14) and inflammatory gene expression in response to environmental stressors, aging, pathogens, endogenous danger signals, and many cytokines (11). DNA methylome changes are implicated in obesity and type 2 diabetes (15), both of which are comorbidities frequently associated with gout (16). Moreover, chemokine *CCL2* promoter hypomethylation was found in patients with gout (17). Therefore, in the present study, we analyzed epigenome-wide DNA methylation changes in patients with gout.

PATIENTS AND METHODS

Subject recruitment. *San Diego cohort*. Study of the San Diego cohort was preapproved by the University of California, San Diego Institutional Review Board, and informed consent was obtained from all subjects. Outpatients diagnosed as having gout, according to the American College of Rheumatology (ACR)/European League Against Rheumatism 2015 gout classification criteria (18), were recruited along with individually matched healthy controls, by rheumatologist referral, invitation via health record searches, and solicitation from pools of healthy subjects. Subjects undergoing hemodialysis or those with stage 4 chronic kidney disease were excluded.

Gout patients (n = 17) were recruited first, followed by individual matching to healthy control subjects who did not have a history of gout (n = 16). Patients and controls were matched by sex, age (≤9 years difference), and body mass index (BMI) (<30 kg/m^2 or ≥30 kg/m^2). One gout patient was not matchable due to advanced age. Sixteen gout patients and 14 healthy controls agreed to DNA testing.

All subjects were evaluated in a single visit, with limited clinical laboratory studies (measurement of serum urate and serum creatinine levels) and examination for palpable tophi. No subjects in the San Diego cohort were smokers. At the time of blood sampling, gout patients had to be ≥7 days removed from treatment with glucocorticoids or colchicine, and ≥14 days from resolution of the last gout flare.

New Zealand cohort. Study of the Māori subjects in the New Zealand cohort was approved by the Institutional Review Board of the New Zealand Health and Disability Ethics Committee. Subjects were part of a diabetes-focused substudy comprising 40 male and 40 female Māori subjects from Aotearoa/New Zealand (19). All subjects gave their informed consent, which allowed extraction of their genomic DNA from whole blood. Forty control subjects who self-reported a diagnosis of type 2 diabetes and 40 healthy control subjects were matched to the patients with gout by sex, age, and BMI. Selection was prioritized to those control subjects who had a normal fasting glucose level, or secondarily to those control subjects who had normal lipid levels and were without evident heart disease. In all, 30 patients with gout (23 male, 7 female), whose diagnosis met the ACR 1977 preliminary gout classification criteria (20), and 23 control subjects without gout (16 male, 7 female) were considered for inclusion in the New Zealand cohort. After excluding subjects who had missing data for smoking status and presence or absence of chronic kidney disease, we assembled a cohort of 13 gout patients and 16 control subjects without gout.

Isolation of PBMCs and DNA, methylation assay, and data preprocessing. PBMCs from EDTA-anticoagulated blood samples collected from subjects in the San Diego cohort were isolated using Ficoll-Paque Plus (GE Healthcare). DNA was isolated from washed PBMC pellets using a DNeasy blood kit (Qiagen), and thereafter concentrated using Amicon Ultra (Millipore). Whole blood DNA from subjects in the New Zealand cohort was isolated using a GuCl-guanadinium–based method. DNA was converted by sodium bisulfite modification, and then hybridized to Illumina

Infinium MethylationEPIC BeadChip arrays. Samples were randomized across plates and chip positions.

Illumina Methylation Beadarray 850K flatform data were processed using the ChAMP package (21) in R. Probes were filtered (using the champ.filter function) based on the following conditions: 1) probes with detected P values of >0.01; 2) probes with <3 beads in ≥5% of samples examined; 3) all non-CpG probes; 4) all single-nucleotide polymorphism (SNP)–related probes; 5) all multi-hit probes; and 6) all probes in chromosomes X and Y. The masking list included probes with SNPs close to the 3′ end, probes with an SNP in the extension base that cause a color channel switch from the designated annotation, and probes with reason to be mapped (22). SNP probes on the array were also removed. Methylation at each locus was reported as a beta value, ranging from β = 0 (unmethylated) to β = 1 (fully methylated). Beta mixture quantile normalization (using the champ.norm function) was applied for normalization of the data. Principal components analysis (PCA) (with the champ.SVD function) showed the top PCs significantly correlated with the slide and array, and batch effects were corrected for normalized data (using the champ.runCombat function).

Adjustment for cell type composition, and identification of gout-associated differential methylation loci (DMLs) and mapping to genes. Cell type composition is a cofactor in DNA methylation from blood samples. We downloaded healthy donor leukocyte type–selective DNA methylation data from the NCBI Gene Expression Omnibus data set (accession no. GSE110554), serving as a reference for DNA methylation measured on the 850K platform. Healthy donor cell types comprised neutrophils (n = 6 reference donors) in the whole blood from the New Zealand cohort, and monocytes (n = 6 reference donors), B lymphocytes (n = 6 reference donors), CD4+ T cells (n = 7 reference donors), CD8+ T cells (n = 6 reference donors), and natural killer cells (n = 6 reference donors) in both cohorts.

Automatic selection using the minfi package was used to identify probes specific to each cell type. The top 50 hyper- and hypomethylated probes were selected for each cell type. In total, 600 probes were used to estimate cell type composition. We estimated the distribution of cell types in a manner previously described (23), and added cell type proportion to the linear regression to adjust for its effects in both cohorts (24). Other cofactors used for adjustment in the linear regression models included age and smoking status (see Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract). For the San Diego cohort, ethnicity was included in the linear regression.

DMLs related to gout were identified using the limma package (25) in R, for differential analysis of microarray, RNA-sequencing, or methylation data using linear models. For both cohorts, the Benjamini-Hochberg procedure (i.e., false discovery rate [FDR]) was used to correct for multiple hypothesis testing. Two criteria were

used to weight the DMLs: 1) FDR cutoff <0.05, and 2) differences in average beta values greater than the 99% quantile. To adjust for effects of allopurinol, we fitted a linear regression model using data from gout cases only. Two criteria to identify DMLs related to allopurinol were applied: 1) P value cutoff <0.05 (a liberal significance threshold was used because of the small sample size), and 2) selection of the top 1% of probes with the largest difference in beta values between those taking and those not taking allopurinol. For the downstream analyses, we used the DML sets derived after removal of the overlapping DMLs related to allopurinol.

Mapping from DML to gene was based on the file "MethylationEPIC_v-1-0_B4.csv" (provided by Illumina). We classified DML genes based on the overall changes in DMLs and the locations of the DMLs. Sites were designated as "hypermethylated" when gout cases had more DNA methylation, and designated as "hypomethylated" when gout cases had less DNA methylation, as compared to controls. DML genes with only 1 differentially methylated probe were designated as either promoter-hypo, promoter-hyper, body-hypo, or body-hyper. DML genes with ≥2 differentially methylated probes were classified based on each DML. For example, if there were 2 hypomethylated DMLs within a gene, the classification was "promoter-hypo;body-hypo." Based on the UCSC RefGene Groups from the Illumina annotation file, the transcriptional start sites TSS200 and TSS1500 were defined as "promoter" sites, while other features were defined as "gene-body" sites.

Hierarchical clustering, PCA, and pathway enrichment analyses, and identification of TF binding sites and TF-gene networks. Based on the DMLs identified, hierarchical clustering was employed via the heatmap.2 function in R, with "maximum" set as the distance function and "ward.D2" as the agglomeration method. PCA was performed for the visualization of DMLs (using the plotMDS function). Because neither hierarchical clustering nor PCA tested a null hypothesis, no P values were used to test significance. KEGG pathway enrichment analysis was performed based on DML gene sets, using ConsensusPathDB (26) (see http://consensuspathdb.org/). The following criteria were applied to define enriched pathways: 1) >2 overlaps between DML genes and pathway genes, and 2) q value cutoff <0.01. R package KEGGprofile was used for visualizing KEGG pathways. The keywords "GWAS" and "gout" were used for searching the literature related to genome-wide association study (GWAS) results, with 110 GWAS gout risk genes tallied (see Supplementary Table 1, available on the Arthritis & Rheumatology web site at http://online library.wiley.com/doi/10.1002/art.41173/abstract).

For TF binding sites in DML regions, sequences from 250 bp upstream to 250 bp downstream of the DML position were extracted; 500-bp sequences were divided into hypo-sets and hyper-sets based on DML status. AME (run from the MEME suite) (27) was employed to find motifs enriched in these 2 sets, separately, with the motif set downloaded from JASPAR (JASPAR2018_CORE_vertebrates_non-redundant.meme file) (28).

We ran AME with the following parameters: –verbose 1 –scoring avg –method fisher –hit-lo-fraction 0.25 –control –shuffle– –kmer 2. Motifs were enriched in the sequence set if the corresponding E value was below 1×10^{-10}.

For analyses of TF-gene networks comparing PBMCs from the San Diego and New Zealand cohorts, ATAC-seq data for B cells, T cells, CD14+ monocytes, CD4+ T helper cells, and CD8+ αβ T cells were downloaded from the ENCODE project database. For each cell type, a TF-gene regulatory network was constructed using Taiji data analysis software (29), to assess the global influence of each TF from the network perspective. The top 10% of edges by weights from Taiji were kept.

The networks were then combined, and the 335 DML genes falling into common classes in both cohorts were mapped (including 2 TFs with differential methylation and hypomethylated sequence enrichment in both cohorts), leaving 224 genes and 1,232 edges. Modules were identified using the Affiliation Graph Model AGMfit algorithm, to detect network communities (i.e., modules) by fitting a probabilistic generative model for given networks using maximum

likelihood estimation. Fourteen TFs were identified on the basis of membership in all 5 different modules, and these 14 TFs were identified as DML genes in both cohorts. Since genes were ranked by the number of modules (without hypothesis testing), no P values were used to test significance. The identified DMLs and beta estimates (with standard errors and P values) in the 2 cohorts are listed in the Supplementary Table 1 and Supplementary Table 2 Appendices; in addition, the Supplementary List of GWAS Genes provides data on the relationship between the GWAS genes, CpG sites, SNP sites, and linkage disequilibrium (LD) blocks in different populations (all available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract).

RESULTS

The gout DNA methylome signature. Most gout patients in the San Diego cohort were being treated with a stable dose of allopurinol, and overall the serum urate levels were comparable to those in healthy controls (characteristics
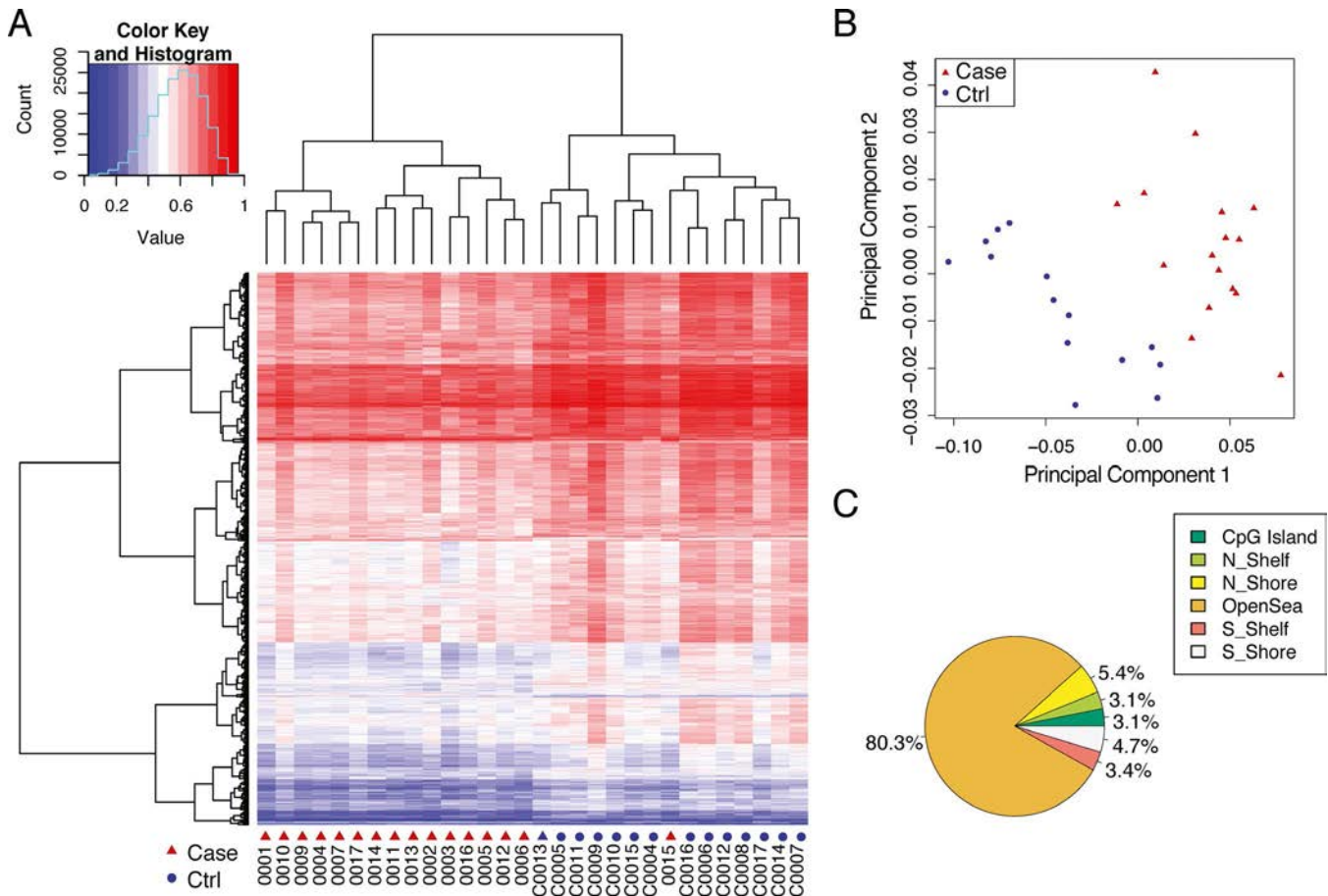
Figure 1. Overview of differential DNA methylation in the San Diego gout cohort. **A**, Separation of gout cases from healthy controls using hierarchical clustering. Rows represent methylation probes; columns show individual DNA samples from cases (red) or controls (blue). **B**, Separation of gout cases from controls using principal components analysis. The proportion of variance was 64.15% for the first principal component (PC1) and 6.50% for the second principal component (PC2). In total, PC1 and PC2 contained 70.65% of the variance. **C**, Genomic locations of the 5,438 differential methylation loci. Shelf = 2–4 kb from CpG island; Shore = 0–2 kb from CpG island; OpenSea = >4 kb away from CpG island; N = upstream (5′) of CpG island; S = downstream (3′) of CpG island.

of the cohort are listed in Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.41173/abstract). The study workflow is depicted in Supplementary Figure 2 (http://online library.wiley.com/doi/10.1002/art.41173/abstract).

Epigenome-wide DNA methylation covered 865,859 loci. After filtering to exclude low-quality probes, 755,223 probes remained, with 5,438 promoter or gene-body DMLs. Of the DMLs detected, 96.7% were hypomethylation changes in gout, mostly (80.3%) in open sea regions (i.e., >4 kb away from CpG islands). The fewest DMLs (3.1% each) were in CpG islands (typically located at promoters with a CG:GC ratio of >0.6) or N Shelf (2–4 kb upstream of CpG islands) (Figure 1C).

The gout DMLs were mapped to 2,488 genes (designated DML genes), with 70.8% demonstrating gene-body hypomethylation, which modulates gene expression, depending on the context and whether or not the specific CpG site is in an enhancer region that is relevant to that specific gene or neighboring genes. The identified DMLs allowed clear separability of gout patients from controls, as shown by hierarchical clustering (Figure 1A) and PCA (Figure 1B).

Enrichment of GWAS gout risk genes among the identified gout DML genes, and identification of differentially methylated pathways. We found 22 DML genes that overlapped with known gout risk genes from past GWAS (Figure 2A and Supplementary Table 3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract); box plots showing DNA methylation levels of the genes are shown in Figure 2D. Several of the identified loci encode gene products involved in vascular function or ion transport (e.g., *SLC2A9*, *ABCC9*). Other loci, such as *PRKG2*, modulate innate and/or adaptive immunity, and *BDKRB2* encodes a bradykinin receptor supporting neutrophil recruitment in experimental gout (30). The *PPARGC1A* product drives mitochondrial biogenesis, inhibits insulin resistance, and is antiinflammatory. The *TRPA1* product mediates pain and neuroinflammation. The probe cg13841979 of *SLC2A9* and the previous GWAS-identified rs1014290 were in the same LD block, as were cg05205932 and rs7688672 and rs6837293 for *PRKG2* (data not shown). DML genes that had DMLs and gout-associated SNPs in the same LD block included *A1CF, CFTR*, *ABCC9*, *MYL2*, *PDZK1*, and *B4GALT1* (data not shown). These findings buttress past evidence of the potential functional impact of GWAS SNPs in gout.

We detected hypomethylation of *IL23R* (cg19227223, cg13549101), *SLC2A9, ABCC9*, *PRKG2*, and *BDKRB2* (Figure 2D). Several of the enriched KEGG pathways that were identified in gout patients (e.g., those for renin and insulin secretion) (Figures 2B and C) are related to hyperuricemia and hypertension, metabolic syndrome, and type 2 diabetes (16). Other pathways included those for phospholipase D, regulation of the actin cytoskeleton and Rap1 signaling, cell migration, cell adhesion, extracellular matrix (ECM)–receptor interaction, and MAPK and calcium signaling pathways. In addition, we detected enrichment of DMLs in circadian entrainment (Figure 2B) and inflammation mediator–regulated transient receptor potential (TRP) channels involved in neuroinflammation (Figure 2B).

AMP-activated protein kinase (AMPK), a nutritional biosensor suppressed by dietary excesses and modulated by inflammation, exerts broad effects on the epigenome and is, in turn, regulated by epigenetic changes, and also limits urate crystal–induced inflammation (31,32). *PRKAG2*, which encodes an AMPK heterotrimer γ-subunit isoform, was gene-body hypomethylated (cg09817217, cg07012178) in gout patients (Figure 2D). AMPK signaling was among the KEGG pathways enriched in both the DML gene set and GWAS gene set (q value 8.1×10^{-4}) (Figure 2C and Supplementary Figure 3A, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract). *PPP2R5C* (for protein phosphatase 2A regulatory subunit B family) negatively regulates AMPK activity via phosphatase activity for the AMPK α-chain catalytic site threonine. In this study, *PPP2R5C* was also gene-body hypomethylated. AMPK-intersecting pathways with DML enrichment in gout included phosphatidylinositol 3-kinase (PI3K)/Akt, circadian entrainment, and longevity regulation (33) (Figures 2B and C and Supplementary Figures 3B–D [http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract]).

TF motif scanning in DML regions and pathway changes. Transcriptional signaling controls the development, differentiation, and function of leukocytes. We identified 75 hypomethylated TF motifs, corresponding to 65 TFs, enriched in the 500-bp sequences centered at the hypomethylated loci. Examples of hypomethylated TF motifs included the circadian clock regulators *CLOCK* and *ARNT* (for aryl hydrocarbon nuclear receptor) and genes encoding *STAT2* and multiple components of *AP1* (for activator protein 1), *IRF1* (for interferon regulatory transcription factor 1) (34), and *NFATC2* and *NFATC3*. Among the top identified TFs (Figure 3A), *NFATC2* was 1 of 2 TFs that were themselves hypomethylated in the gene-body region, and all genes regulated by *NFATC2* were differentially methylated. *MEF2C* (35) also showed gene-body hypomethylation in patients with gout (Figure 3A). For *NFATC2,* 1 DML was intronic. For *MEF2C*, 1 CpG was in an exon region. Specific examples of motifs of *NFATC2* and *MEF2C* that exhibited the conserved DNA binding pattern are shown in Figure 3B. Detected CpG sites and methylation levels for the CpG sites along the *NFATC2* and *MEF2C* genes in gout patients compared to controls are depicted in Figures 3C and D, respectively.

Analyses of gout DNA methylome changes in certain TFs pointed to differential DNA methylation of multiple biologic pathways, including the osteoclast differentiation pathway (Figure 3A and Supplementary Figure 4, available on the *Arthritis & Rheumatology*
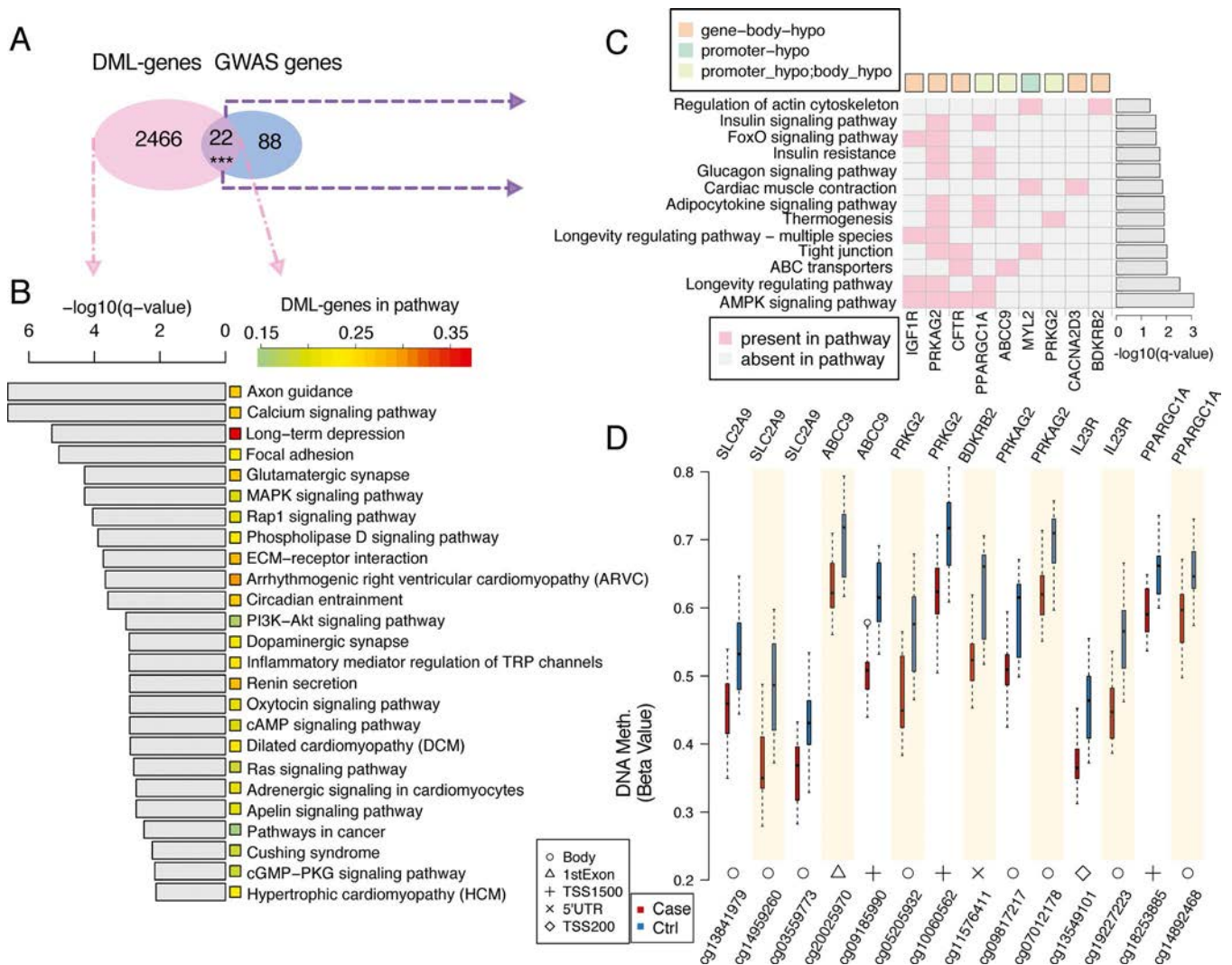
Figure 2. Differential methylation loci (DML) genes and associated pathways in the San Diego gout cohort. **A**, Number of DMLs and overlaps with gout genetic risk loci genes previously identified from genome-wide association studies (GWAS). *** = $P = 3.3 \times 10^{-4}$ from hypergeometric distribution. **B**, Signaling pathways strongly associated with the gout DML genes. Colored squares represent the percentage of DMLs in each pathway. **C**, Pathways associated with the 9 DML genes overlapping with the GWAS gout genetic risk loci genes. Darker green and orange squares represent hypomethylation (hypo) at the promoter or gene-body regions, respectively, while lighter green squares represent DML genes with multiple DMLs. Gray-shaded bars in **B** and **C** indicate the significance (q value) of each association. *P* values for the pathways listed in **B** and **C** were calculated using the hypergeometric distribution and transformed into q values for the multiple hypothesis testing correction. **D**, DNA methylation (Meth.) levels of the CpG sites in 7 DML genes in gout cases and controls. Data are shown as box plots. Each box represents the upper and lower interquartile range (IQR). Lines inside the box represent the median. Whiskers represent 1.5 times the upper and lower IQRs. The circle indicates an outlier. Locations of the CpG sites are shown at the bottom of the box plots. ECM = extracellular matrix; PI3K = phosphatidylinositol 3-kinase; TRP = transient receptor potential; PKG = cGMP-dependent protein kinase; AMPK = AMP-activated protein kinase; TSS1500 = 200–1,500 bases upstream of the transcriptional start site (TSS); 5′UTR = 5′-untranslated region; TSS200 = 0–200 bases upstream of the TSS.

web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract). Furthermore, multiple KEGG pathways that are central to adaptive immunity, specifically the B cell receptor signaling, Th1 and Th2 cell differentiation, Th17 cell differentiation, and IL-17 signaling pathways, were also linked to differential DMLs and DNA methylated TFs in PBMCs from gout patients (Figure 3A).

Validation of the DNA methylation signature of genes, pathways, and TFs in an independent gout cohort. The New Zealand cohort, which was enriched for obesity and hypertension and was previously selected to examine patients with type 2 diabetes, comprised exclusively subjects of New Zealand Māori and Cook Island Māori (East Polynesian) descent
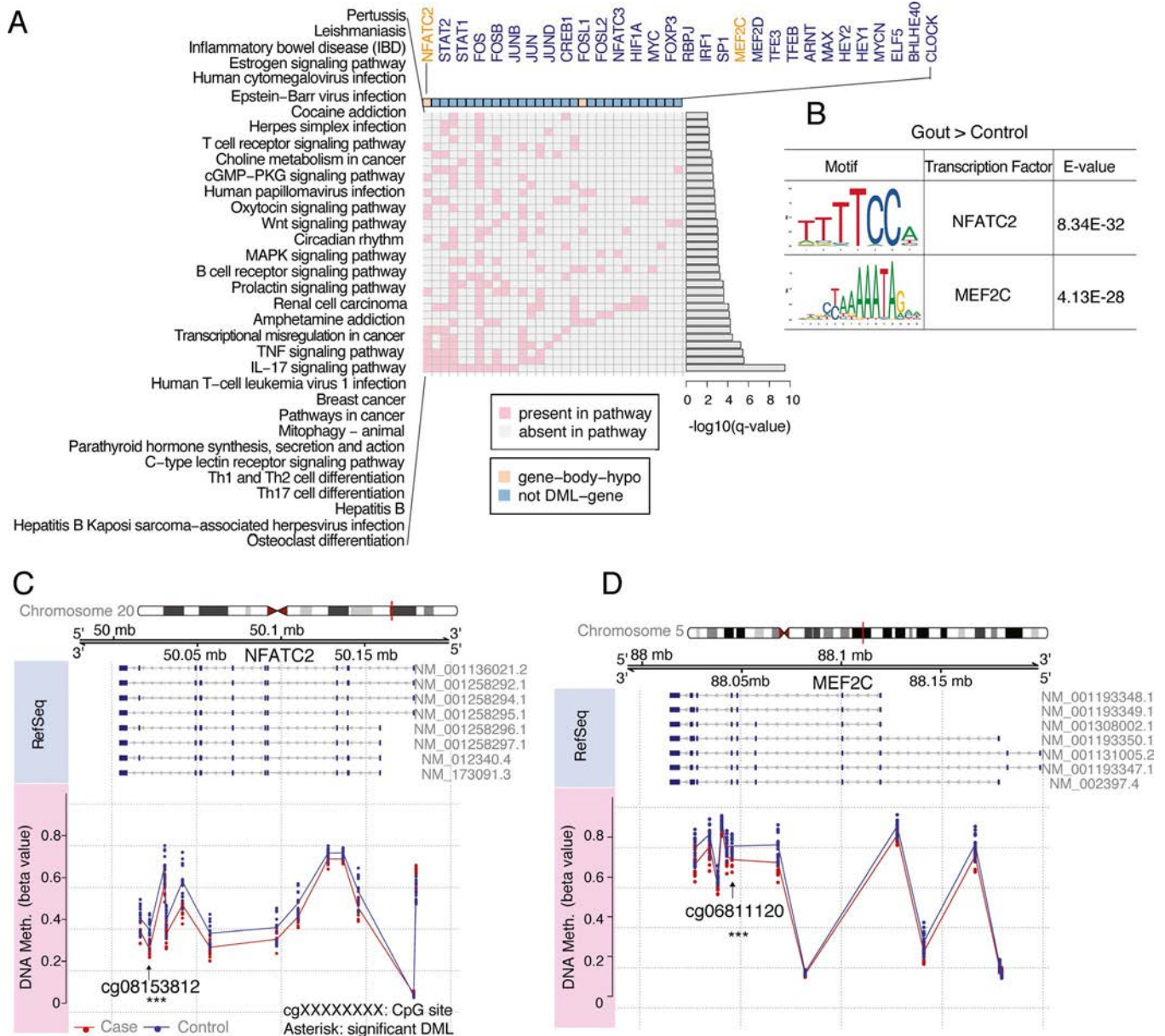
Figure 3. Transcription factor (TF) motifs, CpG sites, and differential DNA methylation (Meth.) in the San Diego gout cohort. **A**, Top biologic pathways in gout identified through TF motif scanning in hypomethylated differential methylation loci (DML) regions. Rows show the pathways, and columns show the TFs associated with the enriched pathways. These TFs in gout cases were either DNA hypomethylated (hypo) at the gene-body (orange) or not differentially DNA methylated (blue). *P* values for pathways were calculated using hypergeometric distribution and transformed into q values. **B**, Examples of motifs of the significantly enriched TFs in gout cases. The E value is the measurement of significance of a motif. **C** and **D**, CpG sites along the TFs *NFATC2* (**C**) and *MEF2C* (**D**). The top line depicts the chromosome, with the gene region highlighted in red. The second line depicts the genome coordinate. The third line lists the RefSeq annotation for each gene. Numbers designated with NM are the transcript IDs for each gene. Exons are shown as dark blue rectangles, and introns as gray lines. The graph at bottom shows the methylation levels for the CpG sites along the genes in the gout and control cohorts. Significant DMLs are labeled with asterisks. PKG = cGMP-dependent protein kinase; TNF = tumor necrosis factor; IL-17 = interleukin-17.

(19) (the characteristics of the cohort are listed in Supplementary Table 4, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract). Using the same cutoff as for the San Diego cohort, 4,845 DMLs were identified, corresponding to 2,530 genes. Hierarchical clus-

tering and PCA clearly separated the New Zealand gout patients from the control subjects without gout (Figures 4A and B).

To probe for the most fundamental methylome changes in gout, we assessed and ranked sharing of methylome findings by looking at DML genes, pathways enriched in DMLs, and TFs and
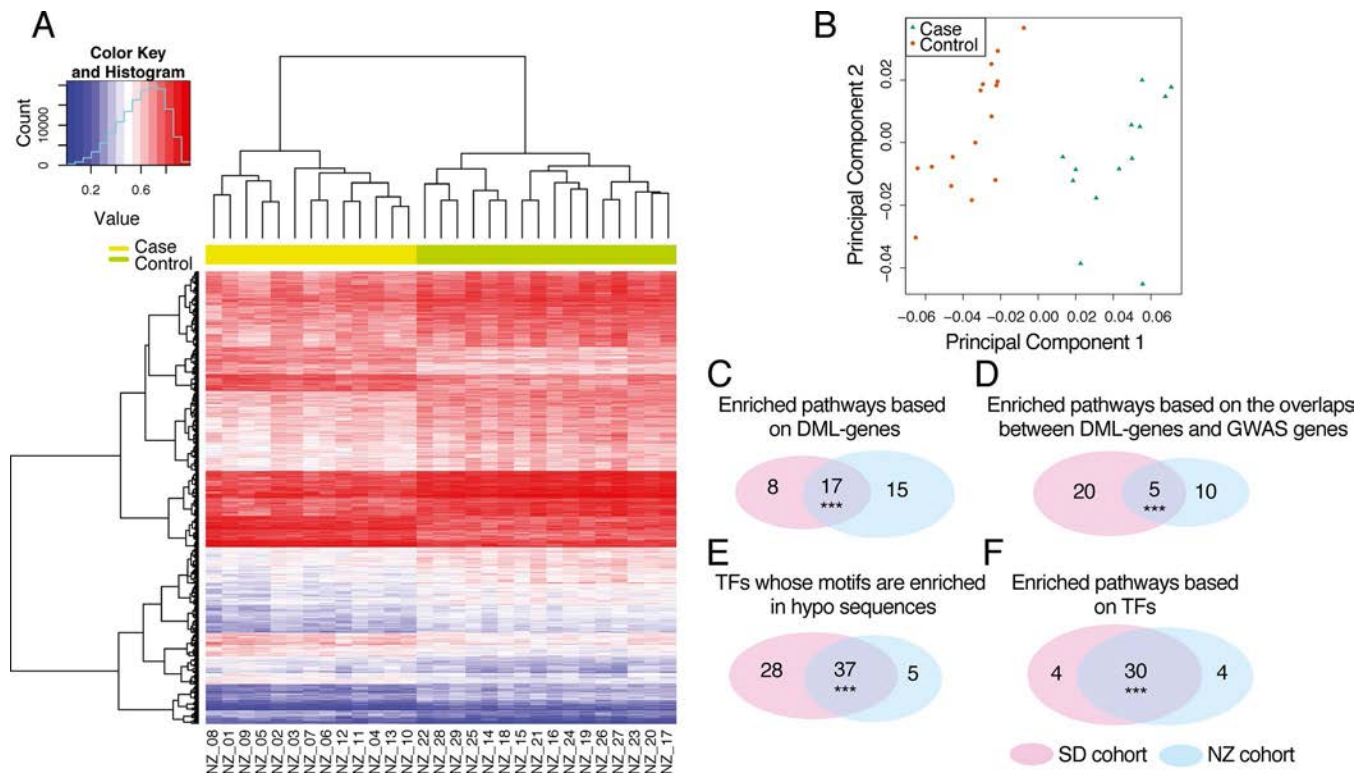
Figure 4. Analysis of the DNA methylome in the New Zealand (NZ) gout validation cohort. **A**, Separation of gout cases from controls using hierarchical clustering. Heatmaps show all of the differential methylation loci (DMLs) in the New Zealand cohort. **B**, Separation of gout cases from controls using principal components analysis in the New Zealand cohort. **C–F**, Venn diagrams illustrating sharing of DMLs, and related changes, between the San Diego (SD) and New Zealand cohorts, for enriched pathways based on DML genes ($P = 2.7 \times 10^{-14}$) (**C**), enriched pathways based on the overlaps between DML genes and genome-wide association study (GWAS) gout risk genes (**D**), transcription factors (TFs) whose motifs are enriched in hypomethylated (hypo) DML regions ($P \leq 5 \times 10^{-324}$) (**E**), and enriched pathways based on TFs ($P \leq 5 \times 10^{-324}$) (**F**). *** = P value is calculated from hypergeometric distribution.

related gene networks. Both gout cohorts shared 656 of 4,362 DML genes ($P \leq 5 \times 10^{-324}$ from hypergeometric distribution) (Figure 5A) and 17 of 40 KEGG pathways (Figures 4C and 5B).

The average $\log_{10}$ (q value) in both cohorts served as an indicator of overlap strength. Shared pathways included those for circadian entrainment, MAPK signaling, ECM–receptor interaction, and focal adhesion. When we looked for overlap between the DML genes in the New Zealand cohort and gout genes in the GWAS, we found that 15 KEGG pathways were enriched, and 5 were shared with the San Diego cohort (Figures 4D and 5C and Supplementary Figure 5A, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract). Strikingly, overlap between the San Diego and New Zealand gout cohorts was seen for 37 (52.9%) of the 70 identified TFs with motifs enriched in hypomethylated sequences (Figure 4E), and for 30 (78.9%) of the 38 enriched KEGG pathways based on TFs (Figures 4F and 5D and Supplementary Figure 5B [http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract]). The strongest overlap was for the osteoclast differentiation pathway (Figure 5D). Moreover,

differential methylome changes in Th17 cell differentiation, Th1 and Th2 cell differentiation, and B cell receptor signaling pathways were supported by the overlap findings (Figure 5D).

Integration of gout-related DMLs with TF-gene regulatory networks. Based on the findings from Taiji data analyses, *NFATC2* and *MEF2C* as well as 12 other TFs, including several involved in retinoid or E26 transformation-specific (ETS) transcriptional programming, exerted effects propagated throughout the entire leukocyte TF-gene regulatory network (Figure 6). The known biologic effects of these 14 TFs, which are pertinent to gout and/or gout comorbidities (as individually described in Supplementary Table 5, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract), included modulation of myeloid cell differentiation, inflammation, and immunity.

DISCUSSION

This study probed the gout PBMC DNA methylome, revealing that the DNA methylome signature in gout patients was clearly
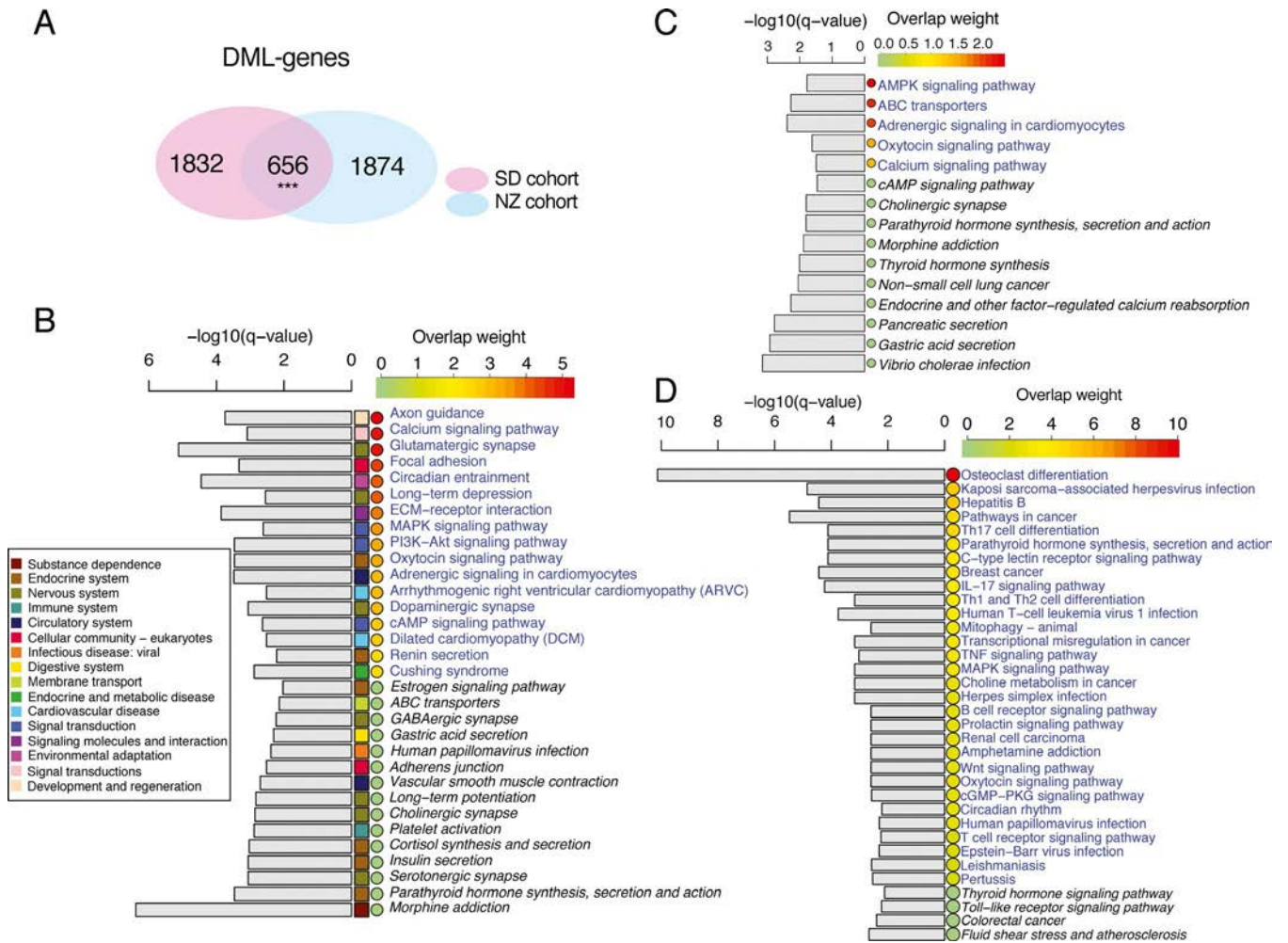
Figure 5. Concordance of differential DNA methylation findings between the San Diego (SD) and New Zealand (NZ) gout sample sets. **A**, Venn diagrams illustrating the differential methylation loci (DML) genes shared between the 2 cohorts. *** = $P \leq 5 \times 10^{-324}$ from hypergeometric distribution. **B**, Enriched KEGG pathways using DML genes in the New Zealand cohort. The *P* value for each pathway was obtained using the hypergeometric distribution and transformed into a q value. Colored squares indicate the subclass for each pathway. Pathways common to the 2 cohorts are indicated in blue, and those in italics are a New Zealand cohort–specific pathway. The overlap weight for the common pathways is the average of the log_{10} (q value) in the 2 cohorts. **C**, Enriched pathways identified from overlaps between DML genes in the New Zealand cohort and genome-wide association study genes. **D**, Enriched pathways in gout, identified from specific transcription factors by motif scanning in hypomethylated DML regions in the New Zealand cohort. ECM = extracellular matrix; PI3K = phosphatidylinositol 3-kinase; AMPK = AMP-activated protein kinase; IL-17 = interleukin-17; TNF = tumor necrosis factor; PKG = cGMP-dependent protein kinase.

separable from that in controls. In gout, epigenome-wide differential DNA methylation was highlighted by changes in gout genetic risk genes, signaling and transcriptional pathways in innate and adaptive immunity, neuroinflammation, and osteoclastogenesis. Many results from the San Diego cohort, particularly for differentially methylated TF motifs and related pathways, were validated in the New Zealand gout cohort. The study design limited potential confounding effects from current hyperuricemia, ongoing acute gouty arthritis, colchicine or glucocorticoid use, age, sex, and some comorbidities.

The gout DNA methylome signature could reflect the combined effects of inherited genetic variants (1,7–9) and past hyperuricemia (13), as well as adaptations to age, diet, lifestyle exposures, and comorbidity stressors (11,15,36). Other contributors likely include innate immunity training effects via past bouts of systemic stress triggered by repeated, self-limiting flares of marked, acute inflammatory arthritis (12). Significantly, the circulating monocyte DNA methylome in rheumatoid arthritis (RA), as identified in a small study of 33 patients with RA, was separable from that in 17 healthy controls and reflected differences in RA disease activity, demonstrated plasticity in response to therapy, and was associated with distinct markers of the inflammatory state in RA peripheral blood (37). Tumor necrosis factor (TNF) and interferons can induce some of the monocyte DNA methylome changes in vitro associated with high disease activity in circulating monocytes from patients with RA (37). Monocyte DNA methylome

Figure 6. Shared transcription factor (TF)–gene regulatory networks between the San Diego and New Zealand cohorts, determined from integration of common gout-related differential methylation loci (DML) genes. The TF-gene regulatory networks were constructed using Taiji data analysis methods, in B cells, T cells, CD14+ monocytes, CD4+ T helper cells, and CD8+ αβ T cells, and then combined into one network. After mapping of DML genes and TFs to the network, 224 DML genes and 1,231 edges remained. Five modules were identified using the AGMfit algorithm. Genes were assigned different colors based on module membership. Notably, olive green indicates presence of a gene in ≥2 modules. Supplementary Table 5 (http://onlinelibrary.wiley.com/doi/10.1002/art.41173/abstract) summarizes the properties of the 14 genes present in all modules (in red). The shape schemes of nodes are as follows: diamond = body-hypermethylation (hyper); ellipse = body-hypomethylation (hypo); triangle = promoter-hypermethylation; V-shape = promoter-hypomethylation; rounded rectangle = DML gene with multiple DMLs.

changes that occurred in conjunction with high disease activity in RA were pinpointed in certain ETS family TFs (37), and effects of 2 ETS TF-gene regulatory networks were propagated throughout the entire leukocyte TF-gene regulatory network in gout patients in this study.

We saw differential DNA methylation in several gout GWAS genes relevant to gouty arthritis, including *IL23R*, which modulates the bridging of innate and adaptive immunity, and promotes granuloma formation and cell invasiveness. *PRKAG2* encodes the γ2 chain of AMPK, a major nutrition biosensor that is a native suppressor of urate crystal–induced inflammation, partly by promoting autophagy and inhibiting NF-κB and NLRP3 inflammasome activation (31). AMPK activation also mediates the antiinflammatory effects of colchicine (31). Through insulin-like growth factor 1 receptor (IGF-1R) and PI3K/Akt signaling, mechanistic target of rapamycin (mTOR) signaling, autophagy and mitophagy, and NF-κB activation, AMPK signaling intersects

with the longevity regulation pathway, also a modulator of innate inflammation (33). AMPK signaling also intersects with circadian rhythm. Circadian entrainment modulates diurnal changes in phagocyte behavior (38), which could promote the nocturnal flare onset that is so characteristic in gout (39). The PI3K/Akt pathway regulates inflammatory monocyte/macrophage activation by downstream signaling via hypoxia-inducible factor 1, the product of a gout genetic risk locus, and via intersection with mTOR activity and AMPK in regulating antiinflammatory autophagy.

Gout DNA methylome changes were detected in multiple pathways fundamental in phagocyte migration and activation, including those for cytoskeletal organization and cell adhesion and migration. The gout DNA methylome also included changes in the pathway for TRP ion channel signaling, which acts in endothelial cells and leukocytes and mediates neuroinflammation, including in response to urate crystals in vivo (40).

Many interconnected TF-gene networks and signaling pathways had differential DNA methylation in gout, including osteoclastogenesis (3). Enhanced in vitro osteoclastogenesis of RANKL- and M-CSF–treated PBMCs from gout patients and joint fluid mononuclear leukocytes was reported in severe, tophaceous gout, being associated with increased circulating levels of RANKL and M-CSF (6). The origins and sites (i.e., in the bone marrow, peripheral blood, or synovial lining) of priming of the osteoclast precursors that develop into osteoclasts remain to be clarified in gout. Characteristically, bone erosion in patients with gout takes years to become evident on imaging (4). Our results suggest that epigenomic training potentially modulating bone erosion is identifiable even in relatively early gout, and even with reasonable control of serum urate levels.

The results of this study identified gout-associated enrichment of differential DNA methylation in pathways for B and T cell receptor signaling and for Th17 cell development and IL-17 signaling, as well as gout-specific enrichment of DMLs for *STAT2*, *IRF1*, *NFATC2*, and *NFATC3*. These findings collectively highlight changes in adaptive immunity programming in gout. Not only IL-1β and IL-17 (41), but also potentially *IRF1* and *STAT2* could help bridge innate and adaptive immunity in gout. For example, *IRF1* activates genes involved in both innate and acquired immune responses, and also promotes inflammatory M1 macrophage commitment (34). Moreover, *NFATC3* regulates polarization of inflammatory M1 macrophages (42). By modulating the effects of *IRF1*, *MYC*, and *STAT1* and the functions of inflammatory myeloid cells, and by suppressing NF-κB hyperactivation, *NFATC2* regulates both acute and chronic inflammation (43). Determining whether such methylome changes mediate the presence of B and T lymphocytes, and plasma cells in the corona surrounding urate crystals in tophi (5), is germane to elucidating the fundamental processes of tissue urate crystal deposition, which might be modulated by crystal-bound immunoglobulin.

This study identified differentially DNA methylated TF-gene regulatory networks in leukocytes. *NFATC2* and *MEF2C* stood out among the TFs propagated throughout the network; notably, both demonstrated gene-body hypomethylation, and both *NFATC2* (also termed *NFAT1* or *NFATp*) and *MEF2C* regulate multiple DML genes. Collective evidence supports a role of *NFATC2* and *MEF2C* in potentially acting in the functional regulation of cascades modulating gouty inflammation. *NFATC2* is a member of an inducible TF family with 5 members, with 4 of these regulated by Ca^{2+} signaling, which is itself a pathway differentially methylated in PBMCs in the San Diego gout cohort. NFATs transduce signaling after ligation of Toll-like receptor 4 (TLR-4), CD14, and dectin-1, to stimulate intracellular Ca^{2+} signaling in myeloid cells. Consequent calcineurin phosphatase activation is permissive for NFAT nuclear translocation, allowing NFATs to act with AP-1 and other TFs to regulate transcription (44). Calcineurin, partly by acting on *NFATC2*, limits NF-κB hyperactivation and signaling and

suppresses TNF expression in macrophages (44). Conversely, calcineurin inhibition promotes hyperactivation of NF-κB and tolerance to the TLR-4 ligand lipopolysaccharide (44). Notably, the calcineurin inhibitor cyclosporine is associated with not only hyperuricemia and gout, but also accelerated tophus development, and particularly severe acute and chronic gouty arthritis (45).

MEF2C is a checkpoint for precursor commitment to monocyte or neutrophil differentiation. In addition, *MEF2C*, via its effects on Jun, promotes monocyte differentiation over granulopoiesis (35). *MEF2C* also enhances activation-induced macrophage apoptosis; and limits NF-κB activation. Furthermore, *MEF2C* promotes the expression of CD14, which is essential for the inflammatory function of classic monocytes (46). CD14 rs2569190 is a gout genetic risk locus in European and New Zealand Polynesian populations (47), and CD14 rs2569190 can modulate transcriptional activity of the CD14 promoter (48). Mononuclear phagocyte CD14, which associates with TLR-2 and TLR-4, is directly engaged by urate crystals, and supports crystal-induced NLRP3 inflammasome activation and the release of IL-1β and CXCL1 in vitro (49). The extent of neutrophil recruitment in experimental gouty inflammation in mice in vivo is diminished by ~75% in global CD14–knockout mice (49). Thus, multiple recognized effects of both *NFATC2* and *MEF2C* could regulate inflammation in gouty arthritis.

A limitation of this study was the fact that the established San Diego cohort of patients with gout were mostly receiving allopurinol and had serum urate levels that were not significantly different from those in healthy controls. There also were differences between the cohorts in the source of the genetic material, with PBMCs in San Diego compared to whole blood in New Zealand. Moreover, we studied only a small number of subjects. Given these limitations, the results of this study are seminal, but differential DNA methylation in even a small cohort of gout patients emphasizes the potential utility of the DNA methylome in discriminating gout from non-gout states. Controlling for the severity of comorbidities was not built into the design of this study. Nevertheless, characterization of the gout DNA methylome brings up the question as to whether epigenetic training of immunity in gout can impact comorbidities such as obesity, type 2 diabetes, and coronary artery disease.

Another limitation is that we focused on unseparated PBMCs in this study. Despite our computational adjustments for cell type composition, future single-cell DNA epigenetic and transcriptomic analyses would be informative. We cannot exclude the possibility that antiinflammatory medications that have been previously used to control gout flares (11) may have contributed to the DNA methylome findings. However, the duration of such a form of epigenetic memory is unknown.

Finally, gout methylome changes might modulate not simply gouty arthritis, but also changes in urate levels (e.g., through *SLC2A9* or *IGF1R*), renal function, and metabolism (e.g., through *PRKAG2* or *A1CF*), which are comorbidities pertinent to gout (16).

In summary, we observed grossly altered DNA methylation in the PBMCs from patients with gout. The disease signature, observed in 2 distinct cohorts, included involvement of AMPK signaling and multiple intersecting pathways, circadian entrainment, neuroinflammatory TRP ion channel signaling, and osteoclast differentiation. Differential DNA methylation in genes involved in B and T cell signaling, Th17 and IL-17 signaling, and genetic loci (including *STAT2* and *IRF1*), in combination with IL-23 signaling, could bridge the processes of innate and adaptive immunity to promote tophaceous granuloma formation and bone erosion in gout. *NFATC2* and *MEF2C* stood out among the PBMC TF–gene regulatory networks potentially involved in gout. Conspicuous absence of the NLRP3 pathway was a notable finding from the DML enrichment analysis. The results suggest a model whereby the DNA methylome can regulate acute and chronic gouty inflammation through signal transduction pathways and transcriptional pathways, including those for leukocyte differentiation and function and for neuroinflammation (50–53). The gout DNA methylome potentially acts on immune, metabolic, and other elements (13–15,31,32) that regulate first-signal NLRP3 inflammasome priming, leading to regulation of phagocyte behavior following second-signal inflammasome activation by phagocytosed urate crystals (54). Future studies to elucidate other novel target genes and pathways in addition to those highlighted herein would help advance the recognition of gout biomarkers and the development of new treatment strategies in patients with gout.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Z. Wang, Zhao, Phipps-Green, Liu-Bryan, Ceponis, Boyle, J. Wang, Merriman, W. Wang, Terkeltaub.

Acquisition of data. Z. Wang, Zhao, Phipps-Green, Liu-Bryan, Boyle, J. Wang, Merriman, W. Wang, Terkeltaub.

Analysis and interpretation of data. Z. Wang, Zhao, Phipps-Green, Liu-Bryan, Ceponis, Boyle, J. Wang, Merriman, W. Wang, Terkeltaub.

REFERENCES

1. Dalbeth N, Merriman TR, Stamp LK. Gout. Lancet 2016;388: 2039–52.

2. Terkeltaub R. What makes gouty inflammation so variable? BMC Med 2017;15:158.

3. Towiwat P, Chhana A, Dalbeth N. The anatomical pathology of gout: a systematic literature review. BMC Musculoskelet Disord 2019;20:140.

4. Dalbeth N, Saag KG, Palmer WE, Choi HK, Hunt B, MacDonald PA, et al. Effects of febuxostat in early gout: a randomized, double-blind, placebo-controlled study. Arthritis Rheumatol 2017;69:2386–95.

5. Dalbeth N, Pool B, Gamble GD, Smith T, Callon KE, McQueen FM, et al. Cellular characterization of the gouty tophus: a quantitative analysis. Arthritis Rheum 2010;62:1549–56.

6. Dalbeth N, Smith T, Nicolson B, Clark B, Callon K, Naot D, et al. Enhanced osteoclastogenesis in patients with tophaceous gout: urate crystals promote osteoclast development through interactions with stromal cells. Arthritis Rheum 2008;58:1854–65.

7. Major TJ, Dalbeth N, Stahl EA, Merriman TR. An update on the genetics of hyperuricaemia and gout. Nat Rev Rheumatol 2018;14:341–53.

8. Köttgen A, Albrecht E, Teumer A, Vitart V, Krumsiek J, Hundertmark C, et al. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. Nat Genet 2013;45:145–54.

9. Matsuo H, Yamamoto K, Nakaoka H, Nakayama A, Sakiyama M, Chiba T, et al. Genome-wide association study of clinically defined gout identifies multiple risk loci and its association with clinical subtypes. Ann Rheum Dis 2016;75:652–9.

10. Berdasco M, Esteller M. Clinical epigenetics: seizing opportunities for translation. Nat Rev Genet 2019;20:109–27.

11. Whyte JM, Ellis JJ, Brown MA, Kenna TJ. Best practices in DNA methylation: lessons from inflammatory bowel disease, psoriasis and ankylosing spondylitis [review]. Arthritis Res Ther 2019;21:133.

12. Netea MG, Schlitzer A, Placek K, Joosten LA, Schultze JL. Innate and adaptive immune memory: an evolutionary continuum in the host's response to pathogens. Cell Host Microbe 2019;25:13–26.

13. Crişan TO, Cleophas MC, Novakovic B, Erler K, van de Veerdonk FL, Stunnenberg HG, et al. Uric acid priming in human monocytes is driven by the AKT-PRAS40 autophagy pathway. Proc Natl Acad Sci U S A 2017;114:5485–90.

14. Gore AV, Weinstein BM. DNA methylation in hematopoietic development and disease. Exp Hematol 2016;44:783–90.

15. Ling C, Rönn T. Epigenetics in human obesity and type 2 diabetes. Cell Metab 2019;29:1028–44.

16. Zhu Y, Pandya BJ, Choi HK. Comorbidities of gout and hyperuricemia in the US general population: NHANES 2007–2008. Am J Med 2012;125:679–87.

17. Li B, Chen X, Jiang Y, Yang Y, Zhong J, Zhou C, et al. CCL2 promoter hypomethylation is associated with gout risk in Chinese Han male population. Immunol Lett 2017;190:15–9.

18. Neogi T, Jansen TL, Dalbeth N, Fransen J, Schumacher HR, Berendsen D, et al. 2015 gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheumatol 2015;67:2557–68.

19. Krishnan M, Major TJ, Topless RK, Dewes O, Yu L, Thompson JM, et al. Discordant association of the CREBRF rs373863828 A allele with increased BMI and protection from type 2 diabetes in Māori and Pacific (Polynesian) people living in Aotearoa/New Zealand. Diabetologia 2018;61:1603–13.

20. Wallace SL, Robinson H, Masi AT, Decker JL, McCarty DJ, Yu TF. Preliminary criteria for the classification of the acute arthritis of primary gout. Arthritis Rheum 1977;20:895–900.

21. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. ChAMP: 450k Chip Analysis Methylation Pipeline. Bioinformatics 2014;30:428–30.

22. Zhou W, Triche TJ Jr, Laird PW, Shen H. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. Nucleic Acids Res 2018;46:e123.

23. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 2012;13:86.

24. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat Biotechnol 2013;31:142–7.

25. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.

26. Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res 2011;39 Suppl 1:D712–7.

27. Buske FA, Bodén M, Bauer DC, Bailey TL. Assigning roles to DNA regulatory motifs using comparative genomics. Bioinformatics 2010;26:860–6.

28. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res 2018;46:D1284.

29. Zhang K, Wang M, Zhao Y, Wang W. Taiji: system-level identification of key transcription factors reveals transcriptional waves in mouse embryonic development. Sci Adv 2019;5:eaav3262.

30. Shaw OM, Harper JL. Bradykinin receptor 2 extends inflammatory cell recruitment in a model of acute gouty arthritis. Biochem Biophys Res Commun 2011;416:266–9.

31. Wang Y, Viollet B, Terkeltaub R, Liu-Bryan R. AMP-activated protein kinase suppresses urate crystal-induced inflammation and transduces colchicine effects in macrophages. Ann Rheum Dis 2016;75:286–94.

32. Sanchez-Lopez E, Zhong Z, Stubelius A, Sweeney SR, Booshehri LM, Antonucci L, et al. Choline uptake and metabolism modulate macrophage IL-1β and IL-18 production. Cell Metab 2019;29:1350–62.

33. Zeng Y, Nie C, Min J, Liu X, Li M, Chen H, et al. Novel loci and pathways significantly associated with longevity. Sci Rep 2016;6:21243.

34. Chistiakov DA, Myasoedova VA, Revin VV, Orekhov AN, Bobryshev YV. The impact of interferon-regulatory factors to macrophage differentiation and polarization into M1 and M2. Immunobiology 2018;223:101–11.

35. Schüler A, Schwieger M, Engelmann A, Weber K, Horn S, Müller U, et al. The MADS transcription factor Mef2c is a pivotal modulator of myeloid cell fate. Blood 2008;111:4532–41.

36. Li E, Zhang Y. DNA methylation in mammals. Cold Spring Harb Perspect Biol 2014;6:a019133.

37. Rodríguez-Ubreva J, de la Calle-Fabregat C, Li T, Ciudad L, Ballestar ML, Català-Moll F, et al. Inflammatory cytokines shape a changing DNA methylome in monocytes mirroring disease activity in rheumatoid arthritis. Ann Rheum Dis 2019;78:1505–16.

38. Heipertz EL, Harper J, Lopez CA, Fikrig E, Hughes ME, Walker WE. Circadian rhythms influence the severity of sepsis in mice via a TLR2-dependent, leukocyte-intrinsic mechanism. J Immunol 2018;201:193–201.

39. Choi HK, Niu J, Neogi T, Chen CA, Chaisson C, Hunter D, et al. Nocturnal risk of gout attacks. Arthritis Rheumatol 2015;67:555–62.

40. Trevisan G, Hoffmeister C, Rossato MF, Oliveira SM, Silva MA, Ineu RP, et al. Transient receptor potential ankyrin 1 receptor stimulation by hydrogen peroxide is critical to trigger pain during monosodium urate-induced inflammation in rodents. Arthritis Rheum 2013;65:2984–95.

41. Raucci F, Iqbal AJ, Saviano A, Minosi P, Piccolo M, Irace C, et al. IL-17A neutralizing antibody regulates monosodium urate crystal-induced gouty inflammation. Pharmacol Res 2019;147:104351.

42. Hu L, He F, Huang M, Peng M, Zhou Z, Liu F, et al. NFATc3 deficiency reduces the classical activation of adipose tissue macrophages. J Mol Endocrinol 2018;61:79–89.

43. Chae CS, Kim GC, Park ES, Lee CG, Verma R, Cho HL, et al. NFAT1 regulates systemic autoimmunity through the modulation of a dendritic cell property. J Immunol 2017;199:3051–62.

44. Bendickova K, Tidu F, Fric J. Calcineurin-NFAT signalling in myeloid leucocytes: new prospects and pitfalls in immunosuppressive therapy. EMBO Mol Med 2017;9:990–9.

45. Stamp LK, Chapman PT. Gout and organ transplantation. Curr Rheumatol Rep 2012;14:165–72.

46. Zheng R, Wang X, Studzinski GP. 1,25-dihydroxyvitamin D_3 induces monocytic differentiation of human myeloid leukemia cells by regulating C/EBPβ expression through MEF2C. J Steroid Biochem Mol Biol 2015;148:132–7.

47. McKinney C, Stamp LK, Dalbeth N, Topless RK, Day RO, Kannangara DR, et al. Multiplicative interaction of functional inflammasome genetic variants in determining the risk of gout. Arthritis Res Ther 2015;17:288.

48. Mertens J, Bregadze R, Mansur A, Askar E, Bickeböller H, Ramadori G, et al. Functional impact of endotoxin receptor CD14 polymorphisms on transcriptional activity. J Mol Med (Berl) 2009;87:815–24.

49. Scott P, Ma H, Viriyakosol S, Terkeltaub R, Liu-Bryan R. Engagement of CD14 mediates the inflammatory potential of monosodium urate crystals. J Immunol 2006;177:6370–8.

50. Immler R, Simon SI, Sperandio M. Calcium signalling and related ion channels in neutrophil recruitment and function [review]. Eur J Clin Invest 2018;48 Suppl 2:e12964.

51. Hawkins PT, Stephens LR. PI3K signalling in inflammation. Biochim Biophys Acta 2015;1851:882–97.

52. Dierickx P, Van Laake LW, Geijsen N. Circadian clocks: from stem cells to tissue homeostasis and regeneration [review]. EMBO Rep 2018;19:18–28.

53. Kurotaki D, Sasaki H, Tamura T. Transcriptional control of monocyte and macrophage development. Int Immunol 2017;29:97–107.

54. Yang Q, Liu R, Yu Q, Bi Y, Liu G. Metabolic regulation of inflammasomes in inflammation. Immunology 2019;157:95–109.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Heritability of the Fibromyalgia Phenotype Varies by Age

Diptavo Dutta,[1] Chad M. Brummett,[2] Stephanie E. Moser,[2] Lars G. Fritsche,[1] Alexander Tsodikov,[1] Seunggeun Lee,[1] Daniel J. Clauw,[2] and Laura J. Scott[1]

Objective. Many studies suggest a strong familial component to fibromyalgia (FM). However, those studies have nearly all been confined to individuals with primary FM, i.e., FM without any other accompanying disorder. The current 2011 and 2016 criteria for diagnosing FM construct a score using a combination of the number of painful body sites and the severity of somatic symptoms (FM score). This study was undertaken to estimate the genetic heritability of the FM score across sex and age groups to identify subgroups of individuals with greater heritability, which may help in the design of future genetic studies.

Methods. We collected data on 26,749 individuals of European ancestry undergoing elective surgery at the University of Michigan (Michigan Genomics Initiative study). We estimated the single-nucleotide polymorphism–based heritability of FM score by age and sex categories using genome-wide association study data and a linear mixed-effects model.

Results. Overall, the FM score had an estimated heritability of 13.9% (SE 2.9%) ($P = 1.6 \times 10^{-7}$). Estimated FM score heritability was highest in individuals ≤50 years of age (23.5%; SE 7.9%) ($P = 3.0 \times 10^{-4}$) and lowest in individuals >60 years of age (7.5%; SE 8.1%) ($P = 0.41$). These patterns remained the same when we analyzed FM as a case–control phenotype. Even though women had an ~30% higher average FM score than men across age categories, FM score heritability did not differ significantly by sex.

Conclusion. Younger individuals appear to have a much stronger genetic component to the FM score than older individuals. Older individuals may be more likely to have what was previously called "secondary FM." Regardless of the cause, these results have implications for future genetic studies of FM and associated conditions.

INTRODUCTION

Fibromyalgia (FM) is a symptom complex characterized by widespread pain accompanied by somatic symptoms such as fatigue and sleep and memory problems. Nearly all recent research studies of FM have focused on what is termed "primary FM," which is FM without any other identifiable autoimmune or structural causes of pain. However, similar symptom complexes are observed in individuals with identifiable causes of pain such as autoimmune disorders and chronic diseases. This form of FM is thought to be more similar to animal and human studies of central sensitization, where ongoing nociceptive input is required to drive the processes of central sensitization, at the level of both the spinal cord and the brain (1–6).

Individuals with primary FM typically begin developing pain in their childhood or teens and are often diagnosed as having regional pain conditions early in their life before finally being diagnosed as having FM. Primary FM occurs preferentially in women, is strongly familial, and coaggregates with other regional pain conditions in both individuals and families (7–12). In contrast to primary FM, pain with identifiable causes can be caused by pain-related diseases such as osteoarthritis, which often occur later in life, and less is known about the heritability of pain with identifiable causes. Understanding the pathogenic differences between FM with and FM without identifiable causes of pain may lead to different treatments for the 2 forms of FM. For example, central nervous system drugs and primary FM therapies may be less effective than identification and treatment of the ongoing nociceptive input (7,13).

Candidate gene and genome-wide association studies (GWAS) comparing variant allele frequencies in FM cases and controls have been performed, but many of the results have been inconsistently noted or replicated (14). However, the studies to date

have been small in size (n < 1,000) and would only have been able to identify common genetic variants with very large effects (14,15).

Population- (16) or hospital-based (17) cohorts provide an opportunity to study the genetics of a wider spectrum of pain in much larger samples, although it may not always be possible to differentiate between patients with and those without any identifiable causes of pain given available survey or medical records and the high prevalence of individuals with peripheral sources of ongoing nociceptive input. Large studies of quantitative pain phenotypes can potentially contribute to the understanding of disease processes. A multisite chronic pain (MCP) GWAS (18) performed using data from the large-scale UK Biobank found 76 independent MCP-associated variants at 39 loci and an estimated MCP heritability of 10.2%.

Estimates of disease or trait heritability are of interest because they give a sense of the genetic contribution to the measured trait. Heritability has traditionally been estimated from family and twin studies, which require intensive participant recruitment. But narrow-sense heritability (additive components) can also be estimated from cohort- or case–control–based GWAS data (19,20). The estimation of heritability from GWAS data is based on the idea that if genetics underlie the predisposition to disease, individuals with more similar levels of a trait or with a disease will tend to share more alleles than individuals with less similar trait levels or without the disorder (21). For example, estimates of heritability based on GWAS of nonfamilial data range from 55% to 81% for height (19,22), 23% to 51% for body mass index (BMI) (20,23), and 37% to 50% for depression (24,25). The GWAS-based estimates are usually smaller than those estimated from familial data or twin studies, likely because they only capture additive effects from the variant classes included in the estimation (narrow-sense heritability) (26).

The genetic contributions to trait level or disease risk can vary by age or sex (27), and these differences can be assessed by estimating heritability in subgroups of samples. For example, many physical measures, including basal metabolic rate, systolic blood pressure, BMI, and neck pain, appear to be more heritable in younger individuals (28), potentially because trait differences at younger ages are less driven by environmental factors or because the processes that cause trait differences at older ages are different from or more diverse than those at younger ages.

This study was undertaken to estimate the heritability of FM and of a continuous measure of FM severity and to investigate whether heritability differs by patient sex and/or age at assessment. To do this, we measured FM severity using patient-completed 2011 Survey Criteria for FM (29,30) from 26,749 individuals of European ancestry undergoing elective surgery in the Michigan Genomics Initiative. Using GWAS data, we estimated the heritability of a continuous phenotype, FM severity (FM score), across age categories. We also dichotomized FM scores as a case–control phenotype according to 2 different definitions and estimated their heritability across age categories. Further, we estimated the genetic correlation of FM score with several psychiatric,

personality, and autoimmune traits using publicly available GWAS summary statistics.

PATIENTS AND METHODS

Patients. Participants were prospectively recruited into the Michigan Genomics Initiative, an institutional biorepository at the University of Michigan. All patients were ≥18 years of age and were scheduled to have an elective surgery on the day of their recruitment. We excluded patients who did not speak English, were unable to provide written informed consent, or were currently imprisoned. We obtained written informed consent from all patients for use of their clinical data and DNA for research purposes. This study was approved by the University of Michigan Institutional Review Board (IRB ID HUM00099605).

Genotyping. We genotyped DNA from blood samples using customized versions of an Illumina HumanCoreExome v12.1 array and applied quality control filters (see Supplementary text, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary. wiley.com/doi/10.1002/art.41171/abstract). After sample quality control, 37,412 samples remained, with 462,868 polymorphic variants. The average genotyping rate for the samples and variants included was 99.96%.

We projected the genotype data for the 37,412 samples on the principal components of Human Genome Diversity Project data using TRACE (31) to infer the genetic ancestry of the samples. Of these samples, 31,730 (84.8%) were inferred to be from individuals of European ancestry. We estimated the sample kinship (32) and retained 30,431 samples from participants who had less than a second-degree relationship. We next performed principal components analysis on their genotype data. We excluded 33 samples that were outliers based on the first and second principal components, resulting in 30,398 samples with genotype data (Supplementary text and Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41171/abstract).

Phenotyping. We phenotyped patients preoperatively using a self-administered questionnaire on widespread pain and psychological status, based on the American College of Rheumatology (ACR) Survey Criteria for FM (30). These criteria for FM, conceptualized in 2011, represent a validated self-report measure based on the presence of widespread pain and comorbid symptoms (29,30). To calculate the Widespread Pain Index (WPI), we assessed 19 specific body areas using the Michigan Body Map (range of possible scores 0–19) as described in the ACR Survey Criteria (30). To calculate a Symptom Severity Index (SSI) we used the comorbid Symptom Severity Scale with questions on fatigue, trouble thinking or remembering, waking up tired, pain or cramps in the lower abdomen, depression, and headache (range of possible scores 0–12). Following the method of Wolfe et al (30), we

summed the WPI and SSI to create an FM score (range of possible scores 0–31).

To dichotomize patients as FM cases or controls based on the FM score and its components, we used the following 2 definitions: 1) any individual with an FM score of ≥13 was defined as a case according to the 2011 criteria for FM; and 2) any individual with pain in 4 of the 5 main body regions, a WPI score of ≥7, and an SSI score of ≥5, or with pain in 4 of the 5 main body regions, a WPI score of 4–6, and an SSI score of ≥9 was defined as a case according to the modified 2016 criteria for FM. The modified 2016 criteria is a modified version of the criteria outlined by Wolfe et al (33) in 2016, adapted to our study according to the availability of data. (For details, see Supplementary text, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41171/abstract.) All individuals who were not classified as cases according to the definitions described above were classified as controls. Of the 30,398 individuals of European ancestry with genotype data, 3,649 (12.0%) did not have an FM score (WPI was missing for 2,708 [8.9%] and SSI was missing for 3,494 [11.5%]), leaving 26,749 individuals for analysis.

Log transformation of FM score, WPI, and SSI. Since the distribution of the FM score was highly skewed, we added a small constant (0.1) to the FM score to retain individuals with a value of 0 in the analysis, and log transformed the adjusted score. We regressed the log-transformed FM score (log FM) on age, age squared, and sex using a linear regression model and inverse normalized the residuals (inverse normalized FM score). The same transformations were applied to the WPI and SSI.

Estimation of heritability and genetic correlation. The genetic contribution to a phenotype can be measured by heritability, the fraction of trait variation explained by genetic variation. We estimated the heritability of the FM score, WPI, and SSI using a linear mixed-effects model (34) defined as:

$$y = X\beta_y + g_y + \varepsilon_y$$

where y is the vector of phenotype values for n individuals, in this case the inverse normalized FM score (or inverse normalized WPI/SSI score); X is a matrix of nongenetic covariates containing the top 10 genetic principal components and a binary variable with levels 0 and 1 indicating the genotype array (UM_HUNT_Biobank or UM_HUNT_Biobank_v1-1); β_y is the vector of corresponding fixed effects; g_y is a random vector of genetic contributions to the phenotypes, random effects, with $g_y \sim N(0, \sigma_y^2 K)$, where K is the genetic relatedness matrix (GRM) between the pairs of individuals (see Supplementary text for the construction of the GRM) and σ_y^2 is the genetic variability contributed by the genetic relatedness of the samples; and $\varepsilon_y \sim N(0, \sigma_{ey}^2 I)$, where σ_{ey}^2 is the residual variance of the model.

The heritability of the phenotype y is estimated as $h^2 = \sigma_y^2 / (\sigma_y^2 + \sigma_{ey}^2)$. We used genome-wide complex trait analysis (35) to fit this model and estimate the heritabilities. We used a likelihood ratio test to evaluate the significance of the estimated heritability. We note that h^2 used here is a narrow-sense heritability measure. Thus, h^2 only measures the fraction of trait variability explained by the additive effects of the variants in the array, which is lower than the total trait heritability. For a case–control phenotype (binary), we used the same linear mixed-effects model but with a liability scale, adjusting for ascertainment probabilities of the cases by the population prevalence (36). We used 2 different ways to dichotomize individuals into cases and controls based on FM score. For both case–control definitions we used a population prevalence of 2% to estimate the heritability of FM as a case–control phenotype (37).

The genetic overlap between 2 phenotypes y and w can be measured by the genetic correlation (20). We estimated the genetic correlation among the FM score, SSI, and WPI using a multiple linear mixed-effects model. For phenotype w, we used the same linear mixed-effects model as for y:

$$w = X\beta_w + g_w + \varepsilon_w; \ g_w \sim N(0, \sigma_w^2 K); \ \varepsilon_w \sim N(0, \sigma_{ew}^2 I)$$

with the additional assumption $cov(g_y, g_w) = \sigma_{yw} K$, where σ_{yw} is defined as the coheritability of the phenotypes y and w. The genetic correlation between the phenotypes is then defined as $r_{yw} = \sigma_{yw} / \sigma_y \sigma_w$. We used Phenix (38) to fit the model and calculate the coheritability and subsequently estimated the genetic correlations. Further, we estimated the genetic correlation of FM score with selected traits using publicly available summary statistics from existing genetic association studies using linkage disequilibrium score regression (39–41).

RESULTS

Our analysis included 26,749 patients of European ancestry who were scheduled for elective surgery. Their mean ± SD age was 54.2 ± 15.9 years, and 53.2% were women. The distribution of the FM scores for the 26,749 individuals is shown in Figure 1. For 10.8% of the samples, the FM score was 0 (both the WPI and SSI scores were 0). (See Supplementary Figure 2, available on the
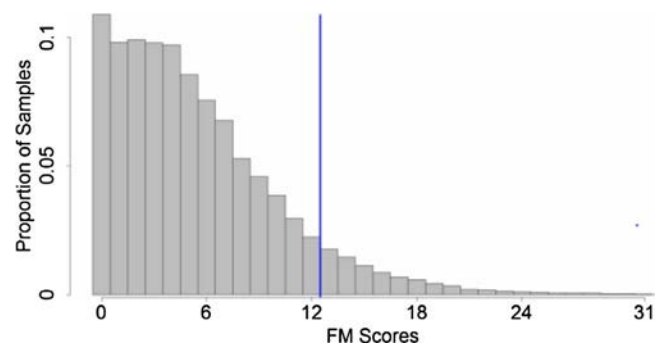
Figure 1. Distribution of fibromyalgia (FM) scores in the sample included for analysis. The range of possible scores is 0–31. Vertical line indicates the cutoff for considering an individual an FM patient (FM score ≥13). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41171/abstract.

Arthritis & Rheumatology web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41171/abstract, for WPI and SSI distributions.)

FM scores by age categories and sex are shown in Figure 2. When divided into age subcategories of 10 years, individuals who were 40–50 years old and those who were 50–60 years old had significantly higher FM scores than those who were ≤40 years old. Individuals who were 60–70 years old, those who were 70–80 years old, and those who were 80–90 years old had lower FM scores than those who were ≤40 years old (Figure 2). This pattern was consistent for women and men (Figure 2) (for regression estimates see Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41171/abstract). In the total sample, the mean ± SD FM score was higher in women than in men (6.2 ± 4.9 versus 4.6 ± 4.2, respectively; mean score in women/mean score in men = 1.34). (See Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41171/abstract, for the FM scores for patients dichotomized into 2 groups by age [≤50 years or >50 years].)

We investigated whether the components of the FM score, measurements of pain at different sites of the body (WPI), and symptoms of pain (SSI) had consistent trends across age and sex (Supplementary Table 3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41171/abstract). Younger individuals (age ≤50 years) had a lower WPI score than older individuals (age >50 years) (mean ± SD 1.8 ± 2.6 versus 2.0 ± 2.6, respectively). In contrast, younger individuals had a higher SSI score than older individuals (mean ± SD 3.8 ± 2.9 versus 3.4 ± 2.8) (Supplementary Table 3). Women had higher WPI and SSI scores than men (2.1 ± 2.8 versus 1.6 ± 2.3 for WPI and 4.1 ± 2.9 versus 3.0 ± 2.6 for SSI).

To evaluate the significance of the effects of age and sex on FM score–related measures, we used a multiple linear regression model simultaneously adjusted for age, age squared, and sex. We found that age, age squared, and sex were significantly associated with the FM score, WPI, and SSI ($P < 0.05$) (Supplementary Table 4, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41171/abstract). To determine if the effect of age category (younger or older) on FM score varied by sex or if the effect of sex on FM score varied by age category (younger or older), we tested for an interaction between age category and sex. We did not find significant
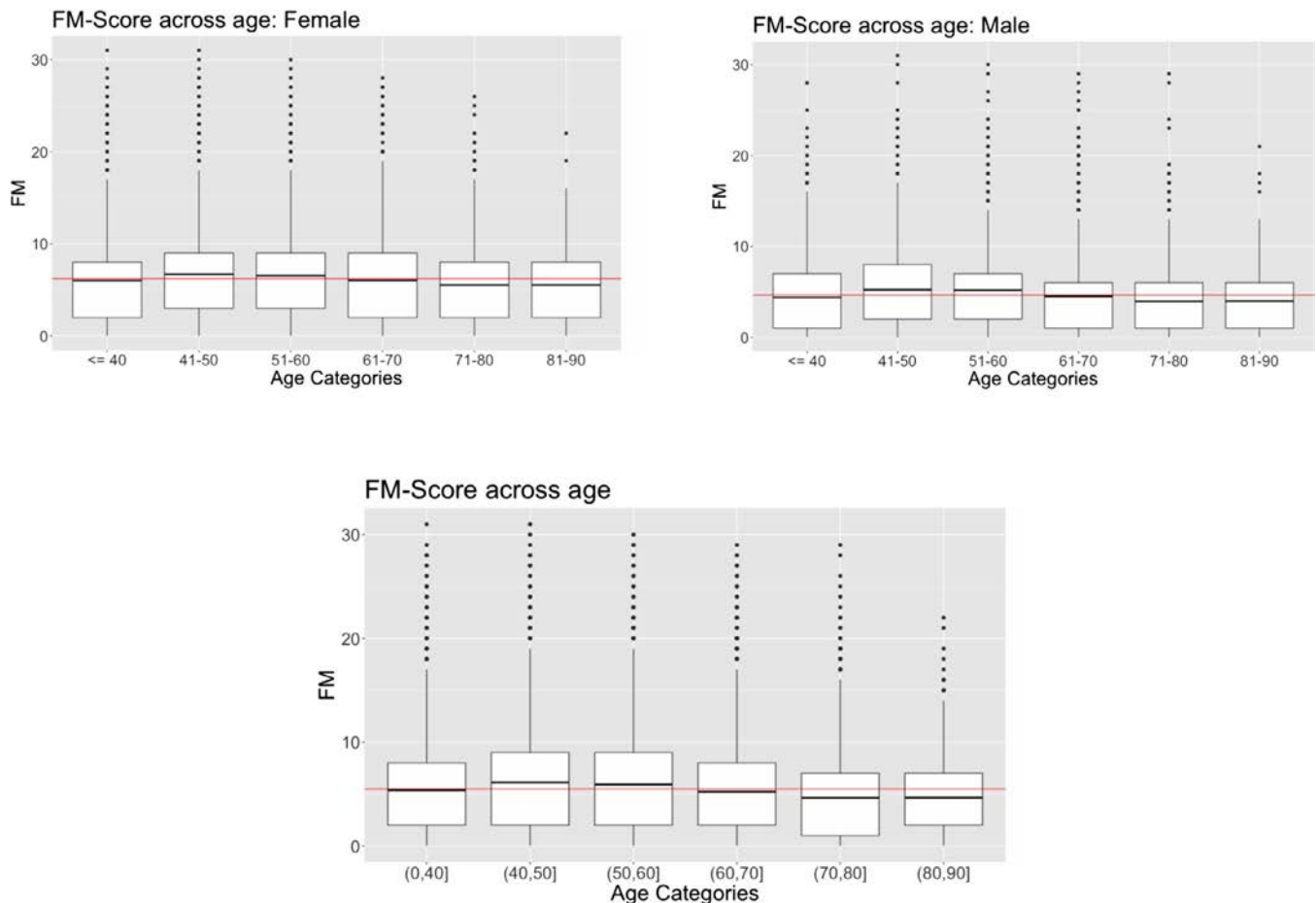
Figure 2. Distribution of fibromyalgia (FM) scores in women, men, and the entire sample, divided into 10-year age categories. Data are shown as box plots, where the boxes represent the 25th to 75th percentiles, the lines within the boxes represent the median, and the lines outside the boxes represent the 10th and 90th percentiles. Circles indicate outliers. Red lines indicate the mean FM score across all age categories.

evidence that the effect of sex varied between older and younger individuals or that the effect of age varied by sex.

To estimate the genetic contribution to FM score, we calculated the genotype-based heritability of the inverse normalized FM score. We constructed the GRM using the common variants (minor allele frequency >5%) and fit a linear mixed-effects model (see Patients and Methods). The estimated heritability of FM was 13.9% (SE 2.9%) ($P = 1.6 \times 10^{-7}$). To examine if heritability differed by age, we divided the sample into 10-year age categories. Patients ≤40 years old had the highest heritability, estimated at 22.8% (SE 13.4%), and those 60–70 years old had the lowest heritability, estimated at 3%, although no age category was significantly heritable on its own. We saw similar trends for WPI and SSI (Figure 3).

To obtain larger age subgroups, we dichotomized patients by age cutoffs and estimated the heritability in the corresponding age groups (Table 1). For each age cutoff we observed that younger individuals consistently had higher heritability of FM than older individuals. For example, individuals age ≤50 years had an estimated heritability of 23.5% (SE 7.9%) ($P = 3.0 \times 10^{-4}$) and those age >50 years had an estimated heritability of 8.6% (SE 5.6%) ($P = 0.12$). This means that the estimated heritability of FM for individuals ≤50 years old is significantly higher than 0. Conversely, the heritability for individuals >50 years old is low and could not be distinguished from 0 in this sample.

When we repeated the analysis by age category for men and women separately, we found that women had slightly higher estimated heritabilities than men in almost all age categories. However, we found no evidence of a significant difference in the estimated heritabilities between men and women (Supplementary text and Supplementary Table 5, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10. 1002/art.41171/abstract).

To assess the heritability in individuals more likely to have FM we used the 2 definitions of FM cases described above. Any individual with an FM score of ≥13 was defined as a case according to the 2011 FM criteria (22) (n = 2,304; sample prevalence 8.6%).
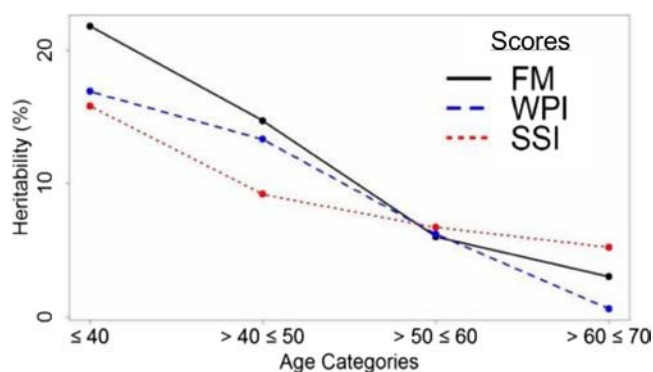
Table 1. Heritability of inverse normalized fibromyalgia score by age

Age, years	n	Heritability estimate, %	SE	P
≤40	5,693	22.8	13.4	0.09
>40	21,056	8.1	4.9	0.06
≤50	10,201	23.5	7.9	3.0×10^{-4}
>50	16,548	8.6	5.6	0.12
≤60	16,687	12.4	5.5	0.01
>60	10,062	7.5	8.1	0.41
All	26,749	13.9	2.9	1.6×10^{-7}

Any individual with pain in 4 of the 5 main body regions, a WPI score of ≥7, and an SSI score of ≥5, or with pain in 4 of the 5 main body regions, a WPI score of 4–6, and an SSI score of ≥9 was defined as a case according to the modified 2016 criteria (adapted from Wolfe et al [33]) (n = 1,319; sample prevalence 4.9%). All individuals not defined as a case were defined as controls. All but 48 of the participants who met the modified 2016 criteria also met the 2011 criteria. We estimated heritabilities as 8.6% ($P = 0.005$) for those who met the 2011 criteria and 7.9% ($P = 0.41$) for those who met the modified 2016 criteria. When participants were divided into age categories, we observed higher estimated heritability for younger individuals than for older individuals, suggesting that the overall trends by age were consistent in the data across varying levels of FM severity measured by different criteria (Supplementary Table 6, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10. 1002/art.41171/abstract).

To understand the contributions of the WPI and SSI to the age-based heritability trends for FM score, we estimated the genetic correlations of FM score with WPI and SSI by 10-year age categories (Figure 4). The estimated genetic correlation between the FM score and WPI score varied from 38% for younger individuals (age ≤40 years), to 56% for individuals ages >40 and ≤50 years and those ages >50 and ≤60 years and 30% for the group of older individuals (ages >60 and ≤70 years). The genetic correlation between the FM score and SSI varied from 57% in individuals age ≤40 years to 88% in individuals ages >60 years and ≤70 years. The estimated genetic correlation between WPI and SSI varied between 55% for individuals ages >40 years and ≤50 years and 6% for older individuals (ages >60 years and ≤70 years). These genetic correlation estimates show that for younger individuals, both WPI and SSI contribute substantially toward the genetic components of FM, while in older individuals SSI appears to be the dominant component. Further, our estimates show that, for younger individuals, WPI and SSI have a substantial shared genetic component, while in older individuals, they have a low genetic correlation. Using the definitions based on the 2011 criteria for FM and the modified 2016 criteria for FM to dichotomize individuals as cases and controls, we found similar patterns of genetic correlations across age categories (Supplementary Figure 3,

Figure 3. Heritability of fibromyalgia (FM) score, Widespread Pain Index (WPI), and Symptom Severity Index (SSI) scores by 10-year age category.
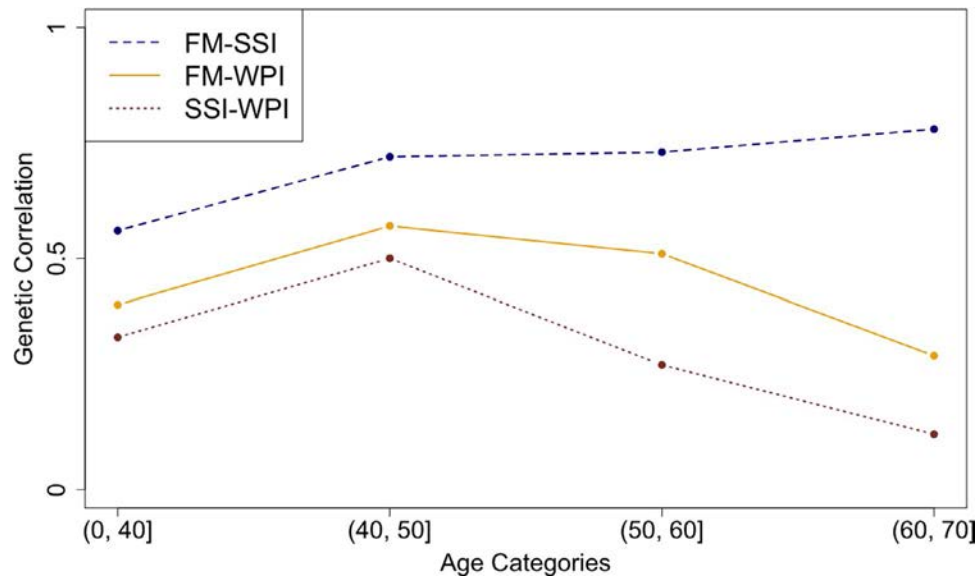
Figure 4. Estimated genetic correlation between the fibromyalgia (FM) score and Symptom Severity Index (SSI), between the FM score and Widespread Pain Index (WPI), and between SSI and WPI, by 10-year age categories.

available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.41171/abstract).

We next tested for coheritability of FM score with traits that might a priori be expected to be correlated with FM score or were found to be significant in the UK Biobank study of MCP. FM score had a significant genetic correlation with psychiatric disorders such as attention deficit hyperactivity disorder (ADHD), neuroticism, major depressive symptoms, subjective well-being, and depressive symptoms (–0.26 to 0.78) (Table 2). We also found significant

genetic correlation of FM score with immune and autoimmune diseases such as asthma and rheumatoid arthritis (RA) (0.31 to 0.35) (Table 2). To understand the contribution of SSI and WPI to the genetic correlations of FM score, we estimated their genetic correlations with the same traits separately. For almost all traits tested, SSI had similar estimated genetic correlations as FM score. WPI had similar estimated genetic correlations as FM score for asthma and RA, and of the traits tested, these were the most strongly genetically correlated with WPI. In contrast, WPI showed much

Table 2. Genetic correlation of FM score, WPI, and SSI with selected traits for the overall sample and by age*

Trait	PubMed ID	All			Age ≤50 years			Age >50 years			Genetic correlation in UK Biobank (MCP)
		FM	WPI	SSI	FM	WPI	SSI	FM	WPI	SSI	
Depressive symptoms†	27089181	0.53‡	0.11	0.49‡	0.63‡	0.12	0.59‡	0.49‡	0.07	0.44‡	0.59‡
Major depressive disorder†	29700475	0.40‡	0.08	0.43‡	0.48‡	0.07	0.50‡	0.36‡	0.09	0.37‡	0.53‡
Bipolar disorder†	21926972	0.08	0.03	0.09	0.13	0.04	0.14‡	0.06	0.04	0.07	0.02
ADHD†	27663945	0.78‡	0.19‡	0.66‡	0.81‡	0.20‡	0.80‡	0.53‡	0.15	0.54‡	NR
Subjective well-being†	27089181	−0.26‡	−0.12	−0.28‡	−0.25‡	−0.11	−0.23‡	−0.21‡	−0.14	−0.22‡	NR
PGC cross-disorder analysis†	24353885	0.28‡	0.15	0.33‡	0.32‡	0.13	0.34‡	0.25‡	0.11	0.30‡	0.13‡
Autism spectrum disorder†	30804558	0.04	−0.01	0.05	−0.01	−0.02	0.04	0.05	0.01	0.07	−0.10‡
Anorexia nervosa†	24514567	−0.09	−0.06	−0.12	−0.13	−0.07	−0.17‡	−0.07	−0.06	−0.10	−0.06‡
RA§	24390342	0.35‡	0.38‡	0.26‡	0.38‡	0.39‡	0.25‡	0.31‡	0.32‡	0.19‡	0.16‡
Asthma§	17611496	0.31‡	0.33‡	0.30‡	0.40‡	0.45‡	0.39‡	0.28‡	0.30‡	0.22	0.22‡
Primary biliary cirrhosis§	26394269	0.13	0.16	0.12	0.15	0.16	0.09	0.11	0.13	0.06	0.10‡
Neuroticism¶	27089181	0.39‡	0.16	0.35‡	0.45‡	0.17‡	0.41‡	0.37‡	0.13	0.38‡	0.40‡
Sleep duration#	27494321	−0.03	0.05	−0.11	−0.02	0.01	−0.08	−0.03	0.02	−0.08	NR
MCP	31194737	0.46‡	0.38‡	0.29‡	0.49‡	0.43‡	0.31‡	0.41‡	0.37‡	0.21‡	–

* FM = fibromyalgia; WPI = Widespread Pain Index; SSI = Symptom Severity Index; MCP = multisite chronic pain; ADHD = attention deficit hyperactivity disorder; NR = not reported; PGC = Psychiatric Genomic Consortium; RA = rheumatoid arthritis.
† Psychiatric trait.
‡ Significant estimate (*P* < 0.05).
§ Immune/autoimmune trait.
¶ Personality trait.
# Sleeping trait.

lower genetic correlation with psychiatric traits (–0.12 to 0.19) than we found for FM score or SSI, although genetic correlation of WPI with ADHD and neuroticism was nominally significant in younger individuals. Multiple phenotypes that were reported to have significant genetic correlation with MCP in the UK Biobank study (18) did not have significant genetic correlation with FM score, WPI, or SSI. However, the directions of genetic correlations for these traits were highly consistent between the UK Biobank and our sample (9 of the 10 reported traits in Table 2 had the same direction of effects). Further, the estimated genetic correlations for these traits were highly correlated with those estimated in the UK Biobank sample for MCP (r > 0.9).

We also estimated the genetic correlation of FM score, WPI, and SSI with the UK Biobank MCP. We found significant genetic correlations of MCP with FM score (0.46), WPI (0.38), and SSI (0.29). For younger individuals the genetic correlations were slightly higher than those estimated for older individuals (Table 2).

DISCUSSION

The present study is the largest to date to examine genetic contributions to the FM score (a composite measure of WPI and SSI). We found that the FM score was more heritable in younger individuals than in older individuals in this hospital-based sample. Thus, the variability in FM score for younger individuals, who are potentially more likely to have primary FM, appears to be driven by genetic factors shared across individuals to a greater degree than the variability in FM score in older individuals. Older individuals may have a greater contribution of environmental factors to pain, a greater diversity of conditions that increase pain, and/or more susceptibility towards nociceptive pain.

We found that there was a substantial genetic correlation of FM with both WPI and SSI for younger individuals, indicating that both the pain component (WPI) and the comorbid symptoms component (SSI) jointly contributed to the genetic architecture of FM in the younger individuals. In contrast, for the older individuals, the heritability of FM was more highly correlated with the comorbidity component (SSI) than with the pain component (WPI). Overall, our results suggest that genetic studies of FM might have differing results depending on the age of the participants.

If FM in younger individuals stayed constant throughout their lives, one would expect FM measures to slowly increase with age as pain from chronic diseases increases, but in this study, the mean FM score (in individuals undergoing elective surgery) was slightly lower in older individuals. Other studies have shown that the incidence and prevalence of FM wanes over time. Wolfe et al (37) showed that the prevalence of chronic widespread pain in Kansas peaked at ages 60–69 and then decreased in older individuals. Vincent et al (42) used the 2011 FM Survey Criteria to show that the prevalence of FM in the general Minnesota population was 8.4% for individuals ages 21–39 years, 6.0% for indi-

viduals ages 40–59 years, and 3.8% for individuals >60 years of age (27).

Given that an individual's genetic information is constant from birth until death (notwithstanding epigenetic modifications), our results suggest that, for a set number of FM cases, inclusion of younger individuals might increase the power to detect primary FM. All previous GWAS or large candidate gene studies in FM either included many individuals >50 years of age or did not report age (15,21,43). Thus, all of the large-scale genetic studies performed to date in FM have included sizable numbers of older individuals, where the genetic contributions to FM or FM symptoms might have been lower or different than in younger cohorts.

Differentiating individuals with primary FM from those with pain from an identifiable source, in a hospital or electronic health record (EHR)–based cohort, is difficult, even with the potentially large arrays of phenotype data. Although EHR-based studies can reduce diagnosis/reporting misclassifications and recall bias compared to cohort studies, ongoing sources of nociceptive pain might still be present without a diagnosis and hence not identified in an EHR-based study. In particular, nociceptive pain might be relatively more common in the population we have considered in this study, which consists of patients undergoing elective surgery. How to distinguish between pain with and pain without identifiable causes in such an EHR-based study remains a largely unanswered question. This in turn impedes the ability of our study and other EHR-based studies to isolate the patients with primary FM. In their UK Biobank–based study of MCP, Johnston et al (18) did not report whether individuals had identifiable or nonidentifiable sources of pain. Given our current measures of FM, we cannot definitively say if the older individuals we classified as being FM cases in this study had primary FM or had pain from an identifiable source. Thus, we do not have data to evaluate whether primary FM has a smaller genetic component in older individuals than it does in younger individuals.

Although women had higher FM scores than men across age categories, we did not find evidence that heritability varies by sex. This is possible because women can have a higher mean FM score value than men, for example, because sex-related factors cause higher FM score values in women, but still have the same amount of FM score *variability* explained by genetic variation. We can interpret this as follows: although the average FM scores in women are higher than those in men, the genetic contributions to FM score variability did not differ significantly by sex in this sample size.

FM co-occurs with multiple diseases, suggesting there are shared genetic factors underlying these diseases. We found that FM score and SSI have a strong genetic correlation with several psychiatric and personality syndromes, indicating a substantial genetic overlap between them, potentially because psychological measures are part of the FM score. However, our genetic correlation findings for FM score are consistent with results from a UK Biobank study (n = 387,649) by Johnston et al (18) that found a

genetic overlap between MCP and several psychiatric conditions. FM score and WPI were more strongly positively genetically correlated with asthma and RA than was MCP from the UK Biobank. For RA this may be due to an enrichment of individuals with RA in our surgical patient population compared to the more general UK Biobank population.

One limitation of our study is that our sample is not population based. Individuals who were scheduled to have surgery were eligible for recruitment in the study and are more likely to suffer from pain and to be enriched for particular FM-related disorders. Additionally, we used quantile-based inverse normalization of the FM score, which can affect the power to detect heritability.

Overall, this study highlights the importance of considering the age distribution of individuals when designing a genetic association study of FM. These data support the notion that there are (at least) 2 different forms of FM: one that occurs in younger individuals and is strongly genetically driven and one that occurs in older individuals and can be driven by a variety of nongenetic factors and other conditions that cause pain.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Scott 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. Brummett, Clauw.
Acquisition of data. Dutta, Scott.
Analysis and interpretation of data. Dutta, Moser, Fritsche, Tsodikov, Lee, Scott.

REFERENCES

1. Schmidt-Wilcke T, Clauw DJ. Fibromyalgia: from pathophysiology to therapy. Nat Rev Rheumatol 2011;7:518–27.

2. Clauw DJ. Fibromyalgia: a clinical review. JAMA 2014;311: 1547–55.

3. Arnold LM, Hudson JI, Hess EV, Ware AE, Fritz DA, Auchenbach MB, et al. Family study of fibromyalgia. Arthritis Rheum 2004;50:944–52.

4. Hudson JI, Goldenberg DL, Pope HG Jr, Keck PE Jr, Schlesinger L. Comorbidity of fibromyalgia with medical and psychiatric disorders. Am J Med 1992;92:363–7.

5. Aaron LA, Buchwald D. A review of the evidence for overlap among unexplained clinical conditions. Ann Intern Med 2001;134:868–81.

6. Kato K, Sullivan PF, Evengård B, Pedersen NL. A population-based twin study of functional somatic syndromes. Psychol Med 2009;39:497–505.

7. Buskila D, Neumann L, Hazanov I, Carmi R. Familial aggregation in the fibromyalgia syndrome. Semin Arthritis Rheum 1996;26: 605–11.

8. Woolf CJ. The pathophysiology of peripheral neuropathic pain: abnormal peripheral input and abnormal central processing. Acta Neurochir Suppl (Wien) 1993;58:125–30.

9. Woolf CJ, Thompson SW. The induction and maintenance of central sensitization is dependent on N-methyl-D-aspartic acid receptor activation: implications for the treatment of post-injury pain hypersensitivity states. Pain 1991;44:293–9.

10. Sluka KA. Pain mechanisms involved in musculoskeletal disorders. J Orthop Sports Phys Ther 1996;24:240–54.

11. Sluka KA, Clauw DJ. Neurobiology of fibromyalgia and chronic widespread pain. Neuroscience 2016;338:114–29.

12. Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. Pain 2011;152 Suppl:S2–15.

13. Ablin JN, Buskila D. Update on the genetics of the fibromyalgia syndrome. Best Pract Res Clin Rheumatol 2015;29:20–8.

14. Lee YH, Choi SJ, Ji JD, Song GG. Candidate gene studies of fibromyalgia: a systematic review and meta-analysis. Rheumatol Int 2012;32:417–26.

15. Docampo E, Escaramís G, Gratacòs M, Villatoro S, Puig A, Kogevinas M, et al. Genome-wide analysis of single nucleotide polymorphisms and copy number variants in fibromyalgia suggest a role for the central nervous system. Pain 2014;155:1102–9.

16. Bahcall OG. UK Biobank—a new era in genomic medicine. Nat Rev Genet 2018;19:737.

17. Wolford BN, Willer CJ, Surakka I. Electronic health records: the next wave of complex disease genetics. Hum Mol Genet 2018;27:R14–21.

18. Johnston KJ, Adams MJ, Nicholl BI, Ward J, Strawbridge RJ, Ferguson A, et al. Genome-wide association study of multisite chronic pain in UK Biobank. PLoS Genet 2019;15:e1008164.

19. Zaitlen N, Pasaniuc B, Sankararaman S, Bhatia G, Zhang J, Gusev A, et al. Leveraging population admixture to characterize the heritability of complex traits. Nat Genet 2014;46:1356–62.

20. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nat Genet 2014;11:407–9.

21. Cohen H, Buskila D, Neumann L, Ebstein RP. Confirmation of an association between fibromyalgia and serotonin transporter promoter region (5-HTTLPR) polymorphism, and relationship to anxiety-related personality traits [letter]. Arthritis Rheum 2002;46:845–7.

22. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 2010;42:565–9.

23. Silventoinen K, Magnusson PK, Tynelius P, Kaprio J, Rasmussen F. Heritability of body size and muscle strength in young adulthood: a study of one million Swedish men. Genet Epidemiol 2008;32:341–9.

24. Lohoff FW. Overview of the genetics of major depressive disorder. Curr Psychiatry Rep 2010;12:539–46.

25. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry 2000;157: 1552–62.

26. Mayhew AJ, Meyre D. Assessing the heritability of complex traits in humans: methodological challenges and opportunities. Curr Genomics 2017;18:332–40.

27. Pan L, Ober C, Abney M. Heritability estimation of sex-specific effects on human quantitative traits. Genet Epidemiol 2007;31:338–47.

28. Ge T, Chen CY, Neale BM, Sabuncu MR, Smoller JW. Phenome-wide heritability analysis of the UK Biobank. PLoS Genet 2017; 13:e1006711.

29. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic

criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10.

30. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RS, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR preliminary diagnostic criteria for fibromyalgia. J Rheumatol 2011;38:1113–22.

31. Wang C, Zhan X, Liang L, Abecasis GR, Lin X. Improved ancestry estimation for both genotyping and sequencing data using projection procrustes analysis and genotype imputation. Am J Hum Genet 2015;96:926–37.

32. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics 2010;26:2867–73.

33. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RL, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–29.

34. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, et al. Variance component model to account for sample structure in genome-wide association studies. Nat Genet 2010;42: 348–54.

35. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82.

36. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. Am J Hum Genet 2011;88:294–305.

37. Wolfe F, Ross K, Anderson J, Russell IJ, Hebert L. The prevalence and characteristics of fibromyalgia in the general population. Arthritis Rheum 1995;38:19–28.

38. Dahl A, Iotchkova V, Baud A, Johansson Å, Gyllensten U, Soranzo N, et al. A multiple-phenotype imputation method for genetic studies. Nat Genet 2016;48:466–72.

39. Bulik-Sullivan B, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium, et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015;47:291–5.

40. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015;47:1236–41.

41. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics 2017;33:272–9.

42. Vincent A, Lahr BD, Wolfe F, Clauw DJ, Whipple MO, Oh TH, et al. Prevalence of fibromyalgia: a population-based study in Olmsted County, Minnesota, utilizing the Rochester Epidemiology Project. Arthritis Care Res (Hoboken) 2013;65:786–92.

43. Offenbaecher M, Bondy B, de Jonge S, Glatzeder K, Krüger M, Schoeps P, et al. Possible association of fibromyalgia with a polymorphism in the serotonin transporter gene regulatory region. Arthritis Rheum 1999;42:2482–8.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Transcutaneous Electrical Nerve Stimulation Reduces Movement-Evoked Pain and Fatigue: A Randomized, Controlled Trial

Dana L. Dailey,[1] Carol G. T. Vance,[2] Barbara A. Rakel,[2] M. Bridget Zimmerman,[2] Jennie Embree,[2] Ericka N. Merriwether,[3] [iD] Katharine M. Geasland,[2] Ruth Chimenti,[2] [iD] Jon M. Williams,[4] Meenakshi Golchha,[4] Leslie J. Crofford,[4] and Kathleen A. Sluka[2] [iD]

Objective. Fibromyalgia (FM) is characterized by pain and fatigue, particularly during physical activity. Transcutaneous electrical nerve stimulation (TENS) activates endogenous pain inhibitory mechanisms. This study was undertaken to investigate if using TENS during activity would improve movement-evoked pain and other patient-reported outcomes in women with FM.

Methods. Participants were randomly assigned to receive active TENS (n = 103), placebo TENS (n = 99), or no TENS (n = 99) and instructed to use it at home during activity 2 hours each day for 4 weeks. TENS was applied to the lumbar and cervicothoracic regions using a modulated frequency (2–125 Hz) at the highest tolerable intensity. Participants rated movement-evoked pain (primary outcome measure) and fatigue on an 11-point scale before and during application of TENS. The primary outcome measure and secondary patient-reported outcomes were assessed at baseline (time of randomization) and at 4 weeks.

Results. After 4 weeks, a greater reduction in movement-evoked pain was reported in the active TENS group versus the placebo TENS group (group mean difference –1.0 [95% confidence interval –1.8, –0.2]; $P = 0.008$) and versus the no TENS group (group mean difference –1.8 [95% confidence interval –2.6, –1.0]; $P < 0.0001$). A reduction in movement-evoked fatigue was also reported in the active TENS group versus the placebo TENS group (group mean difference –1.4 [95% confidence interval –2.4, –0.4]; $P = 0.001$) and versus the no TENS group (group mean difference –1.9 [95% confidence interval –2.9, –0.9]; $P = <0.0001$). A greater percentage of the patients in the active TENS group reported improvement on the global impression of change compared to the placebo TENS group (70% versus 31%; $P < 0.0001$) and the no TENS group (9%; $P < 0.0001$). There were no TENS-related serious adverse events, and <5% of participants experienced minor adverse events from TENS.

Conclusion. Among women who had FM and were on a stable medication regimen, 4 weeks of active TENS use compared to placebo TENS or no TENS resulted in a significant improvement in movement-evoked pain and other clinical outcomes. Further research is needed to examine effectiveness in a real-world setting to establish the clinical importance of these findings.

INTRODUCTION

Fibromyalgia (FM) is a complex condition characterized by widespread pain and fatigue. Pharmacologic interventions are only modestly effective for treating FM, with most individuals experiencing activity-limiting pain despite use of multiple drugs (1,2). It has become increasingly recognized that nonpharmacologic interventions should be considered as first-line treatments

for chronic pain (3–5) and as safe, low-cost treatments that can be added to pharmacologic approaches. While there is strong evidence that exercise is an effective treatment for FM (6,7), individuals report that movement-evoked pain limits activity participation (8,9). Use of nonpharmacologic approaches that reduce movement-evoked pain would theoretically increase activity participation, resulting in a perceived global improvement.

Transcutaneous electrical nerve stimulation (TENS) is a nonpharmacologic intervention that delivers electrical current through the skin for pain control. Animal studies show that TENS activates endogenous inhibitory mechanisms to reduce central excitability (10–14). In contrast, individuals with FM exhibit reduced endogenous inhibition and enhanced central excitability (15,16). Thus, based on the mechanism of action of TENS, it may be useful in individuals with FM.

Although TENS is effective for several pain conditions, recent systematic reviews have shown mixed results (17–20). Johnson and colleagues have noted limitations of the TENS trials currently described in the literature, such as inadequate sample size, limited outcome data, and moderate risk of bias (17,21). We have further suggested that variables not considered in TENS clinical trials also lead to equivocal results (22). Stimulation intensity must be strong, but comfortable or greater for TENS effectiveness (14,23,24). TENS works best for movement-evoked pain (24,25) and provides the greatest effects while the unit is on (22,26), yet prior studies have routinely measured resting pain or assess pain after treatment, when physiologic effects of TENS are no longer optimal (25,27–30). Furthermore, few studies have examined effects on other domains such as fatigue, quality of life, or function.

We designed a double-blind randomized controlled trial to examine the effects of TENS in women with FM, using a study protocol that does not entail weaknesses of prior studies. The primary aim was to test effectiveness of repeated TENS on movement-evoked pain in women with FM following random assignment to 3 groups: active TENS, placebo TENS, or no TENS. Secondary aims were to test the effects of TENS on fatigue, function, and other patient-reported outcomes. We hypothesized that TENS would reduce movement-evoked pain, resulting in perceived global improvement in women with FM.

PATIENTS AND METHODS

Study design and participants. The Fibromyalgia Activity Study with TENS (FAST) is a phase II randomized, double-blind, placebo-controlled, dual-site clinical trial conducted at the University of Iowa and Vanderbilt University Medical Center and approved by the institutional review boards of both universities. The study protocol has been previously described (31), and the study was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrollment.

We examined the effects of TENS home use in women with FM in a trial in which women were treated with TENS or placebo for 4 weeks, followed by a 4-week period during which all subjects received active TENS (Figure 1A). Participants were recruited from 2 sites using a variety of strategies (for details see Supplementary Methods, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract). Eligibility was verified at both visit 1 and visit 2. Inclusion criteria were as follows: female sex, age 18–70 years, FM according to the American College of Rheumatology 1990 criteria (32), on a stable medication regimen during the 4 weeks preceding the study, and projected to be on a stable treatment regimen for the next 2 months. Exclusion criteria included a pain level of <4 on a 10-point numerical rating scale (NRS) at the first and second visits, inability to walk 6 minutes without assistance, TENS use in the last 5 years, presence of a pacemaker, history of neuropathic or autoimmune disorder, history of spinal fusion or metal implants in the spine, allergy to adhesive or nickel, pregnancy, epilepsy, and/or a serious or unstable medical or psychiatric condition that would preclude participation (31). All participants continued current treatments prescribed by their health care provider and were asked not to change medications during the study. Current medications were recorded at each visit. Analgesic use before the second visit did not differ between groups (Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract).

Outcome measures. The primary outcome measure was movement-evoked pain and secondary outcome measures were resting pain, fatigue, function, disease impact, quality of life, fear of movement, and other psychological factors. These measures are described briefly below, with more detail in the published protocol (31). The effects of TENS on pain and fatigue were examined before and during TENS treatment on visits 2, 3, and 4. Patient-reported outcomes were examined before TENS treatment at these same visits.

Assessment of pain, fatigue, and physical function. Pain intensity at rest and during movement was measured with an 11-point NRS before and during TENS. Movement-evoked pain was measured during a 6-minute walk test, which measures the distance a person can walk in 6 minutes, and a 5-time sit-to-stand test, which measures how long it takes for a person to move from a seated position to standing 5 times. Pain intensity and interference were measured using the Brief Pain Inventory (BPI) (33). Fatigue at rest and during movement was measured with an 11-point NRS before and during the 6-minute walk test and 5-time sit-to-stand test and with the Multidimensional Assessment of Fatigue (MAF) (34). Physical function was assessed using the 6-minute walk test, the 5-time sit-to-stand test, physical activity for 1 week recorded via accelerometry (Supplementary Methods, http://onlinelibrary. wiley.com/doi/10.1002/art.41170/abstract), and the International Physical Activity Questionnaire (IPAQ) short form (35).
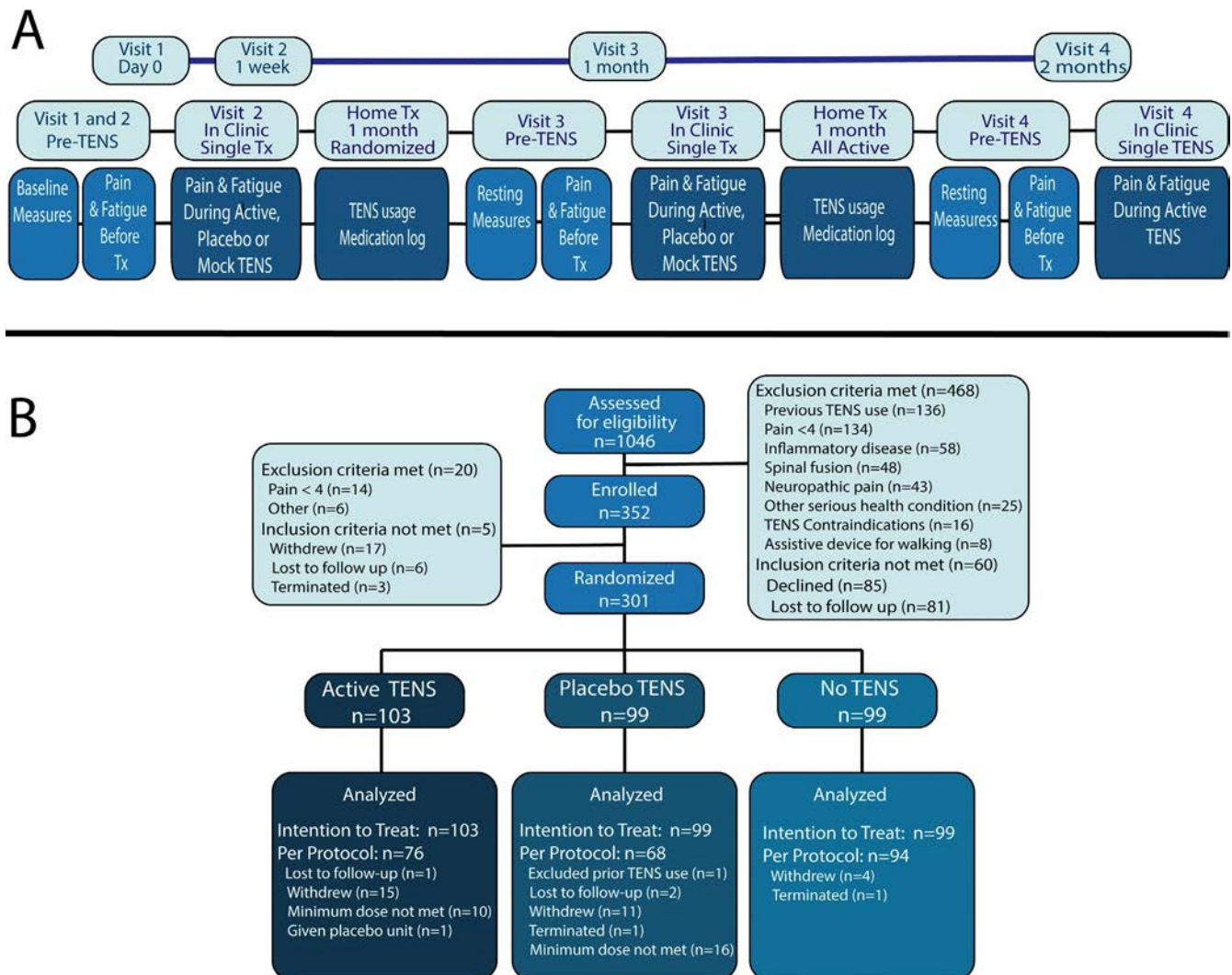
Figure 1. Study design and Consolidated Standards of Reporting Trials (CONSORT) diagram. **A,** Study design for all 4 visits. At visit 1, participants were screened for pain and fibromyalgia (FM) according to the American College of Rheumatology 1990 criteria (32). At visit 2, subjects were re-screened for pain and randomized into treatment (Tx) groups. Baseline questionnaires were assigned, and subjects were assessed for pain and fatigue at rest and during functional tasks. Transcutaneous electrical nerve stimulation (TENS) was applied during visit 2 and remained turned on for 30 minutes prior to reassessment of pain, fatigue, and function. Participants were given the TENS unit for home use over a 4-week period before returning for visit 3. Visit 3 followed the same protocol as visit 2. After visit 3, all subjects received active TENS for 4 weeks and were reassessed with the same protocol as visits 2 and 3. All assessments were the same across treatment arms. **B,** CONSORT diagram. We assessed 1,046 participants for eligibility, with 468 excluded prior to enrollment. The main reasons for exclusion were previous TENS use and a pain level of <4 on a 10-point numerical rating scale. Following enrollment at visit 1, 5 subjects did not meet the criteria for FM, and 1 week later at visit 2, 14 subjects were excluded for having a pain level of <4. After enrollment on visit 1, 17 subjects withdrew from the study due to personal reasons. Thus, the remaining 301 participants were then randomly assigned to receive active TENS (n = 103), placebo TENS (n = 99), or no TENS (n = 99). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract.

Patient-reported outcomes. We examined fear of movement with the Tampa Scale of Kinesiophobia (TSK) (36), pain catastrophizing with the Pain Catastrophizing Scale (PCS) (37), self-efficacy with the Pain Self-Efficacy Questionnaire (PSEQ) (38), and depression and anxiety with Patient-Reported Outcomes Measurement Information System (PROMIS) short forms. Disease impact was measured with the Revised Fibromyalgia Impact Questionnaire (FIQ) (39), and quality of life was assessed with the

Short Form 36 (SF-36) Health Survey (40). Use of rescue pain medication was examined using home logs for opioid and non-opioid analgesic use 1 week before visits 2, 3, and 4. Of the 301 participants enrolled in the study, completed logs were available for 227 patients' pain medication use, opioid and non-opioid, (76 in the active TENS group, 70 in the placebo TENS group, and 81 in the no TENS group). Perceived improvement was examined with the Global Impression of Change (GIC) using a 7-point scale.

Randomization, allocation, and blinding. Participants were randomly assigned to active TENS, placebo TENS, or no TENS groups with permuted blocks sizes of 6 and 9, stratified by site and opioid use status (Proc Plan; SAS/STAT software version 13.1). Subjects were classified as opioid users if they had taken an opioid at least 5 days per week for the last 30 days. The randomization schedule was password-protected, with access granted only to those who were not blinded with regard to the intervention (the statistician [MBZ] who generated the randomization schedule and the TENS allocators [CGTV, JMW] who provided the TENS intervention). Neither the statistician nor the TENS allocators had any role in patient recruitment, scheduling, or assessment of outcomes.

Assessments were performed by a separate person (outcome assessors [DLD, KMG, MG]) than the TENS allocators. Participants were blinded with regard to their treatment group (active TENS or placebo TENS), and outcome assessors were blinded with regard to all 3 groups. TENS allocators, who were not blinded with regard to treatment, were responsible for accessing the randomization schedule to assign participants to groups and for maintaining contact with participants between visits 2 and 3 (blinded phase). A mock TENS unit (a TENS unit with attached electrodes that provided no electric current intensity) was used in the no TENS group during visits 2 and 3 to blind outcome assessors. For all groups, a concealment pouch was used to prevent an outcome assessor from viewing the TENS unit, and participants were asked to not discuss treatment with outcome assessors. A standardized script for each treatment group specific to each visit was utilized so that all participants received the same instructions. The standardized script remained identical (except for 1 line that differed between the active TENS and placebo TENS groups, to reduce bias) (see Supplementary Methods, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary. wiley.com/doi/10.1002/art.41170/abstract). Blinding of outcome assessors was assessed after visit 3 by asking the assessors if the participant had received active TENS, placebo TENS, or no TENS (or if they did not know what treatment the participants received), and blinding of participants was determined by asking if they had received active TENS or placebo TENS, or if they did not know what treatment they received. Additional details on the integrity of blinding procedures performed in this study are available in Supplementary Methods.

TENS intervention. The EMPI Select TENS units (generously provided by DJO Global) delivered both active TENS and placebo TENS interventions via butterfly electrodes placed at the cervicothoracic junction and lower back (see Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.41170/abstract). Active TENS parameters were as follows: asymmetric biphasic waveform with a modulating frequency of 2–125 Hz, pulse duration 200 μsec, and the highest tolerable stimulation intensity. During visits 2, 3, and

4, TENS was applied by a TENS allocator in a clinical setting for 30 minutes prior to an outcome assessor measuring its effects on pain, fatigue, and function. The placebo TENS unit had an appearance identical to that of the active TENS unit and delivered current for 45 seconds, ramping down to 0 in the last 15 seconds (41).

Following completion of visit 2, active TENS or placebo TENS units were sent home with participants with an instruction manual developed by study personnel. TENS allocators used a standardized script to instruct participants on home use and for weekly contact. Participants were instructed to use TENS at least 2 hours per day during physical activity. Both active TENS and placebo TENS units monitored the number of sessions, number of minutes used, and average intensity per channel. All participants received active TENS between visit 3 and visit 4 (the nonblinded phase) with identical instructions.

Statistical analysis. Sample size was determined using data from our pilot study (25), in which a single active TENS treatment was compared to placebo TENS and no TENS (maximum SD 1.96 for movement-evoked pain). Assuming an SD of 2.0, 80% power to detect a P value of <0.05, a correlation (r) of 0.5 between pain measurements in the same subject, and a sample size of 88 per group, linear mixed model analysis of repeated measures at 3 time points (visit 2 pre-TENS, visit 3 pre-TENS, and visit 3 post-TENS) would be able to detect a clinically meaningful mean difference of at least 1.5 (equivalent to a 30% improvement in pain for this sample, which had an average baseline pain score of 5 on a 0–10 NRS), which corresponds to an effect size of 0.75. A 30% improvement in pain is considered clinically significant (42).

Both intent-to-treat (ITT) and per-protocol analyses were used to assess treatment effect. For per-protocol analysis, minimal effective treatment was defined as an average of 30 minutes each day and a minimum of 8 sessions over 4 weeks. Primary and secondary outcome variables, except for rescue medication, were compared among groups using linear mixed models for repeated measures controlling for site, as there were significant differences between sites at baseline (Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/ doi/10.1002/art.41170/abstract) (42). For the outcome variables of movement-evoked pain, resting pain, fatigue, and function during the randomized portion of the trial, the time variable comprised 4 time points: visit 2 pre-TENS, visit 2 during TENS, visit 3 pre-TENS, and visit 3 during TENS. In fitting the linear mixed model, Akaike's information criteria (AIC) and Bayesian information criteria (BIC) were used to select the covariance structure that best fit these longitudinal measures within-subject. The covariance types that were considered included compound symmetry, heterogeneous compound symmetry, first-order autoregressive, and unstructured.

Based on these model parameter estimates and the fitted covariance structure, tests of mean contrast were performed to assess the effect of TENS, compared to placebo, and control on the primary outcome measures. These assessments included 1)

testing within each treatment group for the immediate effect of TENS use (during TENS versus pre-TENS at visits 2 and 3); 2) testing within each treatment group for the long-term effect of TENS (during TENS at visit 3 versus pre-TENS at visit 2, as well as pre-TENS at visit 3 versus pre-TENS at visit 2); and 3) comparison of long-term effect of TENS according to treatment (visit 3 during TENS minus visit 2 pre-TENS, as well as visit 3 pre-TENS minus visit 2 pre-TENS, compared between treatment groups).

The analgesic effect of TENS is produced by release of inhibitory neurotransmitters (endogenous opioids, serotonin, γ-aminobutyric acid [GABA]), and thus the effects are maximal during the actual time the unit is on (14). We therefore tested the primary outcome after 4 weeks of home use at a time point when the TENS unit was active (visit 3 during TENS) and compared this outcome to visit 2 before TENS use (visit 2 pre-TENS). To account for the number of tests performed within each of the 3 sets of tests, P values with Bonferroni adjustment for multiple comparisons were used. Similar analysis was performed, using a linear mixed model for repeated measures, for the other secondary variables that were measured at 2 time points (visit 2 and visit 3). The following analyses were performed: 1) test for change (visit 3 versus visit 2) within each treatment group (Bonferroni-adjusted for 3 tests) and 2) comparison of visit 3 minus visit 2 change between treatment groups (Bonferroni-adjusted for 3 tests).

Table 1. Demographic and baseline clinical characteristics of the study participants*

	Active TENS (n = 103)	Placebo TENS (n = 99)	No TENS (n = 99)
Demographic variables†			
Age, years	44.7 ± 14.3	47.2 ± 12.6	48.6 ± 11.8
White race, %	92	92	92
Ethnicity, not Hispanic, %	95	95	95
Married/living with partner, %	33‡	51	52
Less than college graduate, %	61	61	64
Working, %	55	45	58
Health variables			
Never smoked, %	82	80	70
Body mass index, kg/m^2	34.8 ± 8.7	33.7 ± 8.8	34.0 ± 8.9
Duration of fibromyalgia, median (range) years	7 (3–12)	7 (2–14)	7 (4–15)
Opioid use for pain§, no. (%)	27 (26)	26 (26)	26 (26)
Baseline measures			
Pain with movement during 6MWT (0–10) (primary outcome measure)	6.5 ± 1.9	6.2 ± 1.9	6.4 ± 1.9
Pain with movement during 5STS (0–10) (primary outcome measure)	5.8 ± 2.4	5.5 ± 2.2	5.6 ± 2.2
Pain at rest, NRS (0–10)	6.2 ± 1.5	5.9 ± 1.4	6.1 ± 1.6
Fatigue at rest, NRS (0–10)	6.8 ± 2.0	6.1 ± 1.8	6.4 ± 2.0
Revised FIQ pain score (0–10)	6.7 ± 1.8¶	6.0 ± 1.6	6.15 ± 1.8
Revised FIQ disease impact score (0–100)	59.2 ± 16.8#	53.7 ± 15.9	55.6 ± 16.0
Mental quality of life (SF-36 mental composite score, T score)	38.7 ± 10.0	40.2 ± 10.2	39.5 ± 10.6
Physical quality of life (SF-36 physical composite score, T score)	32.7 ± 6.4	33.3 ± 6.2	32.7 ± 6.6
Pain catastrophizing (PCS, 0–52)	23.1 ± 13.0	20.4 ± 12.5	20.8 ± 12.1
Self-efficacy (PSEQ, 0–60)	28.2 ± 13.3	29.9 ± 13.1	29.0 ± 13.2
Fear of movement (TSK, 17–68)	36.5 ± 7.7	37.1 ± 8.0	37.4 ± 8.3
Anxiety (PROMIS, T score)	58.8 ± 8.7	58.1 ± 8.0	58.3 ± 7.8
Depression (PROMIS, T score)	58.1 ± 8.1	55.7 ± 8.5	56.6 ± 8.1
Function (6MWT, feet walked)	1,386 ± 323	1,358 ± 305	1,316 ± 318
Function (5TSTS, sit-to-stand times in 10 seconds)	4.1 ± 1.5	4.0 ± 1.4	3.9 ± 1.5
Physical activity, median (range) minutes per day of moderate-to-vigorous activity	17.7 (7.4–29.0)	16.5 (6.3–29.1)	15.0 (7.3–36.0)
Physical activity, IPAQ SF, median (range) METs per week	1,290 (504–3,276)	1,108 (198–2,839)	1,386 (297–2,970)
TENS use (n = 94)			
Intensity used on lumbar locations, mA	38.67 ± 7.98	–	–
Intensity used on cervical locations, mA	38.70 ± 7.24	–	–
Minutes used per day, median (range)	77.1 (51.4–109.7)	–	–

* Except where indicated otherwise, values are the mean ± SD. TENS = transcutaneous electrical nerve stimulation; 6MWT = 6-minute walk test; 5TSTS = 5-time sit-to-stand test; NRS = numerical rating scale; FIQ = Fibromyalgia Impact Questionnaire; SF-36 = short-form 36; PCS = Pain Catastrophizing Scale; PSEQ = Pain Self-Efficacy Questionnaire; TSK = Tampa Scale of Kinesiophobia; PROMIS = Patient-Reported Outcomes Measurement Information System; IPAQ SF = International Physical Activity Questionnaire short-form; METs = metabolic equivalents.
† Percentages are based on the number of participants who chose to respond.
‡ P = 0.010 versus placebo TENS and no TENS groups.
§ Groups were stratified for opioid use during randomization.
¶ P = 0.020.
# P = 0.049.

Since Revised FIQ results and marital status differed at baseline between groups, treatment effect on outcome measures was also tested with Revised FIQ results and marital status as a covariate in the model. Estimates of mean change or difference between groups with 95% confidence intervals (95% CIs) were computed. Rescue pain medication (opioid and non-opioid analgesics) was calculated as morphine equivalents for opioids and as the number of pills for non-opioid analgesics. The groups were subdivided by those who reported rescue pain medication use at visit 2 and those who did not use pain medication. We then examined the change in rescue medication between visit 2 and visit 3 to classify study participants as either 1) those who decreased the amount of rescue medication or who remained non-users or 2) those who increased or used the same amount of rescue medication. The percentages of participants in these 2 categories were compared among the treatment groups using the Cochran-Mantel-Haenszel test, controlling for site and rescue pain medication use at visit 2.

The handling of missing data in linear mixed model analysis assumes that data are missing at random. However, a subsequent analysis that examined sensitivity of the findings if data were missing not at random (with a multiple imputation approach) was performed using a control-based pattern imputation and delta-adjustment imputation (Supplementary Methods and Supplementary Table 3, available on the *Arthritis & Rheumatology* web site at http://online library.wiley.com/doi/10.1002/art.41170/abstract).

RESULTS

Participants. Between September 2013 and February 2018, 352 participants were enrolled in this study, with 301 randomly assigned to 1 of 3 groups (active TENS [n = 103], placebo TENS [n = 99], or no TENS [n = 99]), which comprised ITT analysis (Figure 1B). The majority of participants who were enrolled but not randomized failed to meet the pain severity threshold required for inclusion at visit 2. Of the 301 randomized participants, 238 were included in the per-protocol analysis (active TENS [n = 76], placebo TENS [n = 68], and no TENS [n = 94]). Participant demographics and baseline characteristics prior to randomization at visit 2 were similar between all 3 groups except for marital status and Revised FIQ (Table 1 and Supplementary Results, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary. wiley.com/doi/10.1002/art.41170/abstract). Subjects used active TENS for a median of 77.1 minutes (interquartile range [IQR] 51.4– 109.7) each day and placebo TENS for a median of 72 minutes
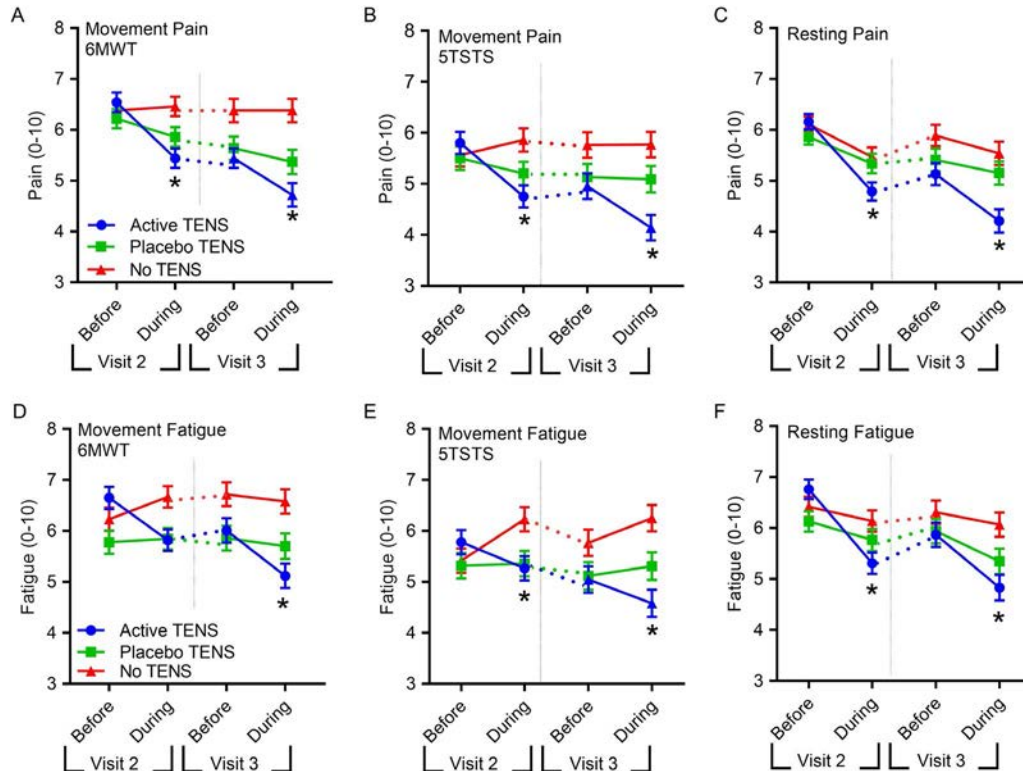
Figure 2. Active transcutaneous electrical nerve stimulation (TENS) use significantly decreased pain and fatigue in women with fibromyalgia during activity and at rest compared to placebo TENS or no TENS use, per intent-to-treat analysis. Graphs show movement-evoked and resting pain and fatigue before and during treatment at visits 2 and 3. Between visits 2 and 3, participants used TENS at home for 4 weeks (dotted lines). **A**, Movement-evoked pain during the 6-minute walk test (6MWT). **B**, Movement-evoked pain during a 5-time sit-to-stand test (5TSTS). **C**, Resting pain. **D**, Movement-evoked fatigue during the 6MWT. **E**, Movement-evoked fatigue during the 5TSTS. **F**, Resting fatigue. * = $P < 0.05$ versus placebo TENS and no TENS. Data are the mean ± SEM.

(IQR 39.4–104.6) each day. The mean ± SD stimulation intensity in the active treatment group was 38.8 ± 7.98 mA for lumbar locations and 38.7 ± 7.2 mA for cervical locations.

Outcome measures. *Pain*. After 4 weeks of active TENS use, within-group movement-evoked pain during a 6-minute walk test was significantly reduced by 1.8 points (95% CI –2.3, –1.2) compared to pre-TENS treatment on visit 2, and the reduction was greater compared to movement-evoked pain observed in the placebo TENS group (mean –0.8 [95% CI –1.4, –0.2]; P = 0.008) and the no TENS group (mean –0.006 [95% CI –0.5, 0.6]; P < 0.0001) (Figure 2 and Table 2), per ITT analysis. Similar results were obtained after adjustment for baseline Revised FIQ and marital status. There were also significant reductions in resting pain (NRS), pain intensity and interference (BPI), and pain based on the Revised FIQ in the active TENS group compared to the placebo TENS or no TENS groups after 4 weeks of home use (P < 0.05) (Figure 2 and Table 2). Per-protocol analysis demonstrated similar results for pain (Supplementary Figure 2 and Supplementary Table 4, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract). There was no significant change observed at visit 3 in rescue pain medication use after 1 month of active TENS versus placebo TENS or no TENS (Supplementary Table 5, http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract).

As part of our recruitment and retention strategy, we provided active TENS units to all participants after visit 3 for 4 weeks of home use and then tested effects on visit 4. The active TENS group (n = 75) continued to exhibit a reduction in resting and movement-evoked pain after an additional 4 weeks of home use (Table 2 and Supplementary Table 6, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract). After 4 weeks of active TENS use, the placebo TENS group (n = 68) and no TENS group (n = 94) had significant decreases in resting and movement-evoked pain (Table 2 and Supplementary Table 6).

Fatigue. After 4 weeks of active TENS use, there was a significant reduction in movement-evoked fatigue in the active TENS group compared to the placebo TENS group (P = 0.001) and the no TENS group (P < 0.0001) (Figure 2 and Table 2), per ITT analysis. Resting fatigue and MAF global fatigue index results showed significant differences between active TENS use and placebo TENS or no TENS use (P < 0.05) (Table 2). The per-protocol analysis demonstrated similar results for fatigue (Supplementary Figure 2 and Supplementary Table 4, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract).

Function, disease impact, quality of life, and pain-related psychological factors. Active TENS produced significant reductions in disease impact (Revised FIQ) and self-reported function (Revised FIQ function subscale) compared to no TENS use but not compared to placebo TENS (Table 2), per ITT analysis. No differences between groups were observed for performance-based function (6-minute walk test and 5-time sit-to-stand test), physical activity (accelerometry and IPAQ short form), fear of movement (TSK), pain catastrophizing (PCS), self-efficacy (PSEQ), anxiety (PROMIS), or quality of life (SF-36), except for a small decrease in depression (PROMIS) with active TENS use (Table 2).

Global impression of change. The GIC showed that 70% of those in the active TENS group reported global improvement compared to 31% of those in the placebo TENS group (P < 0.0001) and 9% of those in the no TENS group (P < 0.0001), by ITT analysis (Figure 3A). The GIC rating was moderately correlated with the change in movement-evoked pain (r = –0.39, P = 0.0001 by Spearman's coefficient; n = 242) (Supplementary Figure 3, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract).

Responders to TENS intervention. We defined TENS responders as subjects who exhibited the following after TENS use: pain reduction of ≥30%, fatigue reduction of ≥20%, and function improvement of ≥20% (based on percentages that have been suggested as clinically meaningful in prior studies [42,43]). For pain, the active TENS group had significantly more responders compared to the placebo TENS and no TENS groups (P = 0.004 and P < 0.001, respectively) (Figure 3B). Similarly, for fatigue, the active TENS group had significantly more responders compared to the placebo TENS and no TENS groups (P = 0.019 and P = 0.004, respectively). For function (measured with the Revised FIQ function scale), the number of responders did not differ between groups (Figure 3B).

Blinding. The outcome assessors were adequately blinded with regard to active TENS, placebo TENS, and no TENS treatment (with correct treatment identification of 45%, 13%, and 20%, respectively). The participants were blinded with regard to placebo TENS treatment (with correct treatment identification of 49%), but correctly identified active TENS treatment 70% of the time. The reduction in movement-evoked pain during the 6-minute walk test after 4 weeks of active TENS use was not significantly different in those who correctly identified active TENS treatment (–1.9 [95% CI –2.6, –1.3]) compared to those who did not correctly identify active TENS treatment (–1.4 [95% CI –2.5, –0.3]), with a mean difference of –0.56 (95% CI –1.7, 0.6) (P = 0.50).

Adverse events. There were 30 adverse events related to TENS intervention in 30 participants on visits 1, 2, or 3. The most common adverse events were pain with TENS (4.8% in the active TENS group, 4% in the placebo TENS group, and 1% in the no TENS group) and skin irritation with electrodes (4.8% in the active TENS group, 1% in the placebo TENS group, and 0% in the no TENS group). Adverse events reported on visit 2 occurred during the first treatment at that visit, and adverse events reported on visit 3 were during treatment at that visit and during the 4-week period of home use.

Table 2. Primary and secondary outcomes in the intent-to-treat analysis (corrected for study site differences at baseline)*

	Change from visit 2			Active TENS versus placebo TENS		Active TENS versus no TENS	
	Active TENS (n = 103)	Placebo TENS (n = 99)	No TENS (n = 99)	Group mean difference (95% CI)	P	Group mean difference (95% CI)	P
Pain (6MWT, 0–10)							
Visit 3 (randomized)	-1.8 (-2.3, -1.2)†	-0.8 (-1.4, -0.2)‡	0.0 (-0.5, 0.6)	-1.0 (-1.8, -0.2)	0.008	-1.8 (-2.6, -1.0)	<0.0001
Visit 4 (all with TENS)	-2.0 (-2.8, -1.3)†	-1.9 (-2.7, -1.2)†	-1.9 (-2.6, -1.2)†				
Pain (5TSTS, 0–10)							
Visit 3 (randomized)	-1.6 (-2.3, -1.0)†	-0.3 (-1.0, 0.3)	0.2 (-0.4, 0.9)	-1.3 (-2.2, -0.4)	0.002	-1.8 (-2.8, -1.0)	<0.0001
Visit 4 (all with TENS)	-1.9 (-2.6, -1.1)†	-1.4 (-2.2, -0.7)†	-1.3 (-2.1, -0.6)†				
Resting pain (0–10)							
Visit 3 (randomized)	-1.9 (-2.5, -1.4)†	-0.7 (-1.3, -0.1)§	-0.5 (-1.1, 0.0)	-1.2 (-2.1, -0.4)	0.0006	-1.4 (-2.2, -0.6)	<0.0001
Visit 4 (all with TENS)	-2.2 (-2.9, -1.6)†	-1.9 (-2.6, -1.2)†	-2.2 (-2.8, -1.5)†				
Fatigue (6MWT, 0–10)							
Visit 3 (randomized)	-1.5 (-2.2, -0.8)†	-0.1 (-0.9, 0.7)	0.4 (-0.3, 1.1)	-1.4 (-2.4, -0.4)	0.001	-1.9 (-2.9, -0.9)	<0.0001
Visit 4 (all with TENS)	-1.3 (-2.0, -0.6)†	-0.9 (-1.7, -0.2)‡	-0.9 (-1.7, -0.2)‡				
Fatigue (5TSTS, 0–10)							
Visit 3 (randomized)	-1.2 (-1.9, -0.5)†	0.0 (-0.8, 0.7)	0.8 (0.1, 1.5)§	-1.2 (-2.2, -0.2)	0.011	-2.0 (-3.0, -1.0)	<0.0001
Visit 4 (all with TENS)	-1.1 (-1.9, -0.4)†	-0.8 (-1.6, -0.1)§	-0.6 (-1.4, 0.1)				
Resting fatigue (0–10)							
Visit 3 (randomized)	-1.9 (-2.6, -1.2)†	-0.8 (-1.5, -0.04)§	-0.4 (-1.0, 0.4)	-1.2 (-2.2, -0.1)	0.016	-1.57 (-2.6, -0.6)	0.0002
Visit 4 (all with TENS)	-2.1 (-2.9, -1.4)†	-1.6 (-2.4, -0.8)†	-1.8 (-2.6, -1.1)†				
Disease impact (Revised FIQ, 0–10)							
Visit 3 (randomized)	-8.5 (-12.9, -4.0)†	-3.4 (-6.5, -0.3)§	-1.39 (-4.4, 1.6)	-5.0 (-10.4, 0.3)	0.074	-7.1 (-12.4, -1.8)	0.005
Visit 4 (all with TENS)	-9.6 (-13.8, -5.4)†	-11.1 (-15.2, 7.0)†	-10.7 (-14.8, -6.6)†				
Pain (Revised FIQ, 0–10)							
Visit 3 (randomized)	-1.3 (-1.8, -0.7)†	-0.4 (-0.9, 0.2)	-0.1 (-0.6, 0.4)	-0.9 (-1.7, -0.1)	0.018	-1.2 (-1.9, -0.4)	0.0006
Visit 4 (all with TENS)	-1.4 (-2.0, -0.8)†	-1.2 (-1.7, -0.6)†	-1.4 (-1.9, -0.8)†				
Pain (BPI interference, 0–10)							
Visit 3 (randomized)	-0.9 (-1.4, -0.5)†	-0.3 (-0.7, 0.2)	-0.3 (-0.7, 0.2)	0.7 (-1.3, 0.01)	0.043	-0.6 (-1.3, 0.0)	0.048
Visit 4 (all with TENS)	-1.1 (-1.6, -0.6)†	-0.9 (-1.4, -0.3)†	-1.2 (-1.7, -0.7)†				
Pain (BPI intensity, 0–10)							
Visit 3 (randomized)	-0.8 (-1.1, -0.4)†	-0.3 (-0.6, 0.1)	0.15 (-0.2, 0.5)	-0.5 (-1.0, 0.0)	0.036	-0.9 (-1.4, -0.4)	<0.0001
Visit 4 (all with TENS)	-1.0 (-1.4, -0.6)†	-0.9 (-1.3, -0.5)†	-0.9 (-1.2, -0.5)†				
Fatigue (MAF GFI, 1–50)							
Visit 3 (randomized)	-4.6 (-6.4, -2.8)†	-1.5 (-3.3, 0.4)	-0.3 (-2.0, 1.5)	-3.2 (-5.7, -0.6)	0.009	-4.4 (-6.8, -1.9)	<0.0001
Visit 4 (all with TENS)	-4.0 (-6.0, -1.9)†	-4.3 (-6.4, -2.2)†	-3.2 (-5.1, 1.2)†				
Self-efficacy (PSEQ, 0–60)¶							
Visit 3 (randomized)	3.2 (0.8, 5.6)†	1.5 (-0.9, 4.0)	0.8 (-1.5, 3.1)	1.6 (-1.8, 5.1)	0.75	2.3 (-1.0, 5.7)	0.28
Visit 4 (all with TENS)	5.3 (2.6, 8.0)†	4.4 (1.7, 7.1)†	4.2 (1.6, 6.8)†				
Pain catastrophizing (PCS, 0–52)							
Visit 3 (randomized)	-3.4 (-5.3, -1.4)†	-3.1 (-5.1, -1.2)†	-1.4 (-3.3, 0.5)	-0.3 (-3.0, 2.5)	>0.99	-2.0 (-4.7, 0.7)	0.23
Visit 4 (all with TENS)	-6.1 (-8.2, -3.9)†	-4.5 (-6.8, -2.3)†	-4.9 (-7.0, -2.8)†				
Fear of movement (TSK, 17–68)							
Visit 3 (randomized)	-0.7 (-2.0, 0.6)	-0.3 (-1.7, 1.0)	-0.2 (-1.4, 1.1)	-0.4 (-2.3, 1.5)	>0.99	-0.6 (-2.4, 1.3)	>0.99
Visit 4 (all with TENS)	-0.3 (-1.6, 1.1)	-2.3 (-3.7, 0.9)†	-3.3 (-4.6, -2.0)†				

(Continued)

Table 2. *(Cont'd)*

	Change from visit 2			Active TENS versus placebo TENS		Active TENS versus no TENS	
	Active TENS (n = 103)	Placebo TENS (n = 99)	No TENS (n = 99)	Group mean difference (95% CI)	P	Group mean difference (95% CI)	P
Mental quality of life (SF-36 mental composite score, T score)¶							
Visit 3 (randomized)	2.3 (0.2, 4.4)§	1.2 (−0.9, 3.4)	−0.04 (−2.1, 2.0)	1.1 (−1.9, 4.1)	>0.99	2.4 (−0.6, 5.3)	0.17
Visit 4 (all with TENS)	2.1 (−0.2, 4.4)	3.6 (1.3, 6.0)†	2.8 (0.6, 5.0)‡				
Physical quality of life (SF-36 physical composite score, T score)							
Visit 3 (randomized)	2.4 (1.0, 3.7)†	1.2 (−0.2, 2.5)	1.4 (0.1, 2.6)	1.2 (−0.7, 3.1)	0.36	1.0 (−0.8, 2.8)	0.58
Visit 4 (all with TENS)	3.5 (2.0, 5.1)†	3.2 (1.6, 4.8)†	4.4 (2.9, 5.9)†				
Anxiety (PROMIS, T score)							
Visit 3 (randomized)	−1.1 (−2.6, 0.5)	−0.6 (−2.1, 1.0)	−0.7 (−2.1, 0.8)	−0.5 (−2.7, 1.7)	>0.99	−0.4 (−2.5, 1.7)	>0.99
Visit 4 (all with TENS)	−0.5 (−2.1, 1.2)	−1.6 (−3.3, 0.1)	−2.2 (−3.8, −0.6)‡				
Depression (PROMIS, T score)							
Visit 3 (randomized)	−2.8 (−4.2, −1.5)†	−0.1 (−1.5, 1.3)	0.4 (−0.9, 1.7)	−2.7 (−4.7, −0.8)	0.002	−3.2 (−5.1, −1.3)	0.0001
Visit 4 (all with TENS)	−2.0 (−3.4, −0.6)‡	−1.3 (−2.7, 0.1)	−1.2 (−2.6, 0.2)				
Self-reported function (Revised FIQ, 0–30)							
Visit 3 (randomized)	−2.7 (−4.0, −1.4)†	−1.4 (−2.7, −0.1)§	−0.6 (−1.8, 0.7)	−1.3 (−3.2, 0.5)	0.25	−2.1 (−3.9, −0.4)	0.013
Visit 4 (all with TENS)	−2.8 (−4.2, −1.4)†	−3.7 (−5.1, −2.3)†	−3.6 (−4.9, −2.3)†				
Self-reported function (SF-36, T score)¶							
Visit 3 (randomized)	1.4 (0.1, 2.7)§	0.5 (−0.8, 1.8)	0.8 (−0.5, 2.0)	0.9 (−1.0, 2.7)	0.79	0.6 (−1.2, 2.4)	>0.99
Visit 4 (all with TENS)	2.9 (1.5, 4.4)†	2.3 (0.9, 3.8)†	3.0 (1.6, 4.4)†				
Self-reported function/physical activity (IPAQ SF, % change in METs per week)							
Visit 3 (randomized)	8.1 (−22.0, 49.9)	52.1 (9.0, 112.2)‡	−14.2 (−37.5, 17.8)		0.24		0.67
Visit 4 (all with TENS)	38.1 (−1.2, 93.0)	46.3 (4.1, 105.5)§	29.9 (−6.3, 80.1)				
Performance-based function (6MWT, feet walked)							
Visit 3 (randomized)	−1 (−55, 54)	−20 (−75, 36)	−42.1 (−95, 11)	19 (−58, 96)	>0.99	42 (−34, 117)	>0.99
Visit 4 (all with TENS)	15 (−42, 71)	16 (−41, 74)	−2 (−57, 53)				
Performance-based function (5TSTS, sit-to-stand times in 10 seconds)							
Visit 3 (randomized)	0.6 (0.3, 1.0)†	0.4 (0.0, 0.7)§	0.1 (−0.3, 0.4)	0.2 (−0.2, 0.7)	0.96	0.6 (0.1, 1.0)	0.008
Visit 4 (all with TENS)	0.6 (0.2, 1.0)†	0.8 (0.4, 1.3)†	0.6 (0.1, 1.0)‡				
Performance-based function (minutes per day of moderate-to-vigorous physical activity, % change)							
Visit 3 (randomized)	−9.4 (−27.8, 13.5)	2.5 (−18.6, 29.1)	−14.1 (−29.5, 4.7)		>0.99		>0.99
Visit 4 (all with TENS)	−6.8 (−27.2, 19.3)	−0.3 (−22.3, 27.9)	−17.0 (−33.9, 4.1)				

* Mean change (from baseline at visit 2) to visit 3 (after 4 weeks of home use of active TENS, placebo TENS, or no TENS) and visit 4 (after 4 weeks of home TENS use in all groups) with adjusted 95% confidence intervals (95% CIs). Change scores are presented for the randomized phase between visit 2 and visit 3, and for the difference from baseline at visit 4 when all subjects received active TENS. P values represent post hoc comparisons between groups. BPI = Brief Pain Inventory; MAF = Multidimensional Assessment of Fatigue; GFI = Global Fatigue Index (see Table 1 for other definitions).
† P = <0.001.
‡ P = <0.01.
§ P = <0.05.
¶ Higher score represents improvement.
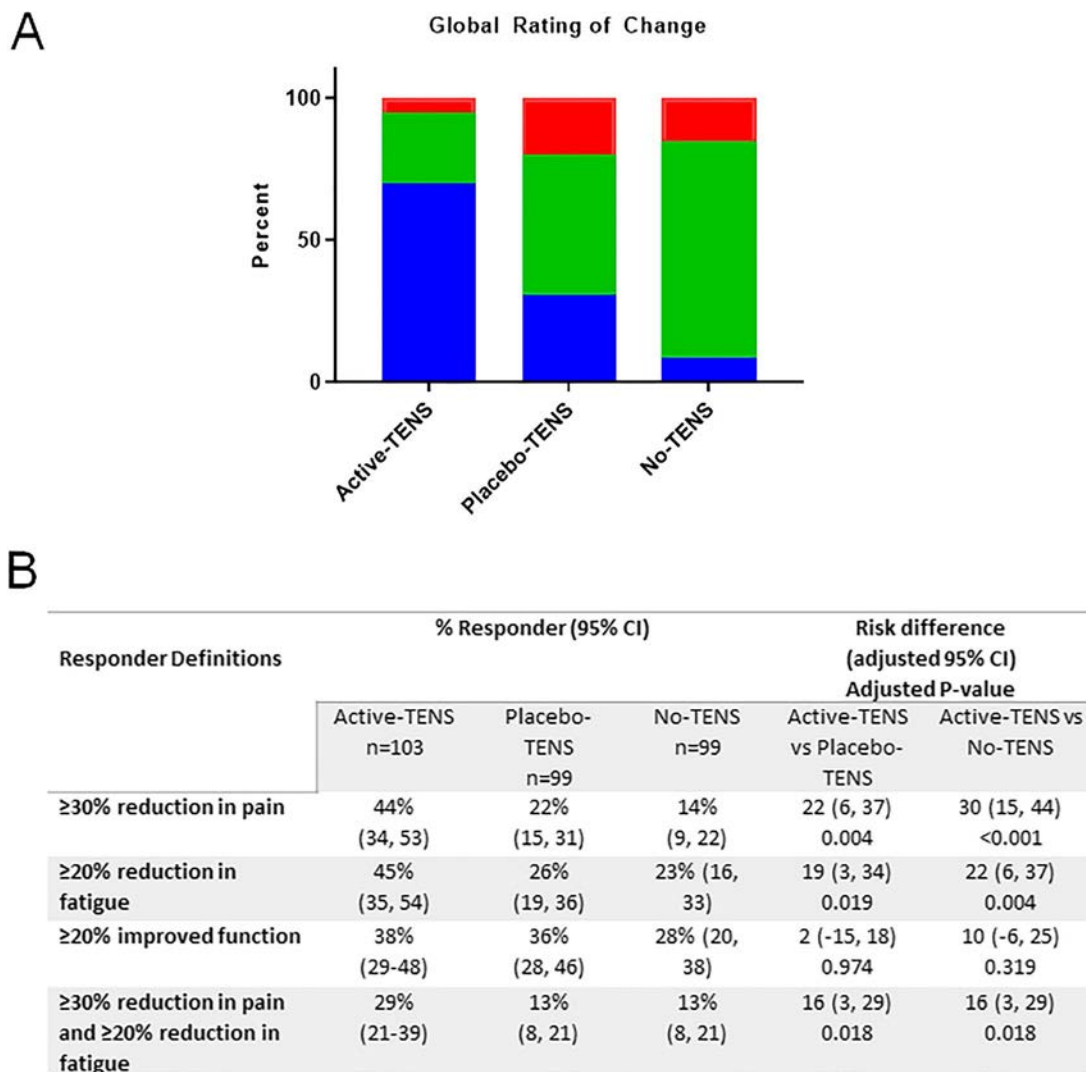
Figure 3. The active TENS group had an improved perception of change and a greater number of responders with regard to the degree of change in movement-evoked pain or fatigue during a 6MWT compared to the placebo TENS or no TENS groups. **A**, The percentage of participants who reported feeling better or much better (blue), no change (green), and worse or extremely worse (red) after 4 weeks of active TENS, placebo TENS, or no TENS treatment. The majority of individuals reported a significant overall improvement after using active TENS compared to those who used placebo TENS or no TENS ($P < 0.0001$). There were no differences between the placebo TENS and no TENS groups ($P = 0.175$). **B**, Percentages (with 95% confidence intervals [95% CIs]) of subjects in each treatment group who had a clinically meaningful response to TENS (as described by Arnold and colleagues [43], i.e., ≥30% reduction [pain], ≥20% reduction [fatigue, function]). There was a significantly greater number of responders for pain, fatigue, and both pain and fatigue in the active TENS group compared to the placebo TENS and no TENS groups. Function was measured according to the Revised Fibromyalgia Impact Questionnaire (39). See Figure 2 for other definitions. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract.

Supplementary Table 7, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41170/abstract, shows rates of TENS-related adverse events by visit. There were 4 serious adverse events, with none related to TENS use (Supplementary Results, http://online library.wiley.com/doi/10.1002/art.41170/abstract).

DISCUSSION

This double-blind, randomized, controlled trial showed that active TENS use significantly reduced movement-evoked and resting pain and fatigue compared to placebo TENS or no TENS use in women with FM. The current study used a classic design to examine active TENS–specific effects compared to placebo TENS use, but also included a unique method of comparison (no TENS use) as a more pragmatic application that includes both specific and nonspecific treatment effects. While participants in the active TENS group correctly identified the intervention 70% of the time, there was no difference in pain reduction between participants who correctly guessed the treatment and participants who did not—an argument against a difference based on inadequate blinding. We also demonstrated adequate blinding with regard to placebo TENS

intervention by utilizing our novel transient placebo TENS unit with repeated use, further validating the placebo TENS intervention from our prior study that showed excellent blinding with a single use (41).

Our primary outcome measure of movement-evoked pain lacks formally validated minimally important differences in the published literature; however, we utilized the general thresholds recommended by the Outcome Measures in Rheumatology group (30% reduction in pain, 20% reduction in fatigue, and 20% reduction in functional impairment) (43). In the active TENS group, 44% of participants exhibited clinically important reductions of 30% and 45% in pain and fatigue, respectively, with 29% exhibiting a reduction in both pain and fatigue, suggesting that a subpopulation of individuals with FM responds well to TENS. The responder rates for pain are similar to those obtained with Food and Drug Administration–approved pharmaceutical agents for FM, such as duloxetine or pregabalin, which demonstrate responder rates of 31–41% based on a 30% reduction in pain (44,45). The comparative reductions in movement-evoked pain with active TENS compared to placebo TENS were small, averaging a 1-point difference on an 11-point scale. However, the comparative reduction was 1.8 when compared to no TENS. A 1.8-point decrease equates to a $\geq$30% reduction for individuals with a pain rating of $\leq$6 (42). It should be noted that some studies also suggest that a 2-point reduction in pain is the clinically relevant threshold (46). Future studies should be conducted to identify which patients are most likely to respond to TENS, which would be an inexpensive, safe, and easy-to-disseminate intervention for pain management.

The reductions in pain and fatigue, 2 common symptoms in FM, likely contributed to improvements in GIC reported by individuals who received active TENS. While pain is a defining characteristic of FM, fatigue is also a common symptom, occurring in the majority of individuals with FM (47). Both pain and fatigue contribute significantly to perceived disability and function (48), and the global rating of change is associated with improvements in pain and fatigue in individuals with FM (49,50). We show in our trial, for the first time, the relationship between global improvement and movement-evoked pain, with similar results to those obtained in prior studies on pain and GIC (49,50). The magnitude of reduction in both pain and fatigue observed in the current study is likely to have a significant impact on the day-to-day experience of FM patients (51,52).

In a recent Cochrane review, Johnson et al examined the efficacy of TENS treatment for individuals with FM and concluded that there was insufficient high-quality evidence (17). The main concerns were inadequately powered studies (n = 5–43 participants per group) with incomplete and limited outcome reporting. Additional concerns in regard to TENS clinical studies include use of an adequate placebo with blinding of participant and assessors, timing of outcome measurements, adequate dosing of TENS, and monitoring of concurrent analgesia (21,22). The current study was designed to address these concerns, as well as the multiple dimensions of FM recommended by professional

societies as clinically important domains (53). While we showed significant effects on pain, fatigue, and global improvement, we failed to detect a change in several FM domains, including function, psychological factors, and quality of life, after 4 weeks of TENS use. The lack of effect on function could be related to the short duration of TENS use, as functional changes may take a longer period of time to change or are harder to achieve, particularly in longstanding conditions such as FM. As an example, most exercise studies require 8–24 weeks for improved function, with improvements varying between 10% and 20% (54). It is also possible that TENS may improve adherence to an exercise task, while not directly improving function. Future studies are needed in order to examine more long-term effects in a pragmatic setting so as to evaluate the interactions between function and adherence with physical activity interventions.

Uniquely, the current study, compared to prior TENS studies (22), examined outcomes during TENS treatment. TENS activates endogenous inhibitory pathways in the central nervous system, releasing the inhibitory neurotransmitters serotonin, opioids, and GABA to reduce sensitization of central neurons (14,55), and the greatest effects occur during stimulation, when endogenous inhibitory neurotransmitters are released. We also demonstrate that active TENS has a cumulative effect because there was a greater reduction after 4 weeks of home use when compared to the first single treatment. Further, active TENS remained effective after 8 weeks of repeated home use in the active TENS group. This is important as we have previously reported the development of analgesic tolerance to the repeated use of high-frequency or low-frequency TENS in animals and human participants, a phenomenon mediated by endogenous opioids (56,57). The lack of tolerance to repeated use is likely a result of the mixed frequency used in the current study, as prior studies in animals have shown decreased tolerance with mixed frequencies (58). Thus, understanding the mechanisms underlying analgesia produced by nonpharmacologic interventions is critical to the design of an adequate clinical trial to detect clinical effectiveness.

Importantly, the current study recorded adverse events in the active TENS and placebo TENS groups and showed minimal adverse events. Fewer than 5% of individuals receiving active TENS or placebo TENS reported pain with TENS or irritation with electrodes. This demonstrates that TENS is safe and could be a useful treatment for home use.

There are several limitations to the study. The inability to achieve full blinding with regard to the active intervention, as noted above, may lead to reporting bias by the subject. However, if TENS is given at an adequate intensity (strong, but comfortable or greater [23,41]) to produce analgesic effects, this limitation may be unavoidable in a clinical trial of TENS. Data on medication usage were collected by self-report, and thus subject recall bias and willingness of the subject to fill out the log can influence results. In the current study, 75% of subjects filled out enough logs for us to examine change in medication

use as an outcome. A greater number of subjects withdrew from the active TENS group (n = 15) and the placebo TENS group (n = 11) group than the no TENS group (n = 4), which might suggest an unwillingness of some individuals to use the units on a regular basis. Additionally, this study was performed only in women with FM and thus the findings may not translate to men with FM. Last, the trial was designed as a 4-week intervention, and while we observed significant effects on pain and fatigue, we did not see significant effects on function, rescue medication usage, or psychological outcomes. It is possible that longer treatment is necessary to see effects on these other outcomes, and thus future experiments will be needed to examine the more long-term impact of TENS use.

In conclusion, among women with FM who were on a stable medication regimen, the use of active TENS compared to placebo TENS for 4 weeks resulted in significantly reduced movement-evoked pain. The use of active TENS compared to no TENS, as would be done clinically, resulted in a statistically and clinically meaningful improvement in movement-evoked pain. Further research is needed for replication, to assess the duration of effect, and to establish clinical importance of these findings.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Sluka 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. Dailey, Vance, Rakel, Zimmerman, Crofford, Sluka.

Acquisition of data. Dailey, Vance, Embree, Merriwether, Geasland, Chimenti, Williams, Golchha, Crofford, Sluka.

Analysis and interpretation of data. Dailey, Vance, Zimmerman, Merriwether, Geasland, Crofford, Sluka.

REFERENCES

1. Okifuji A, Hare BD. Management of fibromyalgia syndrome: review of evidence. Pain Ther 2013;2:87–104.

2. Vincent A, Whipple MO, McAllister SJ, Aleman KM, St Sauver JL. A cross-sectional assessment of the prevalence of multiple chronic conditions and medication use in a sample of community-dwelling adults with fibromyalgia in Olmsted County, Minnesota. BMJ Open 2015;5:e006681.

3. Qaseem A, Wilt TJ, McLean RM, Forciea MA, for the Clinical Guidelines Committee of the American College of Physicians. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017;166:514–30.

4. Macfarlane GJ, Kronisch C, Atzeni F, Häuser W, Choy EH, Amris K, et al. EULAR recommendations for management of fibromyalgia [letter]. Ann Rheum Dis 2017;76:e54.

5. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. MMWR Recomm Rep 2016;65:1–49.

6. Bement MH, Sluka KA. Exercise-induced hypoalgesia: an evidence-based review. In: Sluka KA, editor. Mechanisms and management of pain for the physical therapist. 2nd ed. Philadelphia: Wolters Kluwer Health; 2016. p. 177–202.

7. Bidonde J, Busch AJ, Bath B, Milosavljevic S. Exercise for adults with fibromyalgia: an umbrella systematic review with synthesis of best evidence. Curr Rheumatol Rev 2014;10:45–79.

8. Dailey DL, Keffala VJ, Sluka KA. Do cognitive and physical fatigue tasks enhance pain, cognitive fatigue, and physical fatigue in people with fibromyalgia? Arthritis Care Res (Hoboken) 2015;67:288–96.

9. Bair MJ, Matthias MS, Nyland KA, Huffman MA, Stubbs DL, Kroenke K, et al. Barriers and facilitators to chronic pain self-management: a qualitative study of primary care patients with comorbid musculoskeletal pain and depression. Pain Med 2009;10:1280–90.

10. Ma YT, Sluka KA. Reduction in inflammation-induced sensitization of dorsal horn neurons by transcutaneous electrical nerve stimulation in anesthetized rats. Exp Brain Res 2001;137:94–102.

11. Sluka KA, Deacon M, Stibal A, Strissel S, Terpstra A. Spinal blockade of opioid receptors prevents the analgesia produced by TENS in arthritic rats. J Pharmacol Exp Ther 1999;289:840–6.

12. Kalra A, Urban MO, Sluka KA. Blockade of opioid receptors in rostral ventral medulla prevents antihyperalgesia produced by transcutaneous electrical nerve stimulation (TENS). J Pharmacol Exp Ther 2001;298:257–63.

13. Radhakrishnan R, King EW, Dickman J, Herold CA, Johnston NF, Spurgin ML, et al. Spinal 5-HT(2) and 5-HT(3) receptors mediate low, but not high, frequency TENS-induced antihyperalgesia in rats. Pain 2003;105:205–13.

14. Vance CG, Dailey DL, Rakel BA, Sluka KA. Using TENS for pain control: the state of the evidence. Pain Manag 2014;4:197–209.

15. Clauw DJ. Fibromyalgia: a clinical review. JAMA 2014;311:1547–55.

16. Sluka KA, Clauw DJ. Neurobiology of fibromyalgia and chronic widespread pain. Neuroscience 2016;338:114–29.

17. Johnson MI, Claydon LS, Herbison GP, Jones G, Paley CA. Transcutaneous electrical nerve stimulation (TENS) for fibromyalgia in adults [review]. Cochrane Database Syst Rev 2017;10:CD012172.

18. Johnson MI, Paley CA, Howe TE, Sluka KA. Transcutaneous electrical nerve stimulation for acute pain [review]. Cochrane Database Syst Rev 2015;CD006142.

19. Chen LX, Zhou ZR, Li YL, Ning GZ, Li Y, Wang XB, et al. Transcutaneous electrical nerve stimulation in patients with knee osteoarthritis: evidence from randomized-controlled trials. Clin J Pain 2016;32:146–54.

20. Resende L, Merriwether E, Rampazo ÉP, Dailey D, Embree J, Deberg J, et al. Meta-analysis of transcutaneous electrical nerve stimulation for relief of spinal pain. Eur J Pain 2018;22:663–78.

21. Bennett MI, Hughes N, Johnson MI. Methodological quality in randomised controlled trials of transcutaneous electric nerve stimulation for pain: low fidelity may explain negative findings. Pain 2011;152:1226–32.

22. Sluka KA, Bjordal JM, Marchand S, Rakel BA. What makes transcutaneous electrical nerve stimulation work? Making sense of the mixed results in the clinical literature. Phys Ther 2013;93:1397–402.

23. Moran F, Leonard T, Hawthorne S, Hughes CM, McCrum-Gardner E, Johnson MI, et al. Hypoalgesia in response to transcutaneous electrical nerve stimulation (TENS) depends on stimulation intensity. J Pain 2011;12:929–35.

24. Rakel B, Frantz R. Effectiveness of transcutaneous electrical nerve stimulation on postoperative pain with movement. J Pain 2003;4:455–64.

25. Dailey DL, Rakel BA, Vance CG, Liebano RE, Amrit AS, Bush HM, et al. Transcutaneous electrical nerve stimulation reduces pain, fatigue and hyperalgesia while restoring central inhibition in primary fibromyalgia. Pain 2013;154:2554–62.

26. Cowan S, McKenna J, McCrum-Gardner E, Johnson MI, Sluka KA, Walsh DM. An investigation of the hypoalgesic effects of TENS delivered by a glove electrode. J Pain 2009;10:694–701.

27. Sunshine W, Field TM, Quintino O, Fierro K, Kuhn C, Burman I, et al. Fibromyalgia benefits from massage therapy and transcutaneous electrical stimulation. J Clin Rheumatol 1996;2:18–22.

28. Löfgren M, Norrbrink C. Pain relief in women with fibromyalgia: a cross-over study of superficial warmth stimulation and transcutaneous electrical nerve stimulation. J Rehabil Med 2009;41:557–62.

29. Lauretti GR, Chubaci EF, Mattos AL. Efficacy of the use of two simultaneously TENS devices for fibromyalgia pain. Rheumatol Int 2013;33:2117–22.

30. Carbonario F, Matsutani LA, Yuan SL, Marques AP. Effectiveness of high-frequency transcutaneous electrical nerve stimulation at tender points as adjuvant therapy for patients with fibromyalgia. Eur J Phys Rehabil Med 2013;49:197–204.

31. Noehren B, Dailey DL, Rakel BA, Vance CG, Zimmerman MB, Crofford LJ, et al. Effect of transcutaneous electrical nerve stimulation on pain, function, and quality of life in fibromyalgia: a double-blind randomized clinical trial. Phys Ther 2015;95:129–40.

32. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: report of the Multicenter Criteria Committee. Arthritis Rheum 1990;33:160–72.

33. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singapore 1994;23:129–38.

34. Belza BL. Comparison of self-reported fatigue in rheumatoid arthritis and controls. J Rheumatol 1995;22:639–43.

35. PloS One. International Physical Activity Questionnaire - short form. URL: https://journals.plos.org/plosone/article/file?type=supplementary&id=info:doi/10.1371/journal.pone.0219193.s010.

36. Miller RP, Kori S, Todd D. The Tampa Scale: a measure of kinesiophobia. Clin J Pain 1991;7:51–2.

37. Sullivan MJ, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychol Assess 1995;7:524–32.

38. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 2007;11:153–63.

39. Bennett RM, Friend R, Jones KD, Ward R, Han BK, Ross RL. The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. Arthritis Res Ther 2009;11:R120.

40. 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.

41. Rakel B, Cooper N, Adams HJ, Messer BR, Frey Law LA, Dannen DR, et al. A new transient sham TENS device allows for investigator blinding while delivering a true placebo treatment. J Pain 2010;11:230–8.

42. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94:149–58.

43. Arnold LM, Williams DA, Hudson JI, Martin SA, Clauw DJ, Crofford LJ, et al. Development of responder definitions for fibromyalgia clinical trials. Arthritis Rheum 2012;64:885–94.

44. Straube S, Derry S, Moore RA, Cole P. Cervico-thoracic or lumbar sympathectomy for neuropathic pain and complex regional pain syndrome [review]. Cochrane Database Syst Rev 2013;CD002918.

45. Moore RA, Wiffen PJ, Derry S, Toelle T, Rice AS. Gabapentin for chronic neuropathic pain and fibromyalgia in adults [review]. Cochrane Database Syst Rev 2014;CD007938.

46. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004;8:283–91.

47. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RL, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–29.

48. Dailey DL, Frey Law LA, Vance CG, Rakel BA, Merriwether EN, Darghosian L, et al. Perceived function and physical performance are associated with pain and fatigue in women with fibromyalgia. Arthritis Res Ther 2016;18:68.

49. Arnold LM, Zlateva G, Sadosky A, Emir B, Whalen E. Correlations between fibromyalgia symptom and function domains and patient global impression of change: a pooled analysis of three randomized, placebo-controlled trials of pregabalin. Pain Med 2011;12:260–7.

50. Rampakakis E, Ste-Marie PA, Sampalis JS, Karellis A, Shir Y, Fitzcharles MA. Real-life assessment of the validity of patient global impression of change in fibromyalgia. RMD Open 2015;1:e000146.

51. Arnold LM, Crofford LJ, Mease PJ, Burgess SM, Palmer SC, Abetz L, et al. Patient perspectives on the impact of fibromyalgia. Patient Educ Couns 2008;73:114–20.

52. Okifuji A, Bradshaw DH, Donaldson GW, Turk DC. Sequential analyses of daily symptoms in women with fibromyalgia syndrome. J Pain 2011;12:84–93.

53. Mease PJ, Clauw DJ, Christensen R, Crofford LJ, Gendreau RM, Martin SA, et al. Toward development of a fibromyalgia responder index and disease activity score: OMERACT module update. J Rheumatol 2011;38:1487–95.

54. Bidonde J, Busch AJ, Schachter CL, Overend TJ, Kim SY, Góes SM, et al. Aerobic exercise training for adults with fibromyalgia [review]. Cochrane Database Syst Rev 2017;6:CD012700.

55. Sluka KA, Walsh DM. Transcutaneous electrical nerve stimulation and interferential therapy. In: Sluka KA, editor. Pain mechanisms and management for the physical therapist. 2nd ed. Philadelphia: Wolters Kluwer Health; 2016. p. 203–23.

56. Chandran P, Sluka KA. Development of opioid tolerance with repeated transcutaneous electrical nerve stimulation administration. Pain 2003;102:195–201.

57. Liebano R, Rakel B, Vance CG, Walsh DM, Sluka KA. An investigation of the development of analgesic tolerance to TENS in humans. Pain 2011;152:335–42.

58. DeSantana JM, Santana-Filho VJ, Sluka KA. Modulation between high- and low-frequency transcutaneous electric nerve stimulation delays the development of analgesic tolerance in arthritic rats. Arch Phys Med Rehabil 2008;89:754–60.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Functional and Structural Adaptations of Skeletal Muscle in Long-Term Juvenile Dermatomyositis: A Controlled Cross-Sectional Study

Kristin Schjander Berntsen,[1] [iD] Truls Raastad,[2] Henriette Marstein,[3] Eva Kirkhus,[1] Else Merckoll,[1] Kristoffer Toldnes Cumming,[2] Berit Flatø,[3] Ivar Sjaastad,[3] and Helga Sanner[4]

Objective. To compare muscle strength and endurance of the knee extensors between patients with long-term juvenile dermatomyositis (DM) and controls and between patients with active disease and those with inactive disease, and to explore associations between strength/endurance and 1) clinical parameters, 2) physical activity, and 3) humoral/structural adaptation in the skeletal muscle of patients.

Methods. In a cross-sectional study (44 patients and 44 age- and sex-matched controls), we tested isometric muscle strength (peak torque, in Nm) and dynamic muscle endurance (total work, in Joules) of the knee extensors, physical activity (measured by accelerometer), and serum myokine levels (by enzyme-linked immunosorbent assay). Patients were examined with validated tools (clinical muscle tests and measures of disease activity/damage and inactive disease) and using magnetic resonance imaging of the thigh muscles, which included evaluation of the quadriceps cross-sectional area (CSA). Needle biopsy samples of the vastus lateralis muscle (obtained from 12 patients ages ≥18 years) were assessed by histochemistry.

Results. After a mean ± SD disease duration of 21.8 ± 11.8 years, peak torque was lower in patients with juvenile DM compared to controls (mean difference 29 Nm, 95% confidence interval 13–46; P = 0.001). Similarly, total work of the knee extensors was lower in patients compared to controls (median 738J [interquartile range 565–1,155] versus 1,249J [interquartile range 815–1,665]; P < 0.001). Both peak torque and total work were lower in patients with active juvenile DM compared to those with inactive disease (both P < 0.019); in analyses controlled for quadriceps CSA, only total work remained lower in patients with active disease. Moreover, peak torque and total work correlated with findings from clinical muscle tests in patients with active disease (r = 0.57–0.84). Muscle biopsy results indicated that the fiber type composition was different, but capillary density was similar, between patients with active disease and those with inactive disease.

Conclusion. In patients with long-term juvenile DM, both muscle strength and endurance of the knee extensors were lower when compared to matched controls, and also lower in patients with active disease compared to those with inactive disease. Our results indicate a need for more sensitive muscle tests in this clinical setting. We hypothesize that impaired muscle endurance in patients with active juvenile DM may be influenced by structural/functional adaptations of muscle tissue independent of muscle size.

INTRODUCTION

Reduced muscle strength and endurance are major clinical manifestations of juvenile dermatomyositis (DM), the most common idiopathic inflammatory myopathy (IIM) of childhood (1).

Muscle involvement often comprises the proximal muscles of the extremities as well as truncal muscles, including neck flexors (2).

During the early phase of juvenile DM, decreased muscle function can be severe and is related to active myositis (3). Autoimmune-like mechanisms are believed to contribute to

perivascular inflammation and vasculopathy, resulting in ischemic muscle fiber damage, perifascicular atrophy, and reduced capillary density (4). However, nonimmune mechanisms related to reduced blood flow also seem to be involved in impaired muscle function (3,5). Both immune and nonimmune processes may induce the release of myokines, which are cytokines derived from muscle tissue (6). Myokines are associated with muscle inflammation, but also are believed to mediate antiinflammatory effects related to exercise (6).

With treatment, muscle strength and endurance gradually improve. Yet, long-term outcome studies have shown persistent, mild muscle weakness and decreased muscle endurance, tested clinically using the unilateral Manual Muscle Testing in 8 groups (MMT-8) and the Childhood Myositis Assessment Scale (CMAS) (7–9), both of which are features that are more pronounced in patients with active juvenile DM than in those with inactive disease (10). Severe impairment is rare (7). However, a challenge related to the MMT-8 and CMAS is their frequently observed ceiling effects (11). Mild, but functionally important, muscle weakness may therefore be difficult to detect, especially in patients with inactive disease (10).

Testing of isometric strength and dynamic muscular endurance could provide a more objective and sensitive method for evaluating muscle function in juvenile DM patients who experience mild muscle impairment (12). Thigh muscles, including the knee extensors, are among the most commonly involved muscles in juvenile DM (13), comprising the muscle group most frequently examined by biopsy (14) and magnetic resonance imaging (MRI) (15). Therefore, knee extensors could serve as a representative test localization for proximal muscle function in juvenile DM.

After long-term disease in patients with juvenile DM, scores on the MMT-8 or CMAS have been found to be associated with levels of disease activity (9,16) and damage (16) and elevated serum myokine levels (17). Despite the association with active inflammation, myokines may also be involved in persistent muscle weakness in noninflamed muscle in which inflammation has been suppressed by targeted treatment (6). However, the associations between objective muscle test findings and these parameters in long-term juvenile DM are not known.

To our knowledge, no studies on muscle fiber composition exist in patients after long-term juvenile DM. A study of adult patients with DM found altered muscle fiber composition in those with chronic disease (18). Muscle fiber composition is dynamic, with the size and proportion of slow-twitch oxidative type I and fast-twitch fiber type II muscle fibers changing according to numerous factors, including age, sex, and exercise levels (19,20). Exercise was found to increase the proportion of type I muscle fibers in adult DM (21).

In this study, we aimed to compare isometric muscle strength and dynamic muscle endurance of the knee extensors, measured by sensitive, objective methods, between patients with long-term juvenile DM and controls, and between patients with active juvenile DM and those with inactive disease. Furthermore, we aimed to explore whether changes in thigh muscle strength and endurance in patients are associated with 1) disease parameters, including results of commonly used clinical muscle tests, 2) physical activity, and 3) structural or humoral adaptation of the skeletal muscle.

PATIENTS AND METHODS

Design. We used a controlled cross-sectional study design. The study was part of a larger project on physical fitness in juvenile DM patients at Oslo University Hospital (OUS) and the Norwegian School of Sport Sciences between 2013 and 2015. Patients were recruited from an already established juvenile DM cohort (22). In addition, 8 patients from a prospective juvenile DM cohort at OUS and 3 additional patients were invited.

The inclusion criteria were as follows: 1) a diagnosis of juvenile DM after 1970, 2) a diagnosis of definite or probable dermatomyositis according to the Bohan and Peter criteria (23), 3) having been diagnosed with juvenile DM at age <18 years, and 4) being age ≥10 years at the time of examination. Patients were excluded from the data analyses if they had not completed tests of isometric muscle strength and muscle endurance or had not undergone a muscle biopsy. Patients were scored retrospectively according to the 2017 European League Against Rheumatism (EULAR)/American College of Rheumatology (ACR) classification criteria for adult and juvenile IIMs (24).

Controls were randomly drawn from the Norwegian National Registry, and were age- and sex-matched 1:1 to the patients. Exclusion criteria were as follows: 1) mobility problems, 2) presence of inflammatory rheumatic disease, 3) presence of other active autoimmune disease, 4) presence of other autoimmune disease being treated with immunosuppressive agents, 5) presence of serious lung or heart disease, and 6) exclusion of the matched patient.

All participants (or if age <16 years, their guardians) provided signed informed consent, in accordance with the Declaration of Helsinki (25). The study was approved by the Norwegian Regional Committee for Medical and Health Research Ethics (approval no. 2013/1039).

Clinical examination. In patients with juvenile DM, we used the Disease Activity Score (DAS) (scale 0–20) and the physician global assessment of disease activity visual analog scale (VAS) score (scale 0–10) to assess global disease activity, and used the Myositis Damage Index (MDI) (scale 0–40) and the physician global damage VAS score (scale 0–10) to assess global disease damage (11). We used the MMT-8 (scale 0–80) (including the separate MMT knee extensor component [scale 0–10]) and the CMAS (scale 0–52) to clinically assess muscle strength and endurance (11). We defined an MMT-8 score of <64 or CMAS score of <35 as severe impairment (7). We used the DAS muscle component score (scale 0–11) to assess disease activity in the muscle (26), and the MDI muscle damage extent score (scale 0–3) and MDI muscle VAS severity score (scale 0–10) to assess disease damage in the muscle (27). We divided patients into those with active disease and those with inactive disease based on the

original Paediatric Rheumatology International Trials Organisation criteria for inactive disease (28).

Self-reported health. To evaluate physical function, we used the Norwegian version of the 36-item Short-Form health survey physical component score (scale 0–100) in patients and controls who were age 14 years or older, and used the Childhood Health Assessment Questionnaire (C-HAQ) and adult HAQ (each on a scale of 0–3) in patients ages <18 years and ages ≥18 years, respectively (11).

Physical activity. As previously described, we measured physical activity levels in patients and controls using waist-borne accelerometers for 7 consecutive days (29). Physical activity was evaluated according to the number of minutes spent in sedentary, light, or moderate-to-vigorous physical activity (MVPA) for the total wear period, divided by the number of valid days. Each registered minute was labeled as sedentary, light, or MVPA based on the count value for the given minute (<100 counts, 100–1,999 counts, and >2,000 counts, respectively) (30). MVPA bouts were defined as the average daily time of MVPA in bouts of of at least 10 minutes' duration.

Objective muscle testing of patients and controls. We used the maximal voluntary isometric contraction (MVC) force of knee extension, expressed as the peak torque (in Nm), to measure muscle strength, and dynamic knee extensions, expressed as the total work (in Joules), to measure muscle endurance (collectively referred to as objective muscle tests). A custom-made knee extension device (GYM 2000AS; Vikersund) was set up as previously described (31). Following a warm-up protocol, participants performed 3 consecutive unilateral maximal isometric contractions of the knee extensors that, under strong verbal encouragement, lasted 5 seconds, with each separated by rest periods of 60 seconds. We processed the data using LabVIEW software (National Instruments), and used the average of the maximum force for each leg for statistical analyses. We calculated the peak torque (in Nm) as follows: peak torque = force (in Newtons) × lever arm length (in meters).

To measure muscle endurance, a resistance mass of 30% of the MVC force was attached to the knee extension device. Guided by a metronome paced at 1 Hz, participants performed rounds of full knee extension and 90° flection until exhaustion was reached (defined as the point at which the participant was incapable of full knee extension). The average maximal number of extensions between the right and left leg was used for statistical analyses. We calculated dynamic muscle endurance as the total work (in Joules), as follows: total work = 30% of peak torque (in Nm) × sin (90°) × number of repetitions.

MRI. Patients underwent MRI of the thigh muscles using a 1.5T scanner (Siemens) with phased-array body coils, including transversal T1 turbo-spin echo and short tau inversion recovery sequences. Three of the MRIs performed at local hospitals were summoned and scored collectively with the remaining cohort. Two experienced musculoskeletal radiologists (EK and EM) assessed the presence or absence of edema in the muscle and calcinosis in the soft tissue layers, and scored pathologic fatty infiltration in the muscle on a scale of 0–4 (32), in which a score of 0 = normal, 1 = fatty streaks (interpreted as not pathologic), 2 = muscle greater than fat, 3 = muscle equal to fat, and 4 = muscle less than fat. They also measured the maximal cross-sectional area (CSA; in cm^2) of the anterior thigh compartment (quadriceps femoris) separately for each leg (31). We took the maximal CSA of either leg and found the average between these maximal values by adding them and dividing by 2.

Measurement of muscle enzymes and myokines in the blood. We obtained serum samples from all subjects through venous blood sampling, and performed all procedures mentioned below according to the manufacturers' protocols.

In patients and controls, we analyzed serum levels of creatine kinase, lactate dehydrogenase, and aspartate amino transferase in the hospital's routine laboratory. We analyzed circulating levels of interleukin-6 (IL-6), IL-8, IL-15, interferon-γ (IFNγ), IFNγ-inducible protein 10 (IP-10), CCL5, and tumor necrosis factor using Luminex Xmap technology with the Bio-Plex Pro Human Cytokine 27-plex assay (M500KCAF0Y; Bio-Rad). The assay included a highly sensitive standard curve in order to detect very low concentrations of the cytokines. We used enzyme-linked immunosorbent assay kits to measure the levels of decorin (EHDCN; Thermo Scientific) as well as myostatin and monocyte chemotactic protein 1 (DGDF80 and DCP00, respectively; R&D Systems).

Muscle biopsy. We invited patients ages ≥18 years to undergo a percutaneous needle muscle biopsy (a more gentle procedure compared to open biopsy, as it was not intended for clinical purposes [33]) of the left vastus lateralis muscle. With the patient placed in a supine position and given local anesthesia (xylocaine 10 mg/ml + adrenaline 5 μg/ml), we used a 6-mm Pelomi needle with manual suction to obtain the muscle biopsy sample in a sterile procedure. We obtained 30–40 mg muscle tissue for histochemical analysis. Following excision, the muscle biopsy samples were frozen in OCT medium (CellPath) and dispersed in isopentane at freezing point, before storage at –80°C.

Histochemical analysis. In an atmosphere of –20°C, we cut serial 8-μm thick sections of the muscle biopsy tissue using a microtome (CM 1860 UV; Leica), before mounting the sample on microscopic slides (Superfrost Plus; Thermo Scientific). We performed hematoxylin and eosin staining in accordance with a standard protocol, followed by immunohistochemical analysis according to the method of Paulsen et al, using primary antibodies against myosin heavy chain (MHC) type 1, dystrophin, and CD31 to evaluate muscle fibers and capillaries (34), and CD68, tenascin

C, and embryonic MHC to evaluate active inflammation and regeneration (35) (for details, see Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41174/abstract). We determined the distribution of muscle fiber types, fiber CSA, and capillary data using TEMA software (CheckVision). We expressed capillarization as the total number of capillaries per total number of fibers (CF), total number of capillaries around each muscle fiber type (type I and type II) (CAFI and CAFII, respectively), and total number of capillaries around each muscle fiber type related to muscle fiber area (CAFAI and CAFAII, respectively).

Statistical analysis. For statistical analyses, we used IBM SPSS statistical software, version 25. To compare patients and controls, we used a paired-sample *t*-test, Wilcoxon's signed rank test, or McNemar's test, as appropriate. To compare patients with active disease and those with inactive disease, we used an independent-sample *t*-test, Mann-Whitney U test, or chi-square test, as appropriate. Values are presented as the mean ± SD or median with interquartile range (IQR). We performed correlation analyses using Pearson's R or Spearman's rho correlation coefficients, as appropriate, with weak correlation defined as r <

0.3, moderate as r = 0.3–0.69, and strong as r ≥ 0.7. *P* values less than 0.05 were considered statistically significant. We did not correct for multiple comparisons because of the hypothesis-generating nature of our study, nor did we perform statistical analyses of the muscle biopsy results, due to the small number of patients with available muscle biopsy tissue samples.

RESULTS

Patient participation. Of the 72 invited patients, 45 (85%) of the 53 accepting the study invitation fulfilled the inclusion criteria. One patient was later excluded due to a change of diagnosis. We obtained muscle biopsy samples from 17 (46%) of 37 patients ages ≥18 years. All patients but 1 fulfilled the EULAR/ACR classification criteria for juvenile DM, while the remaining patient fulfilled the criteria for IIMs but lacked the classic juvenile DM presentation of rashes.

General characteristics. Among the patients with juvenile DM, 17 (39%) of 44 had active disease (14 [82%] of 17 were age >18 years) and 27 (61%) of 44 had inactive disease (23 [85%] of 27 were age >18 years) (Table 1). Regarding physical activity, patients had less time spent in MVPA than controls (mean 13.2

Table 1. General characteristics, physical activity measures, and disease variables in patients with juvenile dermatomyositis (DM) compared to controls*

	Patients with juvenile DM			
	Active disease (n = 17)	Inactive disease (n = 27)	Total (n = 44)	Controls (n = 44)
General characteristic				
Age, mean ± SD years	32.2 ± 14.6	28.7 ± 10.3	30.1 ± 12.1	30.49 ± 12.1
Female, no. (%)	11 (65)	16 (59)	27 (61)	27 (61)
Height, mean ± SD cm	165.7 ± 13.1	170.2 ± 9.3	168.5 ± 11.0	171.5 ± 10.1
Weight, mean ± SD kg	65.8 ± 19.1	69.2 ± 16.1	67.9 ± 17.0	67.8 ± 14.7
Physical activity				
Level, mean ± SD minutes/day				
Sedentary	551.6 ± 75.9	579.7 ± 67.0	569.2 ± 70.8	571.0 ± 625
LPA	168.2 ± 60.8	171.4 ± 57.7	170.2 ± 58.1	169.2 ± 60.3
MVPA	48.3 ± 28.0†	42.0 ± 22.1	44.4 ± 24.3‡	57.6 ± 21.1
Counts per minute, mean ± SD	389.4 ± 177.2†	335.2 ± 132.6	352.6 ± 149.6‡	438.7 ± 196.4
Disease variable				
Disease duration, mean ± SD years	23.2 ± 13.4	21.0 ± 10.8	21.8 ± 11.8	NA
Taking medication for juvenile DM, no. (%)	4 (24)	6 (22)	10 (23)	NA
DAS muscle (scale 0–20), mean ± SD§	5.2 ± 3.1	3.8 ± 2.0	4.3 ± 2.5	NA
MDI (scale 0–40), median (IQR)§	5.0 (2.0–5.5)	2.0 (1.0–4.0)	3.0 (1.0–5.0)	NA
PhGA score (scale 0–10), median (IQR)§	0.4 (0.0–0.8)¶	0.0 (0.0–0.3)	0.2 (0.0–0.6)	NA
PhGD score (scale 0–10), median (IQR)§	1.0 (0.5–2.2)	0.8 (0.2–2.1)	1.0 (0.2–2.1)	NA
Self-reported physical health				
SF-36 PCS (scale 0–100)#	49.4 (33.2–53.2)†	56.3 (48.9–60.3)	52.4 (46.6–59.0)‡	58.1 (54.3–60.1)
CHAQ/HAQ score >0, no. (%)	9 (53)	9 (33)	18 (41)	NA

* Categorization into active versus inactive disease was based on the Paediatric Rheumatology International Trials Organisation criteria for inactive disease. LPA = light physical activity; MVPA = moderate-to-vigorous physical activity; NA = not applicable; DAS = Disease Activity Score; MDI = Myositis Damage Index; IQR = interquartile range; PhGA = physician global assessment of disease activity; PhGD = physician global assessment of damage; SF-36 PCS = Short-Form 36 physical component summary score; CHAQ/HAQ = childhood/adult Health Assessment Questionnaire.
† *P* < 0.01 versus patients with inactive disease.
‡ *P* < 0.01 versus age- and sex-matched controls.
§ Higher scores indicate more impairment/worse function.
¶ *P* < 0.05 versus patients with inactive disease.
# Lower scores indicate more impairment/worse function.

fewer minutes/day, 95% confidence interval [95% CI] 4.6–21.8; $P = 0.003$). While there was no significant difference in the MVPA between patients with active disease and respective controls, nor between the control groups of active/inactive disease (data not shown), patients with inactive disease had less time spent in MVPA than respective controls (mean 16.4 fewer minutes/day, 95% CI 6.5–26.2; $P = 0.002$). In patients ages <18 years (n = 6), the MVPA was a mean ± SD 58.5 ± 34.3 minutes/day, while in adults ages ≥18 years (n = 37), it was 341.8 ± 131.6 minutes/day.

A DAS score >0 was found in 42 (95%) of 44 patients with juvenile DM. The MDI score of muscle damage in patients was a mean ± SD 3.3 ± 2.4, and 27 (61%) of 44 patients had an MDI global VAS muscle score of >0.2 cm. In the whole patient group, there were no correlations between age and any physical activity variables determined by accelerometer.

Muscle characteristics. Patients with juvenile DM had a lower peak torque compared to controls (mean difference 29 Nm, 95% CI 13–46; $P = 0.001$) (Table 2). The total work of the knee extensors was a median 738J (IQR 565–1,155) in patients compared to a median 1,249J (IQR 815–1,665) in controls ($P < 0.001$). Peak torque and total work were also lower in patients

Table 2. Muscle characteristics in patients with juvenile dermatomyositis (DM) compared to controls*

	Patients with juvenile DM			
	Active disease (n = 16)	Inactive disease (n = 27)	Total (n = 43)	Controls (n = 44)
Objective muscle test variable				
Peak torque, mean ± SD Nm	98.0 ± 37.8†	127.0 ± 35.8	116.2 ± 38.8‡	145.5 ± 46.6
Repetitions, mean ± SD	22.2 ± 8.7	26.2 ± 7.9	24.7 ± 8.3‡	29.5 ± 8.2
Total work, median (IQR) Joules	565 (350–1,032)†	994 (651–1,175)	738 (565–1,155)§	1,249 (815–1,665)
Peak torque/CSA, median (IQR) Nm/cm^2	2.0 (1.6–2.2)	2.1 (2.0–2.2)	2.0 (1.8–2.2)	NA
Total work/CSA, median (IQR) Joules/cm^2	10.0 (8.3–14.3)†	16.5 (12.4–18.9)	14.1 (9.3–18.5)	NA
Clinical variable				
MMT-8 score (scale 0–80), median (IQR)¶	75.0 (73.0–77.0)#	78.0 (77.0–79.0)	77.5 (74.3–79.0)	NA
CMAS score (scale 0–52), median (IQR)¶	48.0 (45.0–51.5)†	50.5 (49.0–52.0)	50.0 (48.0–52.0)	NA
DAS muscle (scale 0–11), median (IQR)**	2.0 (1.0–4.5)	1.0 (1.0–2.0)	1.8 (1.0–2.5)	NA
MDI muscle (scale 0–3), mean ± SD**	1.5 ± 0.7‡	0.7 ± 0.6	1.0 ± 0.7	NA
MDI muscle VAS score (scale 0–10)				
Median (IQR)**	0.6 (0.3–1.6)#	0.2 (0.0–0.4)	0.3 (0.0–1.0)	NA
Score >0.2, no. (%)	13 (77)†	11 (41)	24 (55)	NA
Laboratory parameter				
CK, median (IQR) units/liter	179 (84–266)	108 (67–131)	118 (79–180)	116 (86–182)
LD, mean ± SD units/liter	181 ± 40	160 ± 19	168 ± 30	170 ± 37
ASAT, median (IQR) units/liter	29 (24–31)	24 (22–33)	26 (22–31)	23 (20–29)
Decorin, median (IQR) pg/ml	177 (156–195)	191 (156–252)	182 (156–238)††	157 (135–184)
IP-10, median (IQR) pg/ml	765 (517–1,477)	863 (547–1,175)	799 (531–1,175)††	579 (461–863)
MCP-1, mean ± SD pg/ml	375 ± 213	290 ± 125	327 ± 168	309 ± 110
Myostatin, mean ± SD pg/ml	1,873 ± 720 (n = 12)	1,983 ± 855 (n = 16)	1,844 ± 629 (n = 26)	2,446 ± 908
CCL5, mean ± SD pg/ml	6,224 ± 2,072	6,046 ± 2,058 (n = 26)	6,122 ± 1,822	6,530 ± 1,610
IL-6, median (IQR) pg/ml	0.6 (0.4–1.1) (n = 16)	1.0 (0.5–1.4) (n = 26)	0.9 (0.4–1.4)	1.0 (0.6–1.5)
IL-8, median (IQR) pg/ml	8.8 (4.6–9.7) (n = 16)	6.5 (5.1–9.8) (n = 26)	6.7 (5.0–9.6)	8.3 (5.1–12.1)
TNF, mean ± SD pg/ml	19.9 ± 11.3	20.9 ± 12.3	20.0 ± 10.6	18.9 ± 7.8
MRI muscle				
Edema, no. (%)‡‡	1 (6)	2 (7)	3 (7)	NA
Fatty infiltration, no. (%)	10 (63)	11 (41)	21 (49)	NA
Grade 2	8 (50)	9 (33)	17 (40)	NA
Grade 3	2 (13)	2 (7)	4 (9)	NA
Calcinosis, no. (%)‡‡	1 (6)	3 (11)	4 (9)	NA
Quadriceps CSA, median (IQR) cm^2	48.5 (39.8–58.7)†	56.7 (50.6–62.5)	53.2 (45.0–62.3)	NA

* Peak torque/cross-sectional area (CSA) and total work/CSA represent the peak torque and total work per maximal quadriceps CSA. Fatty infiltration is defined as the presence of pathologic fatty infiltration. IQR = interquartile range; NA = not applicable; MMT-8 = unilateral Manual Muscle Testing in 8 muscle groups; CMAS = Childhood Myositis Assessment Scale; DAS = Disease Activity Score; MDI = Myositis Damage Index (score for muscle damage extent); MDI VAS = MDI visual analog scale (score for muscle damage severity); CK = creatine kinase; LD = lactate dehydrogenase; ASAT = aspartate amino transferase; IP-10 = interferon-γ–inducible protein 10; MCP-1 = monocyte chemotactic protein 1; IL-6 = interleukin-6; TNF = tumor necrosis factor; MRI = magnetic resonance imaging.
† $P < 0.05$ versus patients with inactive disease.
‡ $P < 0.01$ versus age- and sex-matched controls.
§ $P < 0.001$ versus age- and sex-matched controls.
¶ Lower scores indicate more impairment/worse function.
# $P < 0.01$ versus patients with inactive disease.
** Higher scores indicate more impairment/worse function.
†† $P < 0.05$ versus age- and sex-matched controls.
‡‡ Due to the small numbers of patients, data on edema and calcinosis were not statistically analyzed.

with active disease and patients with inactive disease compared to their respective controls (each $P < 0.034$ versus controls), and in patients with active disease compared to those with inactive disease ($P = 0.016$ for peak torque and $P = 0.019$ for total work) (Table 2). When these muscle strength and endurance values were normalized to the quadriceps femoris CSA, only total work/CSA remained significantly lower in patients with active disease compared to those with inactive disease ($P = 0.027$).

In the total patient group, 38 patients (86%) had an MMT-8 score <80, and 27 (61%) had a CMAS score <52; only 1 (2.3%) had an MMT-8 score <64 and a CMAS score <35 (indicating severe muscle impairment). Muscle dysfunction, as measured by the MDI, was found in 9 patients (21%), muscle weakness was found in 31 patients (71%), and muscle atrophy was found in 5 patients (11%). The MMT-8 and CMAS scores (included in the definition of active/inactive disease) were lower and the extent of muscle damage (MDI muscle scores) and severity of muscle damage (MDI VAS muscle scores) were higher in patients with active disease compared to patients with inactive disease (all $P < 0.005$) (Table 2).

The total patient group had higher serum levels of decorin and IP-10 (myokines related to inflammation) compared to controls.

However, there were no significant differences in any of the serum myokine levels between patients with active disease and those with inactive disease (Table 2). Myokine levels were not found to be correlated with the total DAS scores or MDI muscle scores.

Of the MRI-assessed variables in the thigh muscle, none showed a significant difference between patients with active disease and those with inactive disease. Nevertheless, a numerically larger proportion of patients with active disease had pathologic muscle fatty infiltration as compared to patients with inactive disease (Table 2). No patients had fatty infiltration of the muscle of more than 50% (grade 4). The quadriceps CSA was smaller in patients with active disease compared to patients with inactive disease ($P = 0.017$) (Table 2).

Associations between objective muscle test findings and disease variables/physical activity measures. In patients with active disease, peak torque and total work of the knee extensors showed moderate-to-strong correlations with the MMT-8 score, MMT knee extensor component score, and CMAS score (Table 3 and Figure 1). There were no significant correlations between the findings on clinical tests and findings on objective tests

Table 3. Correlations between peak torque or total work of the knee extensors and general, disease-related, and muscle characteristics in patients with DM*

	Peak torque			Total work		
	Active juvenile DM	Inactive juvenile DM	Total	Active juvenile DM	Inactive juvenile DM	Total
General characteristic						
Height†	0.662‡	0.602‡	0.650§	0.735‡	0.518‡	0.678§
Weight†	0.360	0.532‡	0.464‡	0.206	0.394¶	0.412‡
Physical activity						
MVPA†	0.092	0.395¶	0.227	0.066	0.228	0.139
MVPA bouts	0.260	0.203	0.133	0.254	−0.005	0.066
Disease or muscle variable						
Disease duration†	0.272	0.150	0.165	0.387	0.107	0.220
MMT-8	0.770§	0.270	0.521§	0.838§	0.194	0.533§
CMAS	0.574¶	0.140	0.456‡	0.574¶	0.309	0.519§
DAS muscle	−0.667‡	−0.436¶	−0.605§	−0.706‡	−0.385¶	−0.576§
MDI muscle†	−0.144	−0.292	−0.359¶	−0.091	−0.182	−0.228
MRI fatty infiltration	0.315	−0.223	−0.153	0.347	−0.184	−0.090
CSA†	0.560¶	0.469¶	0.450‡	0.609¶	0.770§	0.756§
MCP-1†	0.515	−0.395¶	−0.076	0.309	−0.048	0.063
Myostatin†	0.247	0.428	0.345	0.259	0.341	0.294
Decorin	0.182	−0.416¶	0.019	0.338	−0.138	0.138
IP-10	−0.025	0.020	0.056	−0.071	−0.234	−0.093
CCL5†	0.324	−0.197	−0.021	0.214	−0.141	0.034
IL-6	0.536¶	−0.053	0.203	0.105	0.054	0.184
IL-8	0.096	−0.277	−0.098	0.063	−0.079	−0.007
TNF†	0.466	0.113	0.236	0.014	−0.002	0.043

* Correlations were determined using Pearson's R and Spearman's rho correlation tests. Moderate-to-vigorous physical activity (MVPA) bouts refer to the average time (in minutes) of MVPA spent in bouts lasting 10 minutes each. The myokines presented were selected on the basis of associations seen in the present and previous studies of myokines in patients with juvenile dermatomyositis (DM). MMT-8 = unilateral Manual Muscle Testing in 8 muscle groups; CMAS = Childhood Myositis Assessment Scale; DAS = Disease Activity Score; MDI = Myositis Damage Index; MRI = magnetic resonance imaging; CSA = (quadriceps) cross-sectional area; MCP-1 = monocyte chemotactic protein 1; IP-10 = interferon-γ–inducible protein 10; IL-6 = interleukin-6; TNF = tumor necrosis factor.
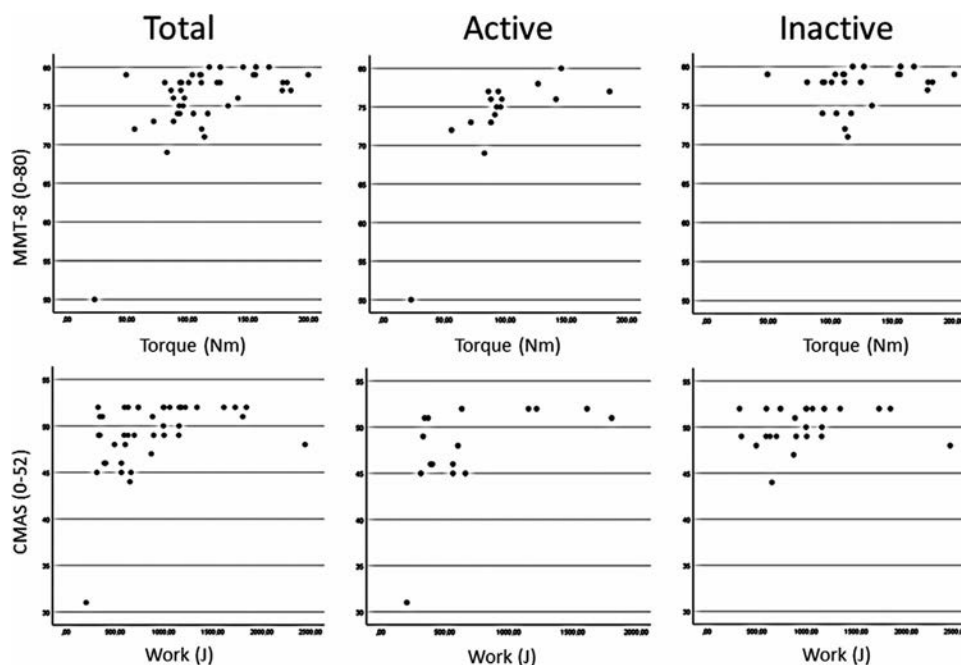† Normally distributed variable.
‡ $P < 0.01$.
§ $P < 0.001$.
¶ $P < 0.05$.

Figure 1. Correlations between objective muscle measures (peak torque of muscle strength and total work of dynamic muscle endurance) and clinical muscle test findings (both general and knee extensor) in patients with juvenile dermatomyositis in total, and in patients with active disease and those with inactive disease. Correlation analyses were performed using Pearson's R or Spearman's rho. MMT-8 = Manual Muscle Testing of 8 muscle groups (unilateral); CMAS = Childhood Myositis Assessment Scale.

of muscle strength or endurance in patients with inactive disease (Table 3 and Figure 1). Peak torque correlated weakly with the MVPA in patients with inactive disease. Both peak torque and total work correlated with DAS muscle scores both in patients with active disease and in patients with inactive disease. In the whole patient group, peak torque correlated negatively with the MDI muscle score. No significant correlations were found between peak torque or total work and MDI muscle scores or MRI findings in the patient subgroups. In patients with active disease, peak torque showed a moderate, positive correlation with serum IL-6 levels (Table 3).

Characteristics of the muscle biopsy tissue. Twelve (71%) of 17 muscle biopsy samples from patients with juvenile DM were of adequate quality for muscle fiber assessment, and 11 (65%) of 17 samples were adequate for capillary assessment. The muscle biopsy samples that were deemed to be of inadequate quality had tissue resembling muscle tissue, but had a texture unsuitable for slicing. Among the patients with adequate-quality muscle biopsy samples, the median time from clinical examination to the needle muscle biopsy was 11.0 months (IQR 9.0–16.3), and none reported major changes in lifestyle or disease activity during this time.

General characteristics of the 12 patients with adequate-quality muscle biopsy samples were not significantly different from the remaining cohort of patients ages ≥18 years (see Supplementary Table 2, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41174/abstract). Seven (58%) of the 12 patients had active disease, and 5 (42%)

had inactive disease, with comparable age and sex distribution between the groups. None of the 12 patients were taking antiinflammatory medications at the time of biopsy. Eight (67%) of the 12 patients had fatty infiltration evident on MRI of the thigh muscle.

Muscle biopsy tissue samples were stained with hematoxylin and eosin, with results showing that 1 patient had increased variability of muscle fiber size and 2 patients had muscle fibers with centralized nuclei (signs of muscle degeneration/regeneration). None of the 12 muscle biopsy samples had pathologic fatty infiltration or definite cell infiltration.

Results of immunohistochemical analyses of the muscle tissue demonstrated that 1 patient with active disease had abnormally large muscle fibers and showed signs of perifascicular atrophy. None of the patients had inflammatory infiltrates (accumulation of CD68+ cells) or the inflammation markers tenascin C or embryonic MHC. There was a numeric trend toward patients with active disease having a larger area of type I muscle fibers compared to type II muscle fibers (Table 4). In addition, patients with active disease tended to have a larger area of type I muscle fibers compared to patients with inactive disease. This trend was reversed in patients with inactive disease, as they had a larger area of type II fibers compared to type I fibers, and had larger type II fibers compared to patients with active disease. There was also a numerically smaller proportion of type I fibers in patients with active compared to inactive disease (Figure 2). There were no significant differences in capillary features between the active disease and inactive disease groups.

Table 4. Muscle biopsy results in patients with juvenile DM*

	Active juvenile DM (n = 7)	Inactive juvenile DM (n = 5)	Total (n = 12)
Type I fiber area, μm^2			
Total	4,497 (2,701–8,560)	4,045 (2,643–4,576)	4,212 (3,233–4,600)
Men	4,608 (3,143–8,560) (n = 3)	4,311 (4,045–4,576) (n = 2)	4,576 (3,143–8,560)
Women	4,098 (2,701–5,387) (n = 4)	3,503 (2,643–4,379) (n = 3)	3,699 (2,643–5,387)
Type II fiber area, μm^2			
Total	4,030 (1,951–12,275)	5,081 (2,319–5,873)	4,255 (2,939–5,434)
Men	4,720 (2,922–12,275) (n = 3)	5,713 (5,552–5,873) (n = 2)	5,552 (2,922–12,275)
Women	3,611 (1,951–4,479) (n = 4)	2,990 (2,319–5,081) (n = 3)	3,191 (1,951–5,081)
Type I fibers/total number of fibers, %	39 (33–64)	47 (33–52)	43 (33–53)
Capillarization†			
CF	1.7 (1.1–1.9)	1.8 (1.5–2.0)	1.7 (1.4–2.0)
CAFI	4.3 (3.3–4.7)	4.0 (3.7–4.4)	4.2 (3.7–4.5)
CAFII	3.6 (3.4–3.9)	3.7 (3.0–4.5)	3.6 (3.3–4.1)
CAFAI	0.9 (0.8–1.2)	0.9 (0.8–1.2)	0.9 (0.8–1.2)
CAFAII	1.0 (0.7–1.3)	0.9 (0.6–1.1)	1.0 (0.7–1.3)

* No statistical analyses of the muscle biopsy data were performed due to the small number of needle biopsy samples available. Type I and type II fibers refer to muscle fiber types. Values are the median (interquartile range).
† Data on capillarization are expressed as the total number of capillaries per total number of fibers (CF), total number of capillaries around type I and type II muscle fibers (CAFI and CAFII, respectively), and total number of capillaries around type I and type II muscle fibers per fiber area (CAFAI and CAFAII, respectively). The number of samples for capillary data were as follows: active juvenile dermatomyositis (DM) n = 7, inactive juvenile DM n = 4, total patient group n = 11.

DISCUSSION

In our study focusing on long-term skeletal muscle outcomes in patients with juvenile DM, we found lower muscle strength and lower muscle endurance in the knee extensors of patients compared to age- and sex-matched controls. Moreover, patients with active disease had lower muscle strength and endurance compared to patients with inactive disease. Corrected for muscle size (the quadriceps CSA), only muscle endurance remained
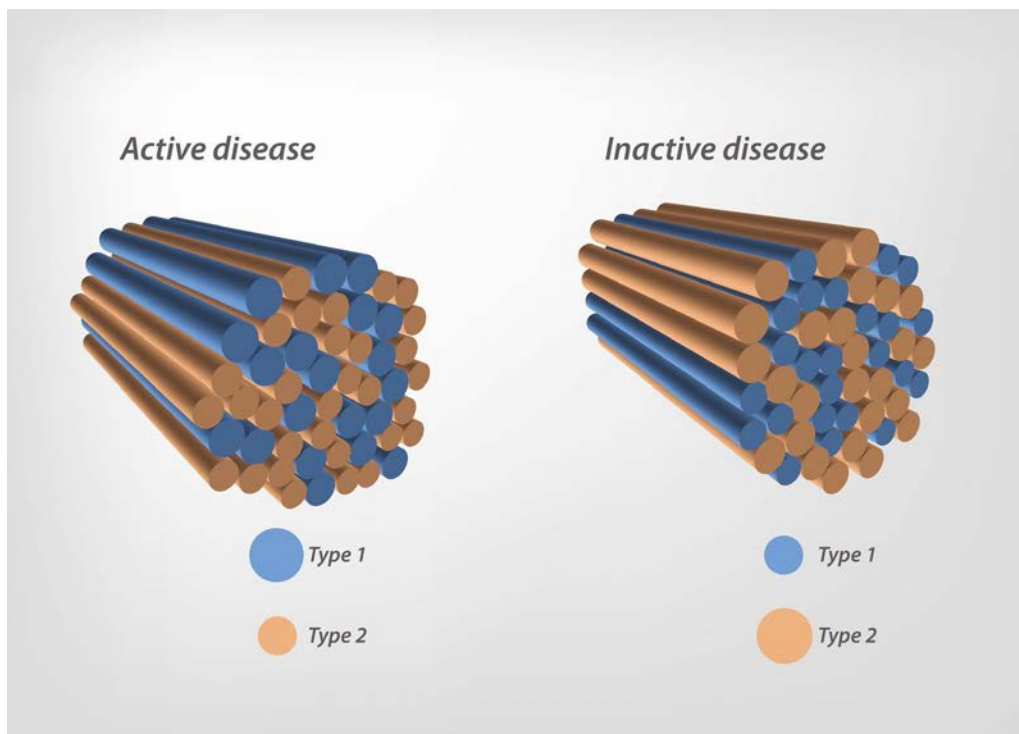
Figure 2. Visual images of the muscle fiber composition (types I and II muscle fibers) and relative size of the muscle fibers in patients with active juvenile dermatomyositis (DM) and those with inactive juvenile DM.

significantly lower in patients with active disease compared to patients with inactive disease. Clinically assessed muscle damage was higher in patients with active disease compared to those with inactive disease. The results of objective muscle tests of the knee extensors correlated with the findings from clinical tests of muscle strength and endurance only in patients with active disease. Muscle biopsy results indicated that capillary density was similar but muscle fiber composition was different between the active and inactive disease groups. To our knowledge, this is the first long-term study to assess functional, laboratory, serologic, radiographic, and histologic muscle outcomes simultaneously in patients with juvenile DM.

Our patients were older, had a longer disease duration, and were comparable in sex distribution when compared to patients in other studies of juvenile DM outcomes (7,8). Compared to those other studies (7,8), more patients in our study had a DAS >0, and our study had a comparable proportion of patients with an MDI global VAS muscle score >0.2; however, the MDI scores were higher in our patients. The physical activity levels of our patients were similar to those in a Danish juvenile DM cohort and higher than those in a Brazilian juvenile DM cohort (36,37). Based on our findings, the patients in our study presented with fairly high levels of physical activity despite frequently having relatively high levels of disease activity and damage.

The control subjects were randomly selected from the Norwegian National Registry, which is a strength of our study. They were age- and sex-matched to our patients in order to account for the large age dispersion and to exclude the potential for age-related confounders in our main results. The physical activity levels of our adult control subjects resembled those of the general Norwegian adult population (38), supporting the representativeness of our control group.

Isometric muscle strength and muscle endurance of the knee extensors as well as MMT-8 scores and CMAS scores were lower in patients with active disease compared to patients with inactive disease, but the findings of these clinical and objective muscle tests correlated only in patients with active disease. Although objective muscle tests were only assessed in the knee extensors, peak torque also correlated with the MMT knee component scores only in patients with active disease. Taken together, these results support the ceiling effects of the MMT-8 and CMAS scoring systems. A precise scoring of mild muscle weakness and dysfunction, especially in patients with inactive disease, may therefore require more objective and sensitive muscle testing.

The MMT-8 and CMAS scores were mildly reduced (median MMT-8 score 77.5 and median CMAS 50.0), similar to that in a Danish study of juvenile DM patients assessed after a disease duration of 13.9 years, in which the mean values were 78.0 and 48.8, respectively (8). An MMT-8 score <80 and CMAS score <52 were, however, more frequent in our study (86% and 61%, respectively) than in a multinational outcome study of 490 juvenile DM patients whose mean disease duration was 7.7 years (41% and 53%,

respectively) (7). Nevertheless, severe muscle weakness/dysfunction was rare; only 1 patient (2.3%) had serious muscle weakness and dysfunction based on having an MMT-8 score <64 and CMAS score <35, as compared to 7% having an MMT-8 score <64 and 8% having a CMAS score <35 in the multinational study. A longer disease duration in our study could be the reason for this difference. Even though the CMAS has been used in mixed pediatric/adult juvenile DM populations (7–9), it has not been validated for adults with juvenile DM. However, our group has shown moderate correlations between the CMAS score and disease measures (MMT-8, DAS muscle scores, and the MDI) in patients ages >18 years (9), supporting the use of the tool in this adult age group.

Objective muscle strength and muscular endurance were lower in patients (both in those with active disease and in those with inactive disease) compared to controls. Multiple factors may have contributed to these results, including exercise habits and disease-related features (39) (age/sex effects were controlled for by matching). We did not systematically collect data on strength-training habits, a limitation to our study. However, only patients with inactive disease had a lower MVPA compared to controls, and there was a correlation between the MVPA and mean peak torque in patients with inactive disease. This suggests that deconditioning may play a larger role in explaining the lower muscle strength in these patients compared to patients with active disease, although muscle disease activity and damage were also present in this patient group.

Patients with active disease had higher muscle damage (based on MDI muscle and MDI VAS muscle scores, as well as numeric values of MRI-detected damage, including fatty infiltration), but not muscle activity (based on DAS muscle scores, muscle enzyme levels, or MRI-detected edema), compared to patients with inactive disease. However, muscle damage scores did not correlate with isometric muscle strength or muscle endurance in either patient group. This suggests that the difference in peak torque and total work between active and inactive juvenile DM may be attributed to disease-related factors other than the classic juvenile DM measures of disease damage.

Muscle CSA was lower in patients with active disease compared to patients with inactive disease; this might represent a larger reduction in volume due to muscle atrophy in patients with active disease. The CSA correlated with findings on objective muscle tests in both patient groups. However, when correcting the measurements of muscle strength and muscle endurance for the quadriceps CSA, only total work remained lower in patients with active disease compared to those with inactive disease. This could indicate that muscle endurance associated with active disease is influenced by structural or functional differences within the muscle tissue independent of muscle size. This hypothesis was supported by the muscle biopsy results. We found no signs of muscle inflammation in the biopsy samples. However, although not statistically tested, there were numeric trends of different muscle fiber composition between patients with active disease and

those with inactive disease. In patients with active disease, type I muscle fibers were relatively larger and type II muscle fibers were relatively smaller compared to patients with inactive disease, and patients with active disease had a lower proportion of type I fibers. Given the numerically higher MVPA in patients with active disease, this difference in fiber composition was unexpected, as exercise is found to increase the size of type II fibers and the proportion of type I fibers in adult patients with DM (21). However, there is evidence that long-term physical inactivity or chronic disease can cause a greater percentage of type II fibers (40).

Hypoxia has been suggested as a possible mechanism for the reduction in muscle endurance in patients with juvenile DM, attributable to the fact that blood flow increases more poorly in response to exercise (41). In patients with severe chronic obstructive pulmonary disease, which leads to chronic hypoxia, the proportion of type I muscle fibers has been found to be decreased, together with an increase in the type II muscle fiber area (42), similar to the findings in our patients with active disease. With regard to capillarization, we found that the histologic capillary density was similar to that in studies of healthy populations (43), and there was no difference in capillary features between patients with active disease and those with inactive disease. However, the presented biopsy results do not tell us anything about the potential for functional impairment of the capillaries in juvenile DM (41).

Myokines are known to be secreted from muscle tissue in response to exercise or inflammatory stimuli (6). We found higher levels of IP-10 in patients compared to controls, similar to previous data from our own cohort (17), supporting the notion that this myokine is up-regulated even after long-term disease. IP-10 was recently validated as a strong, reliable, and sensitive biomarker for active juvenile DM (44). We did not, however, find significant differences in IP-10 levels between the active and inactive disease groups, suggesting that our study may be underpowered, or that the myokine may be a less stable marker of disease activity over a longer disease duration. In addition, our patients had higher circulating levels of decorin compared to controls. Decorin is known to be both antifibrotic and proinflammatory (45,46). Higher levels of decorin may be attributable to increased levels of visceral fat (VAT) depots, as has been previously described (47) (VAT is a greater source of decorin than subcutaneous fat [46]), and may reflect the inflammatory state of juvenile DM. Surprisingly, we found a positive association between the serum IL-6 levels and the peak torque in patients with active disease. IL-6 is known as a biomarker for active inflammation in juvenile DM (48). Interestingly, it was also recently found to be both expressed and secreted from type I muscle fibers in mice (49), while torque was associated with type I fibers in female athletes (50). Thus, we could speculate that the association between IL-6 levels and torque is related to an increase in the area of type I muscle fibers in patients with active disease.

In addition to the limitations to our study already mentioned, the time delay between muscle biopsies and other examinations may have affected the interpretation of the biopsy results, although none of the patients who underwent a muscle biopsy reported experiencing lifestyle changes, including physical activity habits. The small number of patients in each group (active and inactive disease) may have created Type II errors in the statistical analyses. For the lowest numbers, therefore, we chose not to perform statistical analyses, but rather we described numeric differences. We also isolated the muscle subscores of the validated tools DAS and MDI, and the knee extensor component of MMT-8, although these subscores have not been validated separately.

In conclusion, after an average disease duration of almost 22 years, objectively measured muscle strength and muscle endurance of the knee extensors were lower in patients with juvenile DM compared to controls, and in patients with active disease compared to those with inactive disease. The results of objective muscle tests and clinical muscle tests correlated only in patients with active disease, suggesting the need for more objective and sensitive muscle tests in this clinical setting. Based on the results of the present study, we can hypothesize that impaired muscle endurance of the knee extensors in patients with active disease may be influenced by structural and functional adaptions of muscle tissue independent of muscle size. Further study of these concepts should be carried out in patients with juvenile DM.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Berntsen 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. Bernsten, Raastad, Sjaastad, Sanner.
Acquisition of data. Bernsten, Raastad, Marstein, Kirkhus, Merckoll, Sanner.
Analysis and interpretation of data. Bernsten, Raastad, Marstein, Kirkhus, Merckoll, Cumming, Flatø, Sjaastad, Sanner.

REFERENCES

1. Feldman BM, Rider LG, Reed AM, Pachman LM. Juvenile dermatomyositis and other idiopathic inflammatory myopathies of childhood. Lancet 2008;371:2201–12.

2. Harris-Love MO, Shrader JA, Koziol D, Pahlajani N, Jain M, Smith M, et al. Distribution and severity of weakness among patients with

polymyositis, dermatomyositis and juvenile dermatomyositis. Rheumatology (Oxford) 2009;48:134–9.

3. Rayavarapu S, Coley W, Kinder TB, Nagaraju K. Idiopathic inflammatory myopathies: pathogenic mechanisms of muscle weakness [review]. Skelet Muscle 2013;3:13.

4. Wedderburn LR, Rider LG. Juvenile dermatomyositis: new developments in pathogenesis, assessment and treatment. Best Pract Res Clin Rheumatol 2009;23:665–78.

5. Park JH, Niermann KJ, Ryder NM, Nelson AE, Das A, Lawton AR, et al. Muscle abnormalities in juvenile dermatomyositis patients: P-31 magnetic resonance spectroscopy studies. Arthritis Rheum 2000;43:2359–67.

6. Lightfoot AP, Cooper RG. The role of myokines in muscle health and disease. Curr Opin Rheumatol 2016;28:661–6.

7. Ravelli A, Trail L, Ferrari C, Ruperto N, Pistorio A, Pilkington C, et al. Long-term outcome and prognostic factors of juvenile dermatomyositis: a multinational, multicenter study of 490 patients. Arthritis Care Res (Hoboken) 2010;62:63–72.

8. Mathiesen P, Hegaard H, Herlin T, Zak M, Pedersen FK, Nielsen S. Long-term outcome in patients with juvenile dermatomyositis: a cross-sectional follow-up study. Scand J Rheumatol 2012;41:50–8.

9. Sanner H, Kirkhus E, Merckoll E, Tollisen A, Røisland M, Lie BA, et al. Long-term muscular outcome and predisposing and prognostic factors in juvenile dermatomyositis: a case–control study. Arthritis Care Res (Hoboken) 2010;62:1103–11.

10. Sanner H, Sjaastad I, Flatø B. Disease activity and prognostic factors in juvenile dermatomyositis: a long-term follow-up study applying the Paediatric Rheumatology International Trials Organization criteria for inactive disease and the myositis disease activity assessment tool. Rheumatology (Oxford) 2014;53:1578–85.

11. Rider LG, Werth VP, Huber AM, Alexanderson H, Rao AP, Ruperto N, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: physician and patient/parent global activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBM-FRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S118–57.

12. Andersen H, Jakobsen J. A comparative study of isokinetic dynamometry and manual muscle testing of ankle dorsal and plantar flexors and knee extensors and flexors. Eur Neurol 1997;37:239–42.

13. Malattia C, Damasio MB, Madeo A, Pistorio A, Providenti A, Pederzoli S, et al. Whole-body MRI in the assessment of disease activity in juvenile dermatomyositis. Ann Rheum Dis 2014;73:1083–90.

14. Wedderburn LR, Varsani H, Li CK, Newton KR, Amato AA, Banwell B, et al. International consensus on a proposed score system for muscle biopsy evaluation in patients with juvenile dermatomyositis: a tool for potential use in clinical trials. Arthritis Rheum 2007;57:1192–201.

15. Davis WR, Halls JE, Offiah AC, Pilkington C, Owens CM, Rosendahl K. Assessment of active inflammation in juvenile dermatomyositis: a novel magnetic resonance imaging-based scoring system. Rheumatology (Oxford) 2011;50:2237–44.

16. Barth Z, Witczak BN, Flatø B, Koller A, Sjaastad I, Sanner H. Assessment of microvascular abnormalities by nailfold capillaroscopy in juvenile dermatomyositis after medium- to long-term followup. Arthritis Care Res (Hoboken) 2018;70:768–76.

17. Sanner H, Schwartz T, Flatø B, Vistnes M, Christensen G, Sjaastad I. Increased levels of eotaxin and MCP-1 in juvenile dermatomyositis median 16.8 years after disease onset; associations with disease activity, duration and organ damage. PLoS One 2014;9:e92171.

18. Loell I, Helmers SB, Dastmalchi M, Alexanderson H, Munters LA, Nennesmo I, et al. Higher proportion of fast-twitch (type II) muscle fibres in idiopathic inflammatory myopathies: evident in chronic but not in untreated newly diagnosed patients. Clin Physiol Funct Imaging 2011;31:18–25.

19. Suter E, Hoppeler H, Claassen H, Billeter R, Aebi U, Horber F, et al. Ultrastructural modification of human skeletal muscle tissue with 6-month moderate-intensity exercise training. Int J Sports Med 1995;16:160–6.

20. Staron RS, Hagerman FC, Hikida RS, Murray TF, Hostler DP, Crill MT, et al. Fiber type composition of the vastus lateralis muscle of young men and women. J Histochem Cytochem 2000;48:623–9.

21. Dastmalchi M, Alexanderson H, Loell I, Ståhlberg M, Borg K, Lundberg IE, et al. Effect of physical training on the proportion of slow-twitch type I muscle fibers, a novel nonimmune-mediated mechanism for muscle impairment in polymyositis or dermatomyositis. Arthritis Rheum 2007;57:1303–10.

22. Sanner H, Gran JT, Sjaastad I, Flatø B. Cumulative organ damage and prognostic factors in juvenile dermatomyositis: a cross-sectional study median 16.8 years after symptom onset. Rheumatology (Oxford) 2009;48:1541–7.

23. Bohan A, Peter JB. Polymyositis and dermatomyositis: first of two parts. N Engl J Med 1975;292:344–7.

24. Lundberg IE, Tjärnlund A, Bottai M, Werth VP, Pilkington C, de Visser M, et al, and the International Myositis Classification Criteria Project Consortium, the Euromyositis Register, and the Juvenile Dermatomyositis Cohort Biomarker Study and Repository (UK and Ireland). 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Arthritis Rheumatol 2017;69:2271–82.

25. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310:2191–4.

26. Bode RK, Klein-Gitelman MS, Miller ML, Lechman TS, Pachman LM. Disease activity score for children with juvenile dermatomyositis: reliability and validity evidence. Arthritis Rheum 2003;49:7–15.

27. Rider LG, Lachenbruch PA, Monroe JB, Ravelli A, Cabalar I, Feldman BM, et al. Damage extent and predictors in adult and juvenile dermatomyositis and polymyositis as determined with the Myositis Damage Index. Arthritis Rheum 2009;60:3425–35.

28. Lazarevic D, Pistorio A, Palmisani E, Miettunen P, Ravelli A, Pilkington C, et al, for the Paediatric Rheumatology International Trials Organisation (PRINTO). The PRINTO criteria for clinically inactive disease in juvenile dermatomyositis. Ann Rheum Dis 2013;72:686–93.

29. Berntsen KS, Edvardsen E, Hansen BH, Flatø B, Sjaastad I, Sanner H. Cardiorespiratory fitness in long-term juvenile dermatomyositis: a controlled, cross-sectional study of active/inactive disease. Rheumatology (Oxford) 2019;58:492–501.

30. Dalene KE, Anderssen SA, Andersen LB, Steene-Johannessen J, Ekelund U, Hansen BH, et al. Secular and longitudinal physical activity changes in population-based samples of children and adolescents. Scand J Med Sci Sports 2018;28:161–71.

31. Røren Nordén K, Dagfinrud H, Løvstad A, Raastad T. Reduced appendicular lean body mass, muscle strength, and size of type II muscle fibers in patients with spondyloarthritis versus healthy controls: a cross-sectional study. ScientificWorldJournal 2016;2016:6507692.

32. Goutallier D, Postel JM, Gleyze P, Leguilloux P, Van Driessche S. Influence of cuff muscle fatty degeneration on anatomic and functional

outcomes after simple suture of full-thickness tears. J Shoulder Elbow Surg 2003;12:550–4.

33. Dietrichson P, Coakley J, Smith PE, Griffiths RD, Helliwell TR, Edwards RH. Conchotome and needle percutaneous biopsy of skeletal muscle. J Neurol Neurosurg Psychiatry 1987;50:1461–7.

34. Paulsen G, Cumming KT, Holden G, Hallén J, Rønnestad BR, Sveen O, et al. Vitamin C and E supplementation hampers cellular adaptation to endurance training in humans: a double-blind, randomised, controlled trial. J Physiol 2014;592:1887–901.

35. Baumann M, Gumpold C, Mueller-Felber W, Schoser B, Haberler C, Loescher WN, et al. Pattern of myogenesis and vascular repair in early and advanced lesions of juvenile dermatomyositis. Neuromuscul Disord 2018;28:973–85.

36. Pinto AJ, Yazigi Solis M, de Sá Pinto AL, Silva CA, Maluf Elias Sallum A, Roschel H, et al. Physical (in)activity and its influence on disease-related features, physical capacity, and health-related quality of life in a cohort of chronic juvenile dermatomyositis patients. Semin Arthritis Rheum 2016;46:64–70.

37. Mathiesen PR, Ørngreen MC, Vissing J, Andersen LB, Herlin T, Nielsen S. Aerobic fitness after JDM: a long-term follow-up study. Rheumatology (Oxford) 2013;52:287–95.

38. Hansen BH, Anderssen SA, Steene-Johannessen J, Ekelund U, Nilsen AK, Andersen ID, et al. Fysisk aktivitet og sedat tid blant voksne og eldre i Norge : nasjonal kartlegging 2014–2015. Oslo: Helse-direktoratet; 2015.

39. Leblanc A, Pescatello LS, Taylor BA, Capizzi JA, Clarkson PM, Michael White C, et al. Relationships between physical activity and muscular strength among healthy adults across the lifespan. Springerplus 2015;4:557.

40. Wasserman K, Hansen JE, Sue DY, Stringer WW, Sietsema KE, Xing-Guo S, et al. Principles of exercise testing and interpretation: including pathophysiology and clinical applications. 5th ed. Philadelphia: Lippincott Williams & Wilkins, a Wolters Kluwer Business; 2012.

41. Habers GE, De Knikker R, Van Brussel M, Hulzebos E, Stegeman DF, Van Royen A, et al. Near-infrared spectroscopy during exercise

and recovery in children with juvenile dermatomyositis. Muscle Nerve 2013;47:108–15.

42. Eliason G, Abdel-Halim S, Arvidsson B, Kadi F, Piehl-Aulin K. Physical performance and muscular characteristics in different stages of COPD. Scand J Med Sci Sports 2009;19:865–70.

43. Toft I, Lindal S, Bønaa KH, Jenssen T. Quantitative measurement of muscle fiber composition in a normal population. Muscle Nerve 2003;28:101–8.

44. Wienke J, Bellutti Enders F, Lim J, Mertens JS, van den Hoogen LL, Wijngaarde CA, et al. Galectin-9 and CXCL10 as biomarkers for disease activity in juvenile dermatomyositis: a longitudinal cohort study and multicohort validation. Arthritis Rheumatol 2019;71:1377–90.

45. Zhu J, Li Y, Shen W, Qiao C, Ambrosio F, Lavasani M, et al. Relationships between transforming growth factor-β1, myostatin, and decorin: implications for skeletal muscle fibrosis. J Biol Chem 2007;282:25852–63.

46. Bolton K, Segal D, McMillan J, Jowett J, Heilbronn L, Abberton K, et al. Decorin is a secreted protein associated with obesity and type 2 diabetes. Int J Obes (Lond) 2008;32:1113–21.

47. Witczak BN, Godang K, Schwartz T, Olarescu NC, Flatø B, Bollerslev J, et al. Body composition and adipose tissue distribution in juvenile dermatomyositis and associations with cardiac function [abstract]. Arthritis Rheumatol 2016;68 Suppl 10. URL: https://acrabstracts.org/abstract/body-composition-and-adipose-tissue-distribution-in-juvenile-dermatomyositis-and-associations-with-cardiac-function/.

48. Reed AM, Peterson E, Bilgic H, Ytterberg SR, Amin S, Hein MS, et al. Changes in novel biomarkers of disease activity in juvenile and adult dermatomyositis are sensitive biomarkers of disease course. Arthritis Rheum 2012;64:4078–86.

49. Liang AP, Drazick AT, Gao H, Li Y. Skeletal muscle secretion of IL-6 is muscle type specific: ex vivo evidence. Biochem Biophys Res Commun 2018;505:146–50.

50. Gregor RJ, Edgerton VR, Perrine JJ, Campion DS, DeBus C. Torque-velocity relationships and muscle fiber composition in elite female athletes. J Appl Physiol Respir Environ Exerc Physiol 1979;47:388–92.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Systemic Autoimmune Disease Among Adults Exposed to the September 11, 2001 Terrorist Attack

Sara A. Miller-Archie,[1] 🔾 Peter M. Izmirly,[2] Jessica R. Berman,[3] Jennifer Brite,[1] Deborah J. Walker,[1]
Renato C. Dasilva,[1] Lysa J. Petrsoric,[1] and James E. Cone[1]

Objective. Autoimmune disease is an emerging condition among persons exposed to the September 11, 2001 attack on the World Trade Center (WTC). Components of the dust cloud resulting from the collapse of the WTC have been associated with development of a systemic autoimmune disease, as has posttraumatic stress disorder (PTSD). We undertook this study to determine whether dust exposure and PTSD were associated with an increased risk of systemic autoimmune disease in a 9/11-exposed cohort.

Methods. Among 43,133 WTC Health Registry enrollees, 2,786 self-reported having a post-9/11 systemic autoimmune disease. We obtained informed consent to review medical records to validate systemic autoimmune disease diagnoses for 1,041 enrollees. Diagnoses of systemic autoimmune diseases were confirmed by classification criteria, rheumatologist diagnosis, or having been prescribed systemic autoimmune disease medication. Controls were enrollees who denied having an autoimmune disease diagnosis (n = 37,017). We used multivariable log-binomial regression to examine the association between multiple 9/11 exposures and risk of post-9/11 systemic autoimmune disease, stratifying by responders (rescue, recovery, and clean-up workers) and community members (e.g., residents, area workers).

Results. We identified 118 persons with systemic autoimmune disease. Rheumatoid arthritis was most frequent (n = 71), followed by Sjögren's syndrome (n = 22), systemic lupus erythematosus (n = 20), myositis (n = 9), mixed connective tissue disease (n = 7), and scleroderma (n = 4). Among 9/11 responders, those with intense dust cloud exposure had almost twice the risk of systemic autoimmune disease (adjusted risk ratio 1.86 [95% confidence interval 1.02–3.40]). Community members with PTSD had a nearly 3-fold increased risk of systemic autoimmune disease.

Conclusion. Intense dust cloud exposure among responders and PTSD among community members were associated with a statistically significant increased risk of new-onset systemic autoimmune disease. Clinicians treating 9/11 survivors should be aware of the potential increased risk of systemic autoimmune disease in this population.

INTRODUCTION

Systemic autoimmune diseases are an emerging health concern among individuals who were exposed to the September 11, 2001 terrorist attack on the World Trade Center (WTC) in New York City. The collapse of the WTC towers after the attack resulted in the release of a cloud of dust and debris that covered parts of Lower Manhattan, New York. Further exposure occurred subsequently through the resuspension of dust particles during recovery and clean-up activities (1). Crystalline silica, a known risk factor for autoimmune diseases (2–8), was a major component of the dust cloud (1,9). Additional components of the dust cloud and the air at the site during the clean-up period (1,10) have been previously associated with autoimmune diseases, including organic hydrocarbon solvents (11–13), fine particulate matter (PM2.5) (14–16), and asbestos (8,17).

James E. Cone, MD, MPH: New York City Department of Health and Mental Hygiene, New York, New York; [2]Peter M. Izmirly, MD: New York University School of Medicine, New York, New York; [3]Jessica R. Berman, MD: Hospital for Special Surgery and Weill Cornell Medical College, New York, New York.

In addition to the physical exposures, many people witnessed or experienced traumatic events on 9/11 as well as continued occupational or personal reminders of the attacks. Posttraumatic stress disorder (PTSD), one of the most common post-9/11 mental health disorders (18), has been associated with the subsequent onset of rheumatoid arthritis (RA) and other autoimmune disorders in both veteran and civilian populations (19–21).

Studies conducted among New York City firefighters and emergency medical services personnel have identified an exposure-response relationship between length of time worked at the WTC site, level of exposure to the dust and debris following the attacks, and systemic autoimmune disease (22,23). Currently, there have been no published studies on autoimmune diseases among community members who were exposed to the attack or among responders and clean-up workers other than those with the Fire Department of the City of New York (FDNY).

This study aimed to identify systemic autoimmune diseases among a cohort of 9/11-exposed adults. We also sought to determine whether high levels of 9/11 dust exposure or PTSD were associated with an increased risk of systemic autoimmune disease and whether this association differed between 9/11 responders and community members.

SUBJECTS AND METHODS

WTC Health Registry. The WTC Health Registry (referred to as the Registry) is a longitudinal, prospective cohort of 71,426 enrollees who were exposed to the events of September 11, 2001 in New York City, or who were involved in the subsequent recovery and clean-up effort, which lasted until July 2002. The Registry comprises individuals who were part of the rescue, recovery, and clean-up response (responders) (43%) and those who worked, resided, attended school, or were in transit in Lower Manhattan the morning of the attack (community members) (57%), and has been estimated to include 17% of eligible exposed individuals (24,25). Approximately 10% of responders in the Registry were FDNY firefighters or emergency medical services personnel on 9/11 (25). Potential enrollees were identified and recruited using employer, organizational, and building occupant lists and an outreach and advertising campaign (18). A more detailed description of the Registry's recruitment methods has been published elsewhere (18). Enrollment occurred in 2003–2004 via telephone interviews (95%) or in-person interviews (5%) (wave 1), at which time interviewers collected demographic, exposure, and physical and mental health data. All enrollees provided verbal informed consent at Registry enrollment. There have been subsequent Registry-wide follow-up surveys, including wave 2 (2006–2007, 68% response rate) and wave 3 (2011–2012, 63% response rate), which was the first Registry survey that included questions on autoimmune diseases. The study was approved by the Institutional Review Board of the New York City Department of Health and Mental Hygiene.

Autoimmune disease study participants. Eligible enrollees included all those age ≥18 years who reported on the wave 3 survey that they had ever been told by a physician or other health professional that they had RA or "(an)other autoimmune disorders (e.g., lupus, scleroderma, polymyositis)." Participants were eligible if they indicated that the year they were first told was 2001 or later. We excluded 8 enrollees who died between their wave 3 survey and the launch of the autoimmune disease study in 2014, leaving 2,786 potential cases of post-9/11 autoimmune disease for further examination. A detailed flow chart of the study population is shown in Figure 1.

Phase I (autoimmune disease enrollee survey). An in-depth autoimmune disease follow-up survey or survey link (enrollee survey) was mailed or e-mailed to all eligible Registry enrollees in May 2014. The enrollee survey asked participants whether they had been told by a clinician that they had RA or other type of arthritis, systemic lupus erythematosus (SLE), scleroderma, polymyositis (PM)/dermatomyositis (DM), mixed connective tissue disease (MCTD), Sjögren's syndrome (SS), or other autoimmune disease, and whether they were taking medication for their condition. The survey included the CTD Screening Questionnaire (CSQ) (26,27), which is a screening tool for the systemic autoimmune diseases listed above, and includes questions on symptoms and blood tests. Additional questions covered topics including general health, symptoms of other systemic autoimmune conditions, and family history of autoimmune disease. We created versions of the paper survey in English, Spanish, and Chinese; the online version was only available in English.

Enrollees who did not respond to the initial e-mail invitation or mailed paper survey were sent up to 1–2 reminder postcards, 6 reminder e-mails, and 2 additional mailings. Registry and New York City Department of Health and Mental Hygiene call center staff then telephoned enrollees who had not yet completed the survey, with up to 10 phone call attempts and 2 telephone messages left on their home or mobile phone.

Phase II (consent for release of medical records and physician survey). Most enrollee survey respondents had some indication of possible post-9/11 autoimmune disease, either because they screened positive via the CSQ, they reported an autoimmune disease diagnosis, or they reported taking a medication consistent with treatment for autoimmune disease (data not shown). Therefore, all enrollees who completed the survey (n = 2,042; response rate 73%) were sent a letter with a consent form authorizing the release of their medical records to the Registry. The consent form requested permission to obtain copies of enrollee medical records and to send a survey (physician survey) to their rheumatologist or to the physician who could provide information regarding their autoimmune disease diagnosis. Letters and consent forms were written in English, Spanish, and Chinese. Of those registrants who completed the enrollee survey,
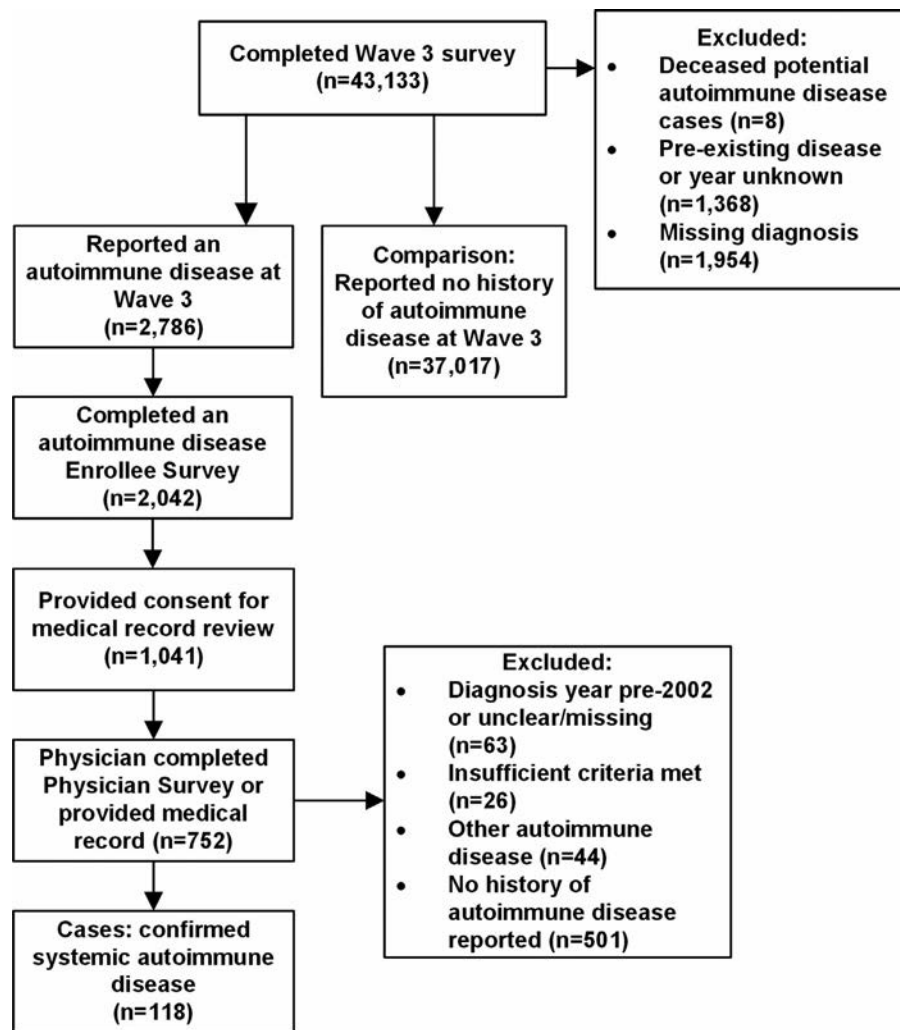
Figure 1. Flow chart of the World Trade Center Health Registry systemic autoimmune disease study population. Systemic autoimmune diseases (rheumatoid arthritis, Sjögren's syndrome, systemic lupus erythematosus, polymyositis/dermatomyositis, mixed connective tissue disease, and scleroderma) were diagnosed in 2002 or later. Cases had to either meet disease-specific classification criteria, have been diagnosed by a board-certified rheumatologist, or have been prescribed a medication commonly used to treat systemic autoimmune disease.

1,041 (51%) provided the Registry with their informed consent. We used methods similar to those used for the enrollee survey to increase participation.

The physician survey, first mailed in May 2015, included checklists adapted from the American College of Rheumatology classification criteria for RA (28), SLE (29), and scleroderma (30); Bohan and Peter's criteria for PM/DM (31,32); Kahn's criteria for MCTD (33); and American–European Consensus Group criteria for SS (34). The survey also queried the enrollee's year of diagnosis for each condition, any medications prescribed to treat the autoimmune disease, and whether the physician was a board-certified rheumatologist.

In response to feedback received from physicians, we also offered the option of having medical records sent directly to the Registry for abstraction. Physicians who did not respond to the initial request were sent additional requests via mail, fax, and office telephone. Medical records retrieval contractors then followed up with physicians who had not responded to multiple requests from the Registry. Of the registrants who submitted consent forms, ~10% listed more than one physician, and attempts were made to contact multiple physicians when listed. Our report is presented based on the medical records and physician surveys obtained as of November 2017.

Cases. Documentation of an autoimmune disease and the year of diagnosis were abstracted either from the completed physician survey or the enrollee's medical records and reviewed independently by 2 Registry research staff, including the Registry's Medical Director (SM-A, JEC). Medical records with unclear diagnoses were reviewed independently by 2 board-certified rheumatologists (PMI, JRB). If the rheumatologists disagreed, they met and discussed the cases until agreement was reached. Agreement was achieved for all records.

Participants' conditions were classified as RA, SLE, scleroderma, SS, MCTD, and PM/DM based on ≥1 of the following 3 criteria: 1) met classification criteria for a specific systemic autoimmune disease; 2) diagnosis determined by a board-certified rheumatologist; and 3) medication prescribed commonly for autoimmune disease.

We only included participants who received their first systemic autoimmune disease diagnosis in 2002 or later. Enrollees for whom we were unable to verify a diagnosis, including those who did not participate in the enrollee survey (n = 744), those who did not provide informed consent (n = 1,001), those for whom a physician survey or medical record was not submitted by November 2017 (n = 289), and those whose survey did not indicate the presence of a systemic autoimmune disease (n = 545) or where the timing or classification criteria were unclear, incomplete, or prior to 2002 (n = 89), were excluded from our analysis.

Comparison group. The comparison study population included all enrollees age ≥18 years who completed the wave 3 survey and who denied having ever been diagnosed with RA or another autoimmune disease.

9/11 exposure. We used 4 metrics of 9/11-related hazards. The first was a measure of intensity of exposure to the cloud of dust and debris on September 11, using responses to the wave 1 survey and the wave 2 survey. Intense dust cloud exposure was defined as having been caught in the dust cloud the morning of 9/11 and also experiencing an additional marker of intense exposure, including the inability to see a few feet ahead, trouble walking or navigating due to dust cloud thickness, taking shelter from the dust cloud, being covered in dust and debris, and the inability to hear anything while in the dust cloud.

The second and third measures were composite dust exposure scores, with one specific to responders and the other specific to community members. These indices are based on those developed by Li et al (35). Briefly, the responder index included the date of arrival to work at the WTC site, duration of work, and exposure to the dust cloud and to the events of the morning of 9/11, and its scoring system was derived using the Delphi method. For this analysis, we defined very high exposure among responders as a score of ≥29.5 (the 75th percentile). The community member index score, dichotomized into categories of "high" and "low" (with "low" defined as low, intermediate, or moderately high), was based on a hierarchy of exposure intensity and duration, including dust cloud exposure and exposures specific to whether the person was a resident (home evacuation and date of return), area worker (date returned to work), or present in Lower Manhattan on 9/11 (for school students/staff members and passersby).

The fourth measure of exposure was the number of months worked at the WTC site. This was calculated based on the work start and end dates provided at wave 1 and was limited to responders.

9/11-related PTSD. PTSD was assessed at wave 1 using the 9/11-specific PTSD checklist and was defined as a score of ≥44 on the checklist. The PTSD checklist is a 17-item questionnaire that asks respondents how much they have been bothered by various symptoms in the past 30 days. The PTSD checklist and, in particular, the use of a cutoff score of 44 have been shown to have high diagnostic efficiency (36). Persons with missing items were included in the appropriate category if their PTSD status could be definitively identified based on their completed items; otherwise, they were categorized as missing.

Covariates. All covariates included in our analysis were self-reported at wave 1, including age, sex, race/ethnicity, educational attainment (high school diploma or less, some college, or college degree and higher), and history of smoking. For our models, we used age at wave 3, which was derived based on date of birth and date of wave 3 completion.

Statistical analysis. This analysis focused on the 6 systemic autoimmune diseases that were specifically queried on the enrollee survey and physician survey. Diagnoses of RA, SLE, PM/DM, scleroderma, SS, and MCTD were aggregated into one outcome, referred to as systemic autoimmune disease. On both surveys, we also included an "other" category, from which we identified additional cases of autoimmune disease. However, these were either organ-specific disorders (e.g., Hashimoto thyroiditis) or conditions for which we were unable to collect sufficient classification criteria for validation.

We conducted a bivariate analysis to compare cases and noncases using chi-square test for independence, Fisher's exact test for sparse cell sizes, t-test, and Wilcoxon's rank sum test, as appropriate. Due to inherent differences in characteristics and exposures between responders and community members, we stratified our bivariate analysis and multivariable log-binomial regression to examine each group separately. In these models, we analyzed the association between 9/11 exposure, PTSD, and the risk of systemic autoimmune disease. Each 9/11 exposure was modeled separately, as was PTSD, and all models were adjusted for sex, age, race/ethnicity, and smoking history. We included educational attainment in our PTSD models to account for potential confounding. Variables were selected for inclusion based on known risk factors for autoimmune disease.

We performed 4 sensitivity analyses. First, we limited our cases to those who met classification criteria, i.e., a more conservative and specific case definition. Second, we excluded diagnoses obtained prior to 2005 to account for possible disease latency effects. We conducted a third analysis examining only RA, the most frequently reported autoimmune disease in this population, as the outcome measure. Finally, we stratified the main analytic exposure models by PTSD status at wave 1. All analyses were performed using SAS Enterprise Guide 7.13.

Table 1. Specific systemic autoimmune disease diagnoses among World Trade Center Health Registry enrollees (n = 118)*

RA	71
SS	22
SLE	20
Myositis (PM or DM)	9
MCTD†	7
Scleroderma	4

* Total number of enrollees with a systemic autoimmune disease will be >118, due to individuals with >1 systemic autoimmune disease diagnosis: rheumatoid arthritis (RA) and Sjögren's syndrome (SS) (n = 4); systemic lupus erythematosus (SLE) and mixed connective tissue disease (MCTD) (n = 3); RA and MCTD (n = 2); RA and SLE (n = 2); SLE and SS (n = 1); RA and myositis (n = 1); RA and myositis, and MCTD (n = 1). PM = polymyositis; DM = dermatomyositis.
† Diagnoses for 6 persons with MCTD, all of whom also had another systemic autoimmune disease, were made based on the opinion of the treating physician, though independent confirmation of classification criteria could not be made.

RESULTS

We identified 118 persons with post-9/11 systemic autoimmune disease in our population, of whom 62 were responders and 56 were community members. Sixty-six cases met disease classification criteria, 43 only met the case criterion of having been diagnosed by a board-certified rheumatologist, and 9 only met the criterion of having been prescribed medication for an autoimmune disease. The most commonly reported diagnosis was RA (n = 71), followed by SS (n = 22), SLE (n = 20), PM/DM (n = 9), MCTD (n = 7), and scleroderma (n = 4) (Table 1). Twelve percent of cases had >1 post-9/11 systemic autoimmune disease. Of enrollees who screened positive on the CSQ, 20% were subsequently verified as having an autoimmune disease, compared with only 4% among those who did not screen positive (*P* < 0.0001).

In both the responder group and the community member group, those who developed autoimmune disease included a higher percentage of females (48% of responders who developed autoimmune disease were female versus 21% of responders who did not, and 86% of community members who developed autoimmune disease were female versus 53% of community members who did not; *P* < 0.0001) and were on average slightly older (age 55 years versus 51 years and age 54 years versus 51 years; each *P* < 0.05) (Table 2). A higher proportion of community members with systemic autoimmune disease had PTSD than those in the

Table 2. Characteristics of the WTC Health Registry study population stratified by systemic autoimmune disease case status*

	Responders		Community members	
	Comparison group (n = 17,284)	Systemic autoimmune disease cases (n = 62)	Comparison group (n = 19,733)	Systemic autoimmune disease cases (n = 56)
Sex				
Female	3,676 (21.3)	30 (48.4)†	10,384 (52.6)	48 (85.7)†
Male	13,608 (78.7)	32 (51.6)†	9,349 (47.4)	8 (14.3)†
Age at wave 3, mean ± SD years	51.3 ± 10.3	55.1 ± 9.5†	51.1 ± 12.6	53.8 ± 10.1†
Race/ethnicity				
Non-Latino white	13,338 (77.2)	45 (72.6)	12,899 (65.4)	32 (57.1)
Non-Latino African American‡	1,133 (6.6)	–	2,410 (12.2)	11 (19.6)
Latino‡	1,896 (11.0)	8 (12.9)	2,134 (10.8)	–
Asian or other‡	917 (5.3)	–	2,290 (11.6)	–
Educational attainment				
High school or less	4,391 (25.6)	21 (33.9)	3,348 (17.1)	13 (23.6)
Some college	5,391 (31.4)	22 (35.5)	3,631 (18.6)	13 (23.6)
College graduate and higher	7,395 (43.1)	19 (30.7)	12,592 (64.3)	29 (52.7)
PTSD				
Yes	1,885 (11.0)	11 (17.7)	2,940 (15.3)	22 (40.0)†
No	15,299 (89.0)	51 (82.3)	16,268 (84.7)	33 (60.0)†
Smoking history				
Ever	7,332 (42.7)	28 (45.2)	8,006 (41.0)	29 (52.7)
Never	9,861 (57.4)	34 (54.8)	11,535 (59.0)	26 (47.3)
9/11-related exposures				
Dust intensity				
Intense	3,208 (19.6)	15 (26.3)	5,991 (33.9)	23 (43.4)
None/some	13,154 (80.4)	42 (73.7)	11,691 (66.1)	30 (56.6)
Dust composite score				
High	4,204 (25.1)	15 (24.6)	1,916 (9.8)	6 (10.7)
Low	12,539 (74.9)	46 (75.4)	17,683 (90.2)	50 (89.3)
Worked at WTC site, median (IQR) months	0.6 (0.2–1.7)	0.8 (0.2–1.8)	–	–

* Except where indicated otherwise, values are the number (%). PTSD = posttraumatic stress disorder; WTC = World Trade Center; IQR = interquartile range.
† *P* < 0.05 versus comparison group.
‡ Data were missing for some subjects/parameters. Values shown are based on the totals with data available.

comparison group (40% versus 15%; *P* < 0.0001). There were no significant differences in race/ethnicity, 9/11-related exposures, or smoking history at the bivariate level.

Our multivariable models revealed that women had a 3–5 times greater risk of systemic autoimmune disease compared with men (Tables 3 and 4). This result was consistent and significant in all models for both responders and community members. Older age, included as a continuous variable, was significant among responders. The adjusted risk ratios (RRs) for Non-Latino African American enrollees and Latino enrollees were consistently elevated compared with Non-Latino white enrollees, and the adjusted RRs for Asians and persons of other race/ethnicities were consistently lower compared with Non-Latino white enrollees, though none of the differences reached statistical significance. Lower levels of educational attainment were associated with an increased risk of systemic autoimmune disease only among responders. Community members with a history of smoking had an elevated, though nonsignificant, risk of systemic autoimmune disease.

Among responders, those who experienced intense dust cloud exposure were at increased risk of systemic autoimmune disease (adjusted RR 1.86 [95% CI 1.02–3.40]). Though the adjusted RRs for dust composite score, PTSD, and for each month worked at the WTC site were elevated, none were statistically significant (Table 3).

Among community members, neither dust cloud intensity nor dust composite score was significantly associated with systemic autoimmune disease (Table 4). However, PTSD was associated with a nearly 3-fold increased risk of systemic autoimmune disease (adjusted RR 2.80 [95% CI 1.60–4.90]).

The results of our sensitivity analyses showed that, when cases were limited to those who met classification criteria (35 responders and 31 community members), there were associations, though not significant, between 9/11 exposures and systemic autoimmune disease among responders. Intense dust cloud exposure and PTSD were both associated with a statistically significant increased risk (>2-fold) for systemic autoimmune disease among community members. When our analysis was restricted to cases diagnosed in 2005 or later, the initial analysis was unchanged, though the magnitude of the association between PTSD and systemic autoimmune disease among community members increased (Table 5).

The results of our analysis examining only RA as an outcome among responders were similar in magnitude to the findings in our analysis of those with the combined outcome, though the association with dust intensity lost significance. Among community members, neither dust exposure variable was significant. However, we observed an elevated association between PTSD and RA (adjusted RR 3.94 [95% CI 1.92–8.07]) (Table 5).

Table 3. Risk of post-9/11 systemic autoimmune disease among responders enrolled in the WTC Health Registry*

	Model 1 adjusted RR (95% CI)	Model 2 adjusted RR (95% CI)	Model 3 adjusted RR (95% CI)	Model 4 adjusted RR (95% CI)
9/11-related exposure				
Dust intensity	–	–	–	–
Intense	1.86 (1.02–3.40)†	–	–	–
None/some	Referent	–	–	–
Dust composite score				
High	–	1.86 (0.98–3.53)	–	–
Low	–	Referent	–	–
Months worked at WTC site (continuous)	–	–	1.10 (0.97–1.24)	–
PTSD				
Yes	–	–	–	1.42 (0.72–2.80)
No	–	–	–	Referent
Demographics				
Age at wave 3 (continuous)	1.03 (1.01–1.06)†	1.04 (1.01–1.06)†	1.04 (1.01–1.06)†	1.03 (1.01–1.06)†
Sex				
Female	3.96 (2.34–6.70)†	4.10 (2.39–7.06)†	3.41 (2.04–5.73)†	3.75 (2.25–6.25)†
Male	Referent	Referent	Referent	Referent
Race/ethnicity				
Non-Latino white	Referent	Referent	Referent	Referent
Non-Latino African American	1.24 (0.49–3.13)	1.70 (0.77–3.79)	1.39 (0.59–3.29)	1.39 (0.62–3.12)
Latino	1.39 (0.65–2.99)	1.37 (0.64–2.93)	1.31 (0.61–2.83)	1.07 (0.49–2.33)
Asian and other	0.70 (0.17–2.91)	0.68 (0.17–2.82)	0.68 (0.17–2.82)	0.70 (0.17–2.90)
Educational attainment				
High school or less	–	–	–	2.21 (1.16–4.21)†
Some college	–	–	–	1.91 (1.03–3.56)†
College graduate and higher	–	–	–	Referent
Smoking history				
Ever	1.01 (0.59–1.70)	1.03 (0.62–1.71)	0.98 (0.59–1.65)	0.98 (0.59–1.63)
Never	Referent	Referent	Referent	Referent

* WTC = World Trade Center; RR = risk ratio; 95% CI = 95% confidence interval; PTSD = posttraumatic stress disorder.
† *P* < 0.05.

Table 4. Risk of post-9/11 systemic autoimmune disease among community members enrolled in the WTC Health Registry*

	Model 1 adjusted RR (95% CI)	Model 2 adjusted RR (95% CI)	Model 3 adjusted RR (95% CI)
9/11-related exposures			
Dust intensity	–	–	–
Intense	1.50 (0.87–2.60)	–	–
None/some	Referent	–	–
Dust composite score			
High	–	1.06 (0.45–2.48)	–
Low	–	Referent	–
PTSD			
Yes	–	–	2.80 (1.60–4.90)†
No	–	–	Referent
Demographics			
Age at wave 3 (continuous)	1.02 (1.00–1.04)	1.02 (1.00–1.04)	1.01 (0.99–1.04)
Sex			
Female	5.04 (2.35–10.79)†	5.18 (2.43–11.05)†	4.69 (2.20–10.04)†
Male	Referent	Referent	Referent
Race/ethnicity			
Non-Latino white	Referent	Referent	Referent
Non-Latino African American	1.18 (0.55–2.50)	1.36 (0.68–2.73)	1.12 (0.55–2.30)
Latino	1.44 (0.66–3.15)	1.52 (0.73–3.19)	1.15 (0.53–2.49)
Asian and other	0.61 (0.19–2.01)	0.57 (0.17–1.87)	0.56 (0.17–1.83)
Educational attainment			
High school or less	–	–	1.23 (0.62–2.45)
Some college	–	–	1.08 (0.55–2.12)
College graduate and higher	–	–	Referent
Smoking history			
Ever	1.47 (0.84–2.56)	1.56 (0.90–2.68)	1.43 (0.83–2.46)
Never	Referent	Referent	Referent

* WTC = World Trade Center; RR = risk ratio; 95% CI = 95% confidence interval; PTSD = posttraumatic stress disorder.
† $P < 0.05$.

After stratifying by PTSD status, we observed that among responders without PTSD, the association between dust composite score and systemic autoimmune disease became statistically significant and increased in magnitude (adjusted RR 2.28 [95% CI 1.13–4.61]) compared with the main results. There were associations, though not significant, between intense dust exposure and systemic autoimmune disease among responders regardless of PTSD status. We also observed effect modification between intense dust cloud exposure and PTSD among community members, with an elevated and significant adjusted RR among those without PTSD and a nonsignificant adjusted RR of <1 among those with PTSD. There was a similar pattern for the dust composite score, though neither association was significant.

DISCUSSION

Intense dust cloud exposure and 9/11-related PTSD were associated with a greater risk of systemic autoimmune disease among Registry enrollees. Among responders, the present analysis found that intense dust cloud exposure was associated with a higher risk of systemic autoimmune disease. Among community members, there was a strong association between PTSD and new-onset systemic autoimmune disease. The responder group and community member group had inherently different demographic compositions and different types of 9/11 exposures (9). In the responder group, 79% were male, and only 47% were male in the community member group. Since autoimmune diseases are generally more prevalent among women (37–39), it is important to identify an elevated risk of autoimmune diseases among mostly male responders.

Our results are similar to 2 FDNY studies of predominantly male firefighters, which showed that increased duration (number of months) and levels of WTC rescue and recovery work were linked to increased odds and incidence of systemic autoimmune disease, respectively (22,23). FDNY researchers found that each additional month worked at the WTC site yielded a 13% increase in the odds of developing a systemic autoimmune disease (23). In our analysis of responders, there was a 10% increased risk of systemic autoimmune disease for each month worked at the WTC site, though this result was not significant. All those who responded to the 9/11 attack were involved in rescue, recovery, and/or clean-up work. Some responders arrived at the site on 9/11 and may have been caught in the cloud of dust and debris, while others arrived days or weeks later and may have been exposed to resuspended 9/11 dust as they sifted through or moved debris (1).

Table 5. Results of sensitivity analyses of risk of systemic autoimmune disease among WTC Health Registry enrollees*

	Cases that met classification criteria, adjusted RR (95% CI)	Cases diagnosed in 2005 or later, adjusted RR (95% CI)	RA only, adjusted RR (95% CI)	With PTSD, adjusted RR (95% CI)	Without PTSD, adjusted RR (95% CI)
Responders					
Dust intensity					
Intense	1.57 (0.66–3.69)	1.85 (0.99–3.45)	1.84 (0.85–3.99)	1.28 (0.37–4.42)	1.88 (0.94–3.75)
None/some	Referent	Referent	Referent	Referent	Referent
Dust composite score					
High	2.03 (0.84-4.93)	1.82 (0.94-3.53)	1.81 (0.79-4.11)	0.60 (0.12-3.07)	2.28 (1.13-4.61)†
Low	Referent	Referent	Referent	Referent	Referent
Months worked at WTC site (continuous)	1.03 (0.86–1.23)	1.10 (0.98–1.25)	0.99 (0.83–1.18)	0.95 (0.73–1.26)	1.13 (0.99–1.29)
PTSD					
Yes	1.84 (0.77–4.38)	1.39 (0.68–2.82)	1.71 (0.76–3.85)	–	–
No	Referent	Referent	Referent	–	–
Community members					
Dust intensity					
Intense	2.12 (1.03–4.35)†	1.45 (0.81–2.61)	1.33 (0.66–2.69)	0.46 (0.18–1.21)	2.40 (1.21–4.75)†
None/some	Referent	Referent	Referent	Referent	Referent
Dust composite score					
High	0.96 (0.29–3.17)	1.00 (0.39–2.53)	0.55 (0.13–2.31)	0.46 (0.06–3.43)	1.51 (0.58–3.93)
Low	Referent	Referent	Referent	Referent	Referent
PTSD					
Yes	2.63 (1.24–5.56)†	3.40 (1.90–6.10)†	3.94 (1.92–8.07)†	–	–
No	Referent	Referent	Referent	–	–

* All models adjusted for age, sex, race/ethnicity, and smoking history. Models with posttraumatic stress disorder (PTSD) were also adjusted for educational attainment. Every 9/11 exposure or characteristic received its own model. WTC = World Trade Center; RR = risk ratio; 95% CI = 95% confidence interval; RA = rheumatoid arthritis.
† *P* = < 0.05.

Responders were not always able to wear appropriate respiratory protection (40), which may have resulted in increased inhalation of fine particulate matter, crystalline silica, asbestos, and organic hydrocarbon solvents (1,9), all of which have a documented association with autoimmune disease (2–8,10–17).

Our analysis used multiple measures of exposure, and mixed results were obtained. Duration of work at the WTC site and an exposure index comprising a dust composite score both had an elevated association, though not statistically significant, with systemic autoimmune disease. However, intense dust cloud exposure was significantly associated with new-onset systemic autoimmune disease among responders. Those who experienced intense dust cloud exposure were also likely exposed to many of the other traumatic events the morning of 9/11, and the definition of intense dust cloud exposure relies on subjective measures that may be associated with trauma or fear (e.g., inability to see or hear, needing to take cover). Given the significant association between PTSD and systemic autoimmune disease identified in this study (possible association between 9/11 exposures and PTSD, and the strong association between dust exposure and systemic autoimmune disease among those without PTSD), the potential mediating or modifying role that PTSD may have had in the development of autoimmune diseases due to 9/11 exposure cannot be discounted. Prior studies conducted among US veterans of the wars in Vietnam, Iraq,

and Afghanistan have identified temporal associations between PTSD and subsequent autoimmune diseases (19,20), as has an analysis conducted among women enrolled in the Nurses' Health Study II (21). Future analyses should examine whether PTSD or other mental health conditions play a mediating role in the association between 9/11 exposure and subsequent autoimmune disease.

Autoimmune disorders result from interactions between the environment, genetics, and the immune system. Inhalation of dust containing crystalline silica, solvents, diesel exhaust, particulate air pollution, and cigarette smoke have been hypothesized to induce systemic autoimmune diseases via inflammatory pathways, dysregulation of the immune response, and increased peptide citrullination (41–46). Environmental risk factors have been associated with many epigenetic DNA methylation changes (47). Evidence from animal studies and other experimental research suggests that reduced methylation is present in immune cell types in persons with SLE and in those with other autoimmune disease, including reduced methylation of the X chromosome, resulting in overexpression of genes specifically among women and increased susceptibility to these diseases (47). Differences in methylation patterns have also been associated with disease severity among persons with SLE (48).

This study has several limitations. Informed consent for medical record release was not obtained from approximately half

of respondents who completed the enrollee survey. Enrollees who provided informed consent to review their medical records were more likely to be male and a 9/11 responder compared with those who did not provide consent (see Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41175/abstract). Given the higher proportion of women who developed autoimmune disease compared with men, a number of cases may have been undetected. Because women in our population were less likely to have experienced high levels of 9/11 exposure, differential availability of data between men and women may have inflated our results. However, after stratifying by sex, the association between 9/11 exposures and autoimmune disease was qualitatively similar in men and women (data not shown). In addition, among responders who were mostly male, there was no difference in the proportion exposed to intense dust among those who provided informed consent and those who did not. This further suggests that selection bias did not greatly impact our results. Enrollees who provided informed consent were more likely to have screened positive for autoimmune disease on the CSQ (73% versus 63%; $P < 0.0001$) (see Supplementary Table 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.41175/abstract). However, given that there were no significant differences in exposure or PTSD among those who provided informed consent and those who did not, this did not likely bias our results.

Another limitation of the study is that due to the different potential 9/11 exposures experienced by responders and community members, we were unable to compare the dust composite scores between the 2 groups. Additionally, while the index for responders relies on a score that was divided into quartiles, the index for community members is based on an ordinal hierarchy of what are likely more intense/chronic dust exposures, which in turn are based on whether a community member was a resident, area worker, passerby, or school student/staff. Therefore, the magnitude of differences between categories are qualitative rather than quantitative. Additionally, the prior occupational exposure history for our responders was not collected, and therefore we were unable to control for potential pre-9/11 exposures to crystalline silica or other components.

Next, due to resource constraints, medical record reviews to confirm the absence of autoimmune disease among our control population were not conducted. However, due to the rarity of autoimmune diseases, it is likely that the number of false negatives in our comparison population is small, and any misclassification among the comparisons would likely have little effect on our findings.

Finally, our study sample included enrollees who completed a wave 3 survey in 2011–2012; therefore, we likely missed autoimmune disease cases among enrollees who did not complete the wave 3 survey, whether due to death, illness, or lack of participation for other reasons. A smaller proportion of wave 3 survey participants had PTSD at baseline compared with those who did not participate (15% versus 19%; $P < 0.0001$). However, it does not appear that persons with PTSD were more likely to overreport or underreport a diagnosis compared with those without PTSD, as there was no significant difference in the likelihood that a self-reported diagnosis would be verified as a case based on PTSD status at wave 1 ($P = 0.51$). Those with PTSD also comprised a similar proportion of both self-reported cases and verified cases (31% and 28%). A prior evaluation of Registry nonresponse bias found that there was no association between wave 3 survey participation status and either chronic health at wave 1 or having been caught in a dust cloud on 9/11 (49).

Numerous studies, including this one, have found that autoimmune diseases are often overreported by study participants. For example, the positive predictive value of self-reported RA is low, ranging between 21% and 34% (50–52). In 2 large population-based studies that investigated risk factors for RA (the Nurse's Health Study and the Iowa Women's Health Study), only 6–7% of self-reported RA diagnoses could be confirmed by medical record review (53,54).

To address the issue of overreporting, only study participants with autoimmune disease verified by a physician or by medical record review were considered. Of the 2,161 enrollees who reported an RA diagnosis after 9/11 on their wave 3 survey, 71 cases (3%) were confirmed. When limited to those who provided permission to have their medical records reviewed (n = 558), our validation rate for enrollees with RA rose to 13%. The reasons for these low rates may include confusion over whether their disease is autoimmune in nature (e.g., mistaking osteoarthritis for RA), having autoimmune-like symptoms but no physician diagnosis, and lack of access to a rheumatologist. The difficulty in diagnosing an autoimmune disease and length of time between symptom onset and diagnosis add to this complexity (55).

In addition to using medically verified diagnoses and classification criteria, strengths of this study included a high response rate to the enrollee survey and the physician survey, a relatively large sample size, and prospective design. Our study identified similar associations, though not significant, between 9/11 exposures and systemic autoimmune disease in enrollees who met classification criteria, which strengthens our findings and highlights the importance of using rigorous diagnostic criteria. The similar findings of our sensitivity analysis limited to cases diagnosed in 2005 or later also helped confirm the main results of this study.

Systemic autoimmune diseases are associated with exposure to 9/11-related dust and debris and PTSD. Given that they are difficult to diagnose, it is not surprising that a pattern of increased risk by level of 9/11 exposure has only emerged in recent years. It also demonstrates the need to monitor the health of populations affected by a disaster over the long term.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

REFERENCES

1. Lioy PJ, Weisel CP, Millette JR, Eisenreich S, Vallero D, Offenberg J, et al. Characterization of the dust/smoke aerosol that settled east of the World Trade Center (WTC) in lower Manhattan after the collapse of the WTC 11 September 2001. Environ Health Perspect 2002;110:703–14.

2. Barbhaiya M, Costenbader KH. Environmental exposures and the development of systemic lupus erythematosus. Curr Opin Rheumatol 2016;28:497–505.

3. Miller FW, Pollard KM, Parks CG, Germolec DR, Leung PS, Selmi C, et al. Criteria for environmentally associated autoimmune diseases. J Autoimmun 2012;39:253–8.

4. Finckh A, Cooper GS, Chibnik LB, Costenbader KH, Watts J, Pankey H, et al. Occupational silica and solvent exposures and risk of systemic lupus erythematosus in urban women. Arthritis Rheum 2006;54:3648–54.

5. Stolt P, Källberg H, Lundberg I, Sjögren B, Klareskog L, Alfredsson L, et al. Silica exposure is associated with increased risk of developing rheumatoid arthritis: results from the Swedish EIRA study. Ann Rheum Dis 2005;64:582–6.

6. Parks CG, Cooper GS, Nylander-French LA, Sanderson WT, Dement JM, Cohen PL, et al. Occupational exposure to crystalline silica and risk of systemic lupus erythematosus: a population-based, case–control study in the southeastern United States. Arthritis Rheum 2002;46:1840–50.

7. Haustein U, Ziegler V, Hermann K, Mehlhorn J, Schmidt C. Silica-induced scleroderma. J Am Acad Dermatol 1990;22:444–8.

8. Ilar A, Klareskog L, Saevarsdottir S, Wiebert P, Askling J, Gustavsson P, et al. Occupational exposure to asbestos and silica and risk of developing rheumatoid arthritis: findings from a Swedish population-based case-control study. RMD Open 2019;5:e000978.

9. Landrigan PJ, Lioy PJ, Thurston G, Berkowitz G, Chen LC, Chillrud SN, et al. Health and environmental consequences of the world trade center disaster. Environ Health Perspect 2004;112:731–9.

10. Pleil JD, Vette AF, Johnson BA, Rappaport SM. Air levels of carcinogenic polycyclic aromatic hydrocarbons after the World Trade Center disaster. Proc Natl Acad Sci U S A 2004;101:11685–8.

11. Barragán-Martínez C, Speck-Hernández CA, Montoya-Ortiz G, Mantilla RD, Anaya JM, Rojas-Villarraga A. Organic solvents as risk factor for autoimmune diseases: a systematic review and meta-analysis. PLoS One 2012;7:e51506.

12. Nietert PJ, Sutherland SE, Silver RM, Pandey JP, Knapp RG, Hoel DG, et al. Is occupational organic solvent exposure a risk factor for scleroderma? Arthritis Rheum 1998;41:1111–8.

13. Diot E, Lesire V, Guilmot J, Metzger M, Pilore R, Rogier S, et al. Systemic sclerosis and occupational risk factors: a case-control study. Occup Environ Med 2002;59:545–9.

14. Bernatsky S, Smargiassi A, Johnson M, Kaplan GG, Barnabe C, Svenson L, et al. Fine particulate air pollution, nitrogen dioxide, and systemic autoimmune rheumatic disease in Calgary, Alberta. Environ Res 2015;140:474–8.

15. Hart JE, Laden F, Puett R, Costenbader KH, Karlson EW. Exposure to traffic pollution and increased risk of rheumatoid arthritis. Environ Health Perspect 2009;117:1065–9.

16. De Roos AJ, Koehoorn M, Tamburic L, Davies HW, Brauer M. Proximity to traffic, ambient air pollution, and community noise in relation to incident rheumatoid arthritis. Environ Health Perspect 2014;122:1075–80.

17. Noonan CW, Pfau JC, Larson TC, Spence MR. Nested case–control study of autoimmune disease in an asbestos-exposed population. Environ Health Perspect 2006;114:1243–7.

18. Farfel M, DiGrande L, Brackbill R, Prann A, Cone J, Friedman S, et al. An overview of 9/11 experiences and respiratory and mental health conditions among World Trade Center Health Registry enrollees. J Urban Health 2008;85:880–909.

19. Boscarino JA, Forsberg C, Goldberg J. A twin study of the association between PTSD symptoms and rheumatoid arthritis. Psychosom Med 2010;72:481–6.

20. O'Donovan A, Cohen BE, Seal KH, Bertenthal D, Margaretten M, Nishimi K, et al. Elevated risk for autoimmune disorders in Iraq and Afghanistan veterans with posttraumatic stress disorder. Biol Psychiatry 2015;77:365–74.

21. Lee YC, Agnew-Blais J, Malspeis S, Keyes K, Costenbader K, Kubzansky LD, et al. Post-traumatic stress disorder and risk of incident rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68:292–8.

22. Webber MP, Moir W, Crowson CS, Cohen HW, Zeig-Owens R, Hall CB, et al. Post-September 11, 2001, incidence of systemic autoimmune diseases in World Trade Center-exposed firefighters and emergency medical service workers. Mayo Clin Proc 2016;91:23–32.

23. Webber MP, Moir W, Zeig-Owens R, Glaser MS, Jaber N, Hall C, et al. Nested case–control study of selected systemic autoimmune diseases in World Trade Center rescue/recovery workers. Arthritis Rheumatol 2015;67:1369–76.

24. Murphy J, Brackbill RM, Thalji L, Dolan M, Pulliam P, Walker DJ. Measuring and maximizing coverage in the World Trade Center Health Registry. Stat Med 2007;26:1688–701.

25. Brackbill R, DiGrande L, Perrin M, Walker D, Wu D, Pulliam P, et al. New York City Department of Health and Mental Hygiene Agency for Toxic Substances and Disease Registry: World Trade Center Health Registry: data file user's manual. New York: RTI International, New York City Department of Health and Mental Hygiene, and the Agency for Toxic Substances and Disease Registry; 2006. URL: https://www1.nyc.gov/assets/911health/downloads/pdf/wtc/wtc-datafile-manual.pdf.

26. Karlson EW, Sanchez-Guerrero J, Wright EA, Lew RA, Daltroy LH, Katz JN, et al. A connective tissue disease screening questionnaire for population studies. Ann Epidemiol 1995;5:297–302.

27. Karlson EW, Costenbader KH, McAlindon TE, Massarotti EM, Fitzgerald LM, Jajoo R, et al. High sensitivity, specificity, and predictive value of the Connective Tissue Disease Screening Questionnaire among urban African-American women. Lupus 2005;14:832–6.

28. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.

29. 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.

30. 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.

31. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). N Engl J Med 1975;292:344–7.

32. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). N Engl J Med 1975;292:403–7.

33. Bennett R. Overlap syndromes. In: Firestein GS, Budd RC, Harris ED Jr, McInnes IB, Ruddy S, Sergent JS, editors. Kelley's textbook of rheumatology. 8th ed. Philadelphia: Saunders; 2008.

34. Vitali C, Bombardieri S, Jonsson R, Moutsopoulos HM, Alexander EL, Carsons SE, et al. Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. Ann Rheum Dis 2002;61:554–8.

35. Li J, Brackbill RM, Liao TS, Qiao B, Cone JE, Farfel MR, et al. Ten-year cancer incidence in rescue/recovery workers and civilians exposed to the September 11, 2001 terrorist attacks on the World Trade Center. Am J Ind Med 2016;59:709–21.

36. Blanchard EB, Jones-Alexander J, Buckley TC, Forneris CA. Psychometric properties of the PTSD checklist (PCL). Behav Res Ther 1996;34:669–73.

37. Izmirly P, Buyon J, Wan I, Belmont HM, Sahl S, Salmon JE, et al. The incidence and prevalence of adult primary Sjögren's syndrome in New York County. Arthritis Care Res (Hoboken) 2019;71:949–60.

38. Izmirly PM, Wan I, Sahl S, Puyon JP, Belmont HM, Salmon JE, et al. The incidence and prevalence of systemic lupus erythematosus in New York County (Manhattan), New York: the Manhattan Lupus Surveillance Program. Arthritis Rheumatol 2017;69:2006–17.

39. Davidson A, Diamond B. General features of autoimmune disease. In: Rose NR, Mackay IR, editors. The autoimmune diseases. 5th ed. San Diego: Elsevier; 2014.

40. Antao VC, Pallos LL, Shim YK, Sapp JH II, Brackbill RM, Cone JE, et al. Respiratory protective equipment, mask use, and respiratory outcomes among World Trade Center rescue and recovery workers. Am J Ind Med 2011;54:897–905.

41. Pollard KM, Christy JM, Cauvi DM, Kono DH. Environmental xenobiotic exposure and autoimmunity. Curr Opin Toxicol 2018;10:15–22.

42. Povey A, Guppy MJ, Wood M, Knight C, Black CM, Silman AJ. Cytochrome P2 polymorphisms and susceptibility to scleroderma following exposure to organic solvents. Arthritis Rheum 2001;44:662–5.

43. Essouma M, Noubiap JJ. Is air pollution a risk factor for rheumatoid arthritis? [review]. J Inflamm (Lond) 2015;12:48.

44. Mastrofrancesco A, Alfè M, Rosato E, Gargiulo V, Beatrice C, Di Blasio G, et al. Proinflammatory effects of diesel exhaust nanoparticles on scleroderma skin cells. J Immunol Res 2014;2014:138751.

45. Yin G, Wang Y, Cen XM, Yang M, Liang Y, Xie QB. Lipid peroxidation-mediated inflammation promotes cell apoptosis through activation of NF-κB pathway in rheumatoid arthritis synovial cells. Mediators Inflamm 2015;2015:460310.

46. Costenbader KH, Karlson EW. Cigarette smoking and autoimmune disease: what can we learn from epidemiology? Lupus 2006;15:737–45.

47. Carnero-Montoro E, Alarcón-Riquelme ME. Epigenome-wide association studies for systemic autoimmune diseases: the road behind and the road ahead. Clin Immunol 2018;196:21–33.

48. Lanata CM, Paranjpe I, Nititham J, Taylor KE, Gianfrancesco M, Paranjpe M, et al. A phenotypic and genomics approach in a multi-ethnic cohort to subtype systemic lupus erythematosus. Nat Commun 2019;10:3902.

49. Yu S, Brackbill RM, Stellman SD, Ghuman S, Farfel M. Evaluation of non-response bias in a cohort study of World Trade Center terrorist attack survivors. BMC Res Notes 2015;8:42.

50. Kvien TK, Glennås A, Knudsrød OG, Smedstad LM. The validity of self-reported diagnosis of rheumatoid arthritis: results from a population survey followed by clinical examinations. J Rheumatol 1996;23:1866–71.

51. Ling S, Fried LP, Garrett E, Hirsch R, Guralanik JM, Hochberg MC. The accuracy of self-report of physician diagnosed rheumatoid arthritis in moderately to severely disabled older women. J Rheumatol 2000;27:1390–4.

52. Star VL, Scott JC, Sherwin R, Lane N, Nevitt MC, Hochberg MC. Validity of self-reported rheumatoid arthritis in elderly women. J Rheumatol 1996;23:1862–5.

53. Karlson E, Mandl LA, Hankinson SE, Grodstein F. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? Results from the Nurses' Health Study. Arthritis Rheum 2004;50:3458–67.

54. Mikuls TR, Cerhan JR, Criswell LA, Merlino L, Mudano AS, Burma M, et al. Coffee, tea, and caffeine consumption and risk of rheumatoid arthritis: results from the Iowa Women's Health Study. Arthritis Rheum 2002;46:83–91.

55. Rose NR, Mackay IR. Autoimmune disease: the consequence of disturbed homeostasis. In: Mackay I, Rose NR, editors. The autoimmune diseases. 5th ed. San Diego: Elsevier; 2014.

LETTERS

DOI 10.1002/art.41227

Cell-bound complement activation products are superior to serum complement C3 and C4 levels to detect complement activation in systemic lupus erythematosus: comment on the article by Aringer et al

To the Editor:

The recently published European League Against Rheumatism/American College of Rheumatology criteria for the classification of systemic lupus erythematosus (SLE) combine clinical criteria with standard laboratory testing (1). One stated advantage of these new criteria is that they perform better than prior criteria in patients with early-onset SLE. Unfortunately, the common "standard" diagnostic laboratory tests included in these criteria—anti-DNA, anti-Sm antibodies, and low serum C3 and C4 levels—have not changed for decades (2,3). Although complement activation is intrinsic to the pathophysiology of SLE, hypocomplementemia is neither a reliable marker for detecting complement activation in SLE nor very sensitive for the diagnosis of lupus (4).

Complement activation in SLE can be detected more accurately by measuring cell-bound complement activation products (CB-CAPs) than by measuring serum complement levels (4). Furthermore, clinical validation studies and a prospective, randomized clinical utility study have demonstrated that CB-CAPs perform better than serum C3 and C4 levels for the diagnosis of early SLE and probable lupus (4–6).

Although classification and diagnostic criteria perform different functions, in the absence of diagnostic criteria the classification criteria often function as a diagnostic surrogate (7,8). Therefore, I believe it is remiss that new classification criteria did not include the option of including complement activation measured by CB-CAPs as an alternative to serum hypocomplementemia.

Dr. Weinstein has received consulting fees and/or honoraria from Exagen (more than $10,000) and owns stock or stock options in Exagen.

Arthur Weinstein, MD, FACP, FRCP, MACR (iD)
Loma Linda University
Loma Linda, CA
Georgetown University
Washington, DC
and Exagen
Vista, CA

1. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71:1400–12.

2. Weinstein A, Bordwell B, Stone B, Tibbetts C, Rothfield NF. Antibodies to native DNA and serum complement (C3) levels: application to diagnosis and classification of systemic lupus erythematosus. Am J Med 1983;74:206–16.

3. Tsokos GC. In the beginning was Sm. J Immunol 2006;176:1295–6.

4. Putterman C, Furie R, Ramsey-Goldman R, Askanase A, Buyon J, Kalunian K, et al. Cell-bound complement activation products in systemic lupus erythematosus: comparison with anti-double-stranded DNA and standard complement measurements. Lupus Sci Med 2014;1:e000056.

5. Wallace DJ, Alexander RV, O'Malley T, Khosroshahi A, Hojjati M, Loupasakis K, et al. Randomized prospective trial to assess the clinical utility of multianalyte assay panel with complement activation products for the diagnosis of SLE. Lupus Sci Med 2019;6:e000349.

6. Ramsey-Goldman R, Alexander RV, Massarotti EM, Wallace DJ, Narain S, Arriens C, et al. Complement activation in patients with probable systemic lupus erythematosus and ability to predict progression to American College of Rheumatology–classified systemic lupus erythematosus. Arthritis Rheumatol 2020;72:78–88.

7. Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res (Hoboken) 2015;67:891–7.

8. Bertsias GK, Pamfil C, Fanouriakis A, Boumpas DT. Diagnostic criteria for systemic lupus erythematosus: has the time come? Nat Rev Rheumatol 2013;9:687–94.

DOI 10.1002/art.41226

Reply

To the Editor:

Dr. Weinstein expresses disappointment that the new European League Against Rheumatism (EULAR)/American College of Rheumatology (ACR) 2019 classification criteria for SLE do not include CB-CAPs.

We find the concept of CB-CAPs interesting in that low complement proteins C3 and C4 can be masked, as they are part of the acute-phase response. We also appreciate that the cited work had prominent involvement of colleagues who also played an active role in the EULAR/ACR SLE classification project (1,2). However, we do not share Dr. Weinstein's disappointment with the EULAR/ACR 2019 classification criteria and believe that it is due, in part, to misunderstanding of the criteria.

Development of new classification criteria takes several years, is unidirectional, and is based on robust evidence. The expert Delphi exercise, which supported the primary choice of candidate criteria, was completed and presented in abstract form

in 2017 and published in 2018 (3). Therefore, 2 of the 3 CB-CAP studies cited by Dr. Weinstein (1,2) could not have been taken into consideration.

Complement fragments (elevated EC4d and BC4d levels) were considered (3) as were many other more novel laboratory parameters, but, as with these other assays, there was insufficient consensus in the expert Delphi exercise. Many laboratory parameters (including CB-CAPs and type I interferon signature) were not selected due to insufficient evidence or unavailability worldwide. This problem is not resolved for CB-CAPs, as of today. Inclusion, in classification criteria approved by EULAR and ACR, of a test that is not broadly available in clinical practice would limit access to scientific studies. This precludes inclusion of potentially interesting tests that have not been proven robust in independent patient cohorts and hence are not fully established for clinical use in different regions of the world.

There is another distinction between diagnosis and classification that appears relevant regarding CB-CAPs. CB-CAP testing was apparently optimized for sensitivity, which at 61% is much higher than the sensitivity of low C4 or C3 levels (23%) (1). This is important in a diagnostic setting, where it is often relevant to exclude an SLE diagnosis. Classification, on the other hand, aims to positively define the disease with high specificity. Therefore, low C3 and C4 complement levels, with a higher specificity (98%) compared to CB-CAPs (86%) (1), are advantageous for classification. Whether the CB-CAP assay becomes available for routine clinical use worldwide and whether its specificity for SLE will be proven to be high remain to be seen, and we would welcome validated demonstrations of improved laboratory test characteristics.

Martin Aringer, MD [iD]
University Medical Center
and Faculty of Medicine Carl Gustav Carus at TU Dresden
Dresden, Germany
Karen H. Costenbader, MD, MPH [iD]
Brigham and Women's Hospital
and Harvard Medical School
Boston, MA
Thomas Dörner, MD
Charité–Universitätsmedizin Berlin
Freie Universität Berlin
Humboldt-Universität zu Berlin
and Berlin Institute of Health
Berlin, Germany
Sindhu R. Johnson, MD, PhD, FRCPC [iD]
Toronto Western Hospital
Mount Sinai Hospital
and University of Toronto
Toronto, Ontario, Canada

1. Ramsey-Goldman R, Alexander RV, Massarotti EM, Wallace DJ, Narain S, Arriens C, et al. Complement activation in patients with probable systemic lupus erythematosus and ability to predict progression to American College of Rheumatology–classified systemic lupus erythematosus. Arthritis Rheumatol 2020;72:78–88.
2. Wallace DJ, Alexander RV, O'Malley T, Khosroshahi A, Hojjati M, Loupasakis K, et al. Randomised prospective trial to assess the clinical utility of multianalyte assay panel with complement activation products for the diagnosis of SLE. Lupus Sci Med 2019;6:e000349.
3. Schmajuk G, Hoyer BF, Aringer M, Johnson SR, Daikh DI, Dörner T. Multicenter Delphi exercise to identify important key items for classifying systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70:1488–94.

Four cases of lupus psychosis: comment on the article by Hanly et al

To the Editor:

We read with great interest the article by Hanly et al discussing lupus psychosis (1). The authors present data from the largest prospective inception cohort study to date (Systemic Lupus International Collaborating Clinics [SLICC] cohort) that demonstrate that psychosis is a relatively rare event, occurring in 1.53% of neuropsychiatric systemic lupus erythematosus (SLE) patients. Of these patients with lupus psychosis, almost one-fourth of psychotic events occurred a year prior to the diagnosis of lupus, and >80% of the psychotic events resolved within 2 years. While this might suggest a benign course of psychosis, each episode of psychosis is a severe, organ-threatening manifestation of SLE and should be treated with immunosuppressants. As emphasized by Hanly et al, when evaluating lupus psychosis, relatively common causes of psychosis should be excluded; such potential causes include: primary psychotic disorder, primary mood disorder with psychotic features, substance-induced psychotic disorder, side effects of medications (e.g., corticosteroids), and encephalitis (e.g., N-methyl-D-aspartate receptor antibodies).

In Table 1 herein, we present 4 cases of lupus psychosis, supplementing a small body of literature describing lupus psychosis as the initial presentation of SLE, sometimes preceding the diagnosis of lupus by many years and thereafter responding to immunosuppressive treatment. After reviewing all of these publications, we have observed some common themes, including the following: 1) most patients with lupus psychosis are initially diagnosed as having major psychiatric disorder (e.g., schizophrenia/schizoaffective disorder, bipolar disorder, or major depressive disorder); 2) initial diagnosis is often followed by eventual unsuccessful treatment with conventional antipsychotics; 3) the diagnosis is confirmed by the finding of lupus-associated antibodies (e.g., antinuclear antibodies [ANAs]) on serologic testing, as well as abnormal neuroimaging results; and 4) symptoms completely or partially respond to steroids and immunosuppressants (2–4). This confirms the phenotype described in the SLICC cohort, in which the lupus psychosis patients responded well to a combination of psychotropic and immunosuppressive agents, emphasizing the essential role of immunosuppression in the treatment of lupus psychosis.

Table 1. Clinical description of the 4 cases of psychosis*

Patient (age/sex)	History	SLICC criteria†	Laboratory/serologic results	CSF	Imaging procedures	Treatment, effect
1 (45/F)	20-year psychiatric decline; initially diagnosed as having schizophrenia; refractory to antipsychotics; overall atypical features of psychosis	Clinical: neurologic psychosis; immunologic: ANA+, anti-DNA+	WBCs 6.99 × 10³/µl, Hgb 11.7 gm/dl (MCV 88.2 gm/dl), platelets 188 × 10³/µl; normal CMP, PT/PTT, urinalysis, ESR, CRP, blood and urine cultures, urine toxicology; C3/C4; ANA+ (titer 1:320; speckled pattern), anti-DNA+, antihistone+, anti-TPO+	Normal opening pressure; cell count 25→50/µl (96% lymphs), RBCs 4→0/µl, protein 44 mg/dl, glucose 55 mg/dl; negative bacterial/fungal cultures; oligoclonal bands, NMDA receptor antibodies	Brain MRI: diffuse volume loss, hyperintense FLAIR in right frontal lobe; FDG-PET: no abnormal radiotracer uptake in right frontal lobe; EEG: normal	CYC (1,000 mg/m²; 6 doses between September 2018 and August 2019), pulse steroids (6 cycles), oral steroids, RTX (2 doses of 1,000 mg); MoCA from 3/30 at baseline to 22/30
2 (19/F)	After starting college, developed psychosis with hallucinations, paranoia, and impulsive sexual behavior	Clinical: chronic CLE, arthritis, neurologic psychosis; immunologic: ANA+, anti-DNA+	WBCs 6.64 × 10³/µl, Hgb 12.9 gm/dl, platelets 401 × 10³/µl; normal CMP, ESR, CRP, C3/C4; urinalysis: negative protein, WBCs 61/hpf, RBCs 2/hpf; ANA (titer 1:640; speckled pattern), anti-DNA, anti-RNP, anti-SSA	Not performed	Brain MRI: no infarct, hemorrhage, mass effect; brain MRA: 2-mm saccular aneurysm from left ICA, no stenosis; CTA of the head and neck: no stenosis or aneurysms; deficits in processing speed, working memory, executive functioning skills	MMF (2 gm/day), oral steroids, pulse steroids deferred due to mania, antipsychotics, HCQ (400 mg/day); resolution of psychotic symptoms and cognitive improvement
3 (5/F)	History of refractory juvenile myoclonic epilepsy, presented with subacute psychosis	Clinical: acute CLE, non-scarring alopecia, neurologic psychosis; immunologic: ANA+, anti-DNA+; other: sicca, RP	WBCs 4.59 × 10³/µl (total lymphocyte count 1.01 × 10³/µl), Hgb 11.4 gm/dl (MCV 85.9 gm/dl), platelets 144 × 10³/µl; normal CMP, urinalysis, ESR, CRP, urine toxicology; C3/C4; ANA (titer 1:1,280; speckled pattern), anti-DNA, anti-SSA, EC4d	Cell count 10→25/µl (84% lymphs.); RBCs 1→11/µl; glucose 27 mg/dl; protein 58 mg/dl; negative bacterial/fungal cultures, NMDA receptor antibodies, paraneoplastic autoantibody panel, 14-3-3 protein	Brain MRI: 1-cm mass-like lesion without abnormal enhancement; video EEG: frequent 3–4-Hz spike wave discharges without epileptic seizures; broad and significant deficits in verbal fluency, attention, memory, executive functioning, fine motor speed	CYC (1,000 mg/m²; 2 doses between April 2019 and June 2019), pulse steroids (2 cycles); improvement of psychosis
4 (67/M)	Developed acute mania	Clinical: chronic cutaneous lupus erythematosus, neurologic psychosis, leukopenia and thrombocytopenia; immunologic: ANA+, anti-Sm+; other: RP with digital ulcerations	WBCs 2.57 ×10³/µl (total lymphocyte count 0.42 × 10³/µl), Hgb 10.5 gm/dl (MCV 85.9 gm/dl), platelets 99 × 10³/µl; normal CMP, PT/PTT, urinalysis, urine toxicology, ESR, CRP, C3/C4; serum paraneoplastic autoantibody panel negative; ANA (titer 1:2,560; speckled pattern), anti-Sm, anti-U1 RNP, anti-SSA	Not performed	Brain MRI: normal-appearing periventricular white matter T2 FLAIR hyperintensity; CTA: diffuse atherosclerosis without stenosis or aneurysms; EEG: normal; deficits in language, psychomotor processing, learning memory, fine motor speed/dexterity	CYC (800–1,000 mg/m²; 6 doses between February 2019 and July 2019), pulse steroids (6 cycles), oral steroids; improvement of psychosis

* SLICC = Systemic Lupus International Collaborating Clinics; CSF = cerebrospinal fluid; ANA = antinuclear antibody; WBCs = white blood cells; Hgb = hemoglobin; MCV = mean cell volume; CMP = complete metabolic panel; PT/PTT = prothrombin time/partial thromboplastin time; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; anti-TPO = anti-thyroid peroxidase; lymphs. = lymphocytes; RBCs = red blood cells; NMDA receptor antibodies = N-methyl-ᴅ-aspartate receptor antibodies; MRI = magnetic resonance imaging; FLAIR = fluid-attenuated inversion recovery; FDG-PET = fluorodeoxyglucose-positron emission tomography; EEG = electroencephalography; CYC = cyclophosphamide; RTX = rituximab; MoCA = Montreal Cognitive Assessment; CLE = cutaneous lupus erythematosus; hpf = high-power field; MRA = magnetic resonance angiography; ICA = internal carotid artery; CTA = computed tomography angiography; MMF = mycophenolate mofetil; HCQ = hydroxychloroquine; RP = Raynaud's phenomenon; EC4d = erythrocyte-bound C4d.

† Neurologic psychosis is defined as delusions, hallucinations, clinical distress, or impairment in social, occupational, or other relevant areas of functioning. Disturbance does not occur exclusively during the course of delirium and is not better accounted for by another mental disorder (e.g., mania). The presence of non-systemic lupus erythematosus (SLE) factors is assessed and not found to be causal (e.g., primary psychotic disorder, substance- or drug-induced psychotic disorder, psychologically mediated reaction to SLE, marked psychosocial stress, or use of corticosteroids).

While serologic testing for ANA positivity is not universally used to screen for lupus, an argument could be made for ANA screening in the evaluation of atypical psychosis. Van Mierlo et al reported a point prevalence of ANA positivity in schizophrenia of 5.9–48.0% (5). In a study by Mantovani et al (6), 3 of 85 patients fulfilling Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition criteria for a first episode of primary psychotic disorder were found to be ANA positive; 2 of the 3 were also positive for anti–ribosomal P and anticardiolipin antibodies and met American College of Rheumatology classification criteria for SLE (7).

The cases presented here illustrate that primary psychosis remains a diagnosis of exclusion, and nothing in the patient's previous history should override the clinical suspicion for a possible autoimmune etiology, especially if atypical psychosis features are present. In patients with atypical or treatment-refractory psychosis, imaging studies, broad assessment of serologic test results and cerebrospinal fluid markers of neuronal autoimmunity, as well as collaboration between rheumatologists and psychiatrists are all critical to establish a diagnosis of lupus psychosis.

Elizabeth Park, MD [ID]
New York–Presbyterian Hospital/
    Columbia University Irving Medical Center
Sander Markx, MD [ID]
New York–Presbyterian Hospital/
    Columbia University Irving Medical Center
        and New York State Psychiatric Institute
Anca Askanase, MD, MPH [ID]
New York–Presbyterian Hospital/
    Columbia University Irving Medical Center
New York, NY

1. Hanly JG, Li Q, Su L, Urowitz MB, Gordon C, Bae SC, et al. Psychosis in systemic lupus erythematosus: results from an international inception cohort study. Arthritis Rheumatol 2019;71:281–9.
2. Montero-Olvera PR, Berebichez-Fridman R, Velázquez-Álvarez L, Ríos-Morales JR, Rodríguez-Guiza MA. Late diagnosis of systemic lupus erythematosus and antiphospholipid syndrome in an older woman with psychosis: a case report and review of the literature. Clin Case Rep 2017;5:1819–25.
3. Perez O, Dave K, Almanzar A, Prodhan T, Concepion L. Primary psychiatric disorder masking the diagnosis of neuropsychiatric lupus in a patient with altered mental status: a case report. Cureus 2017;9:e1793.
4. Alao AO, Chlebowski S, Chung C. Neuropsychiatric systemic lupus erythematosus presenting as bipolar I disorder with catatonic features. Psychosomatics 2009;50:543–7.
5. Van Mierlo HC, de Witte L, Derksen RH, Otten HG, and the GROUP Investigators. The prevalence of antinuclear antibodies in patients with schizophrenia spectrum disorders: results from a large cohort study. NPJ Schizophr 2015;1:15013.
6. Mantovani C, Louzada-Junior P, Nunes EA, de Figueiredo FP, Oliveira GR, Del-Ben CM. Antinuclear antibodies testing as a routine screening for systemic lupus erythematosus in patients presenting first-episode psychosis. Early Interv Psychiatry 2012;6:322–5.
7. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.

Reply

To the Editor:

We would like to thank Dr. Park and colleagues for their interest in our article and for their description of 4 cases of lupus psychosis. Albeit rare in occurrence, psychosis has been reported as an early manifestation of SLE in several studies (1,2). In the SLICC cohort, the earliest psychotic episode occurred 2 months prior to the diagnosis of SLE, so it is incorrect that "almost one-fourth of psychotic events occurred a year prior to the diagnosis of lupus." Prompt recognition of SLE close to the onset of psychosis and the exclusion of other potential causes supports a causal link and a rationale for the introduction of antiinflammatory/immunosuppressive therapies with concurrent antipsychotic drugs. Although speculative, we suggest that this was a factor in the observation that >80% of the psychotic events had resolved by the second annual assessment following onset of psychosis, parallel with an improvement in patient-reported health-related quality of life. In contrast, as illustrated by Park and colleagues, antipsychotic drugs alone were insufficient to restore the patient to full mental health, and a delay in diagnosis may limit the patient's response to lupus therapies.

To our knowledge, screening for systemic autoimmunity is not universally performed by psychiatrists in the initial assessment of patients presenting with new-onset psychosis. The data presented by Park and colleagues support such an approach, at least in patients with atypical psychosis and in those whose disease is refractory to standard antipsychotic drugs. Laboratory screening for ANA is a reasonable and low-cost approach, and a positive result should be followed by a rheumatologic assessment to confirm or rule out a diagnosis of SLE. Although the experienced clinician holds the key role in the diagnosis of neuropsychiatric SLE (NPSLE), advances in neuroimaging, autoantibodies, and insight into pathogenic mechanisms (3,4) hold promise for improved diagnoses and treatments. As suggested by Lockshin (5), state-of-the-art research models may advance the fundamental understanding of neuropsychiatric lupus, whether it presents as psychosis or other manifestations of NPSLE.

John G. Hanly, MD [ID]
on behalf of the Systemic Lupus
    International Collaborating Clinics
Queen Elizabeth II Health Sciences Center
    and Dalhousie University
Halifax, Nova Scotia, Canada

1. Jacobsen S, Petersen J, Ullman S, Junker P, Voss A, Rasmussen JM, et al. A multicentre study of 513 Danish patients with systemic lupus erythematosus. II. Disease mortality and clinical factors of prognostic value. Clin Rheumatol 1998;17:478–84.

2. Pego-Reigosa JM, Isenberg DA. Psychosis due to systemic lupus erythematosus: characteristics and long-term outcome of this rare manifestation of the disease. Rheumatology (Oxford) 2008;47: 1498–502.

3. Hanly JG, Kozora E, Beyea SD, Birnbaum J. Nervous system disease in systemic lupus erythematosus: current status and future directions [review]. Arthritis Rheumatol 2019;71:33–42.

4. Kello N, Anderson E, Diamond B. Cognitive dysfunction in systemic lupus erythematosus: a case for initiating trials [review]. Arthritis Rheumatol 2019;71:1413–25.

5. Lockshin MD. To eat the elephant [editorial]. Arthritis Rheumatol 2019; 71:177–8.

Clinical Images: Facial papular lesions in granulomatosis with polyangiitis
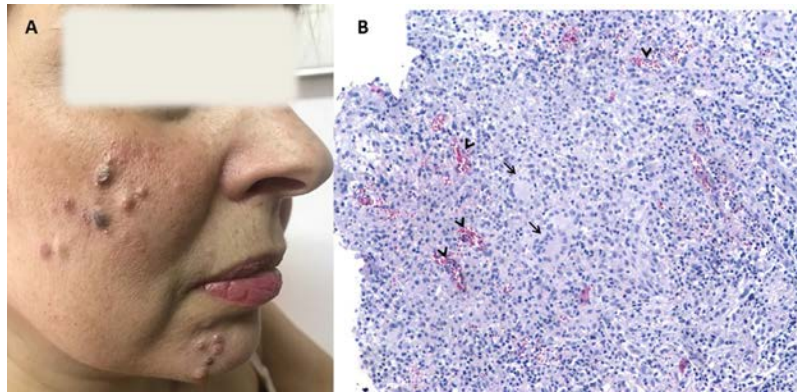
The patient was a 39-year-old woman with a 5-year history of limited pulmonary granulomatosis with polyangiitis (GPA). She was treated with methotrexate (15 mg/week subcutaneously), oral glucocorticoids (0.5 mg/kg/day), and trimethoprim/sulfamethoxazole (160–800 mg/day), with good clinical response. However, after 3 years of receiving the same treatment, she presented with mild anemia (hemoglobin 11.4 gm/dl), elevated serum C-reactive protein level (62 mg/liter [normal <5]), moderate proteinuria (500 mg/24 hours), and high proteinase 3–antineutrophil cytoplasmic antibody (PR3-ANCA) titers. Renal biopsy revealed no abnormalities. High-resolution computed tomography of the chest showed enlargement of the preexisting lung nodules and ground-glass opacities bilaterally. Methotrexate was discontinued; a 4-week course of weekly rituximab (375 mg/m^2 body surface) was initiated and repeated after 6 months. Four months after the second rituximab cycle, the patient presented with a dry cough and red, brown, and black papular lesions on the chin and right cheek (**A**). Titers of cytoplasmic ANCA and PR3-ANCA were high, and lung nodules had recavitated and increased in size. Biopsy of the chin lesions (**B**) showed granulomatous dermal inflammation with multinuclear histiocytic giant cells (**arrows**) without vasculitis, and extravasated red blood cells (**arrowheads**) (hematoxylin and eosin stained; original magnification × 100). Oral cyclophosphamide was initiated. Three months later, the papular lesions had subsided substantially. In immunocompromised individuals presenting with granulomatous cutaneous lesions, mycobacterial or fungal infections should be ruled out. Other conditions seldomly coexisting with GPA, e.g., erythema elevatum diutinum (EED) (1), granuloma annularis (2), and Sweet syndrome (3), should be considered in the differential diagnosis. In EED, papules appear on extensor surfaces, and leukocytoclastic vasculitis is the key diagnostic histologic feature. Granuloma annularis papules appear above/near distal extremity joints and not on the face; histologic analysis reveals collagen fiber degeneration, dermal mucin deposition, and palisaded granulomatous inflammation. Sweet syndrome has been described in only a few patients with GPA; histologic analysis demonstrates neutrophilic dermatosis. GPA causes a wide range of skin lesions in 15–50% of patients. Palpable purpura, due to leukocytoclastic vasculitis, is the most common. Cutaneous manifestations in GPA usually parallel disease activity in other organs and are frequently suggestive of active systemic disease.

1. Kavanagh GM, Colaco CB, Bradfield JW, Archer CB. Erythema elevatum diutinum associated with Wegener's granulomatosis and IgA paraproteinemia. J Am Acad Dermatol 1993;28:846–9.

2. Del Porto F, Proietta M, Muscianese M, Tamburi F, Cifani N, Ferri L, et al. Granuloma annularis revealing Wegener's granulomatosis. Int J Immunopathol Pharmacol 2014;27:273–8.

3. De Boysson H, Martin Silva N, de Moreuil C, Néel A, de Menthon M, Meyer O, et al. Neutrophilic dermatoses in antineutrophil cytoplasmic antibody-associated vasculitis: a French multicenter study of 17 cases and literature review. Medicine (Baltimore) 2016;95:e2957.

Evangelia Zampeli, MD, PhD (iD)
Institute for Autoimmune Systemic and Neurological Diseases
Haralampos M. Moutsopoulos, MD, FACP, FRCPHC, MACR
Academy of Athens
and Institute for Autoimmune Systemic and
Neurological Diseases
Athens, Greece